* *Alternation* is an interdisciplinary journal which publishes interdisciplinary contributions in the fields of Southern African literature and languages.
* Prior to publication, each publication in *Alternation* is refereed by at least two independent peer referees.
* *Alternation* is indexed in The Index to South African Periodicals (ISAP) and reviewed in The African Book Publishing Record (ABPR).
* *Alternation* is published every semester. *Alternation* was accredited in 1996.

EDITOR
Johannes A Smit (UDW)

BOOK REVIEWS EDITOR
Judith Coullie (UDW)

EDITORIAL COMMITTEE
Jaco Alant (UN); Mante Mphahlele (Vista); Shane Moran (UDW); Themba Msimang (Unisa); Priya Narismulu (UDW); Annemarié van Niekerk (Potchefstroom); Johan van Wyk (CSSALL); Jean Philippe Wade (CSSALL).

EDITORIAL BOARD
David Alvarez (Austin); Chukwuma Azuonye (Massachusetts); Richard Bailey (UDW); Neville Choonoo (Oneonta); Ampie Coetzee (UWC); Leon de Kock (Unisa); Marianne de Jong (Unisa); Ian Glenn (UCT); Betty Govinden (UDW); Dorian Haarhoff (Namibia); Sabry Hafez (SOAS); Isabel Hofmeyr (Wits); Dan Izebaye (Ibadan); RK Jain (Jawaharlal Nehru); Philip John (Transkei); Robbie Kriger (NIH, IL); David Lewis-Williams (Wits); Nise Malange (UND); Isaac Mathamba (Unisa); Godfrey Meintjes (Rhodes); Fatima Mendonca (Eduardo Mondlane); Jabulani Mkhize (RAU); Sikhumbuzo Mngadi (RAU); Louis Molamu (Botswana); PT Mtuze (Rhodes); Katwiwa Mule (Pennsylvania); Thengani Ngwenya (UDW); Isidore Okpewho (Binghamton); Andries Oliphant (Unisa); Julie Pridmore (Unisa); Rory Ryan (RAU); Michael Samuel (UDW); Maje Serudu (Unisa); Marilet Sienaert (UCT); Ayub Sheik (Edwin Mellon Postdoctoral Fellow); Graham Stewart (DUT); Liz Thompson (Zululand); Alan Thorold (UN); Cleopas Thosago (UNIN); Marlene van Niekerk (Wits); Helize van Vuuren (UPE); Hildegard van Zweel (Unisa); John van Zyl (Wits).

NATIONAL AND INTERNATIONAL ADVISORY BOARD
Carole Boyce-Davies (Florida Int.); Denis Brutus (Pittsburgh); Simon During (Melbourne); Antony Easthope (Manchester); Vernon February (Leyden); Ulmar Lehmann (Essen); Douglas Killam (Guelph); Andre Lefevere (Austin); Kenneth Lindfors (Austin); Mbulelo Mzamane (Fort Hare); Jeff Opland (Charterhouse); Graham Pechey (Hertfordshire); Eckard Reckwitz (Essen).

COVER DESIGN
AW Kruger

CORRESPONDENCE ADDRESS
The Editor: Alternation, CSSALL, Univ. of Durban-Westville, Priv. Bag X54001, 4000 DURBAN, South Africa; Tel/Fax: 27-(0)31-2044245; e-mail: essall@pixie.udw.ac.za; CSSALL web-site: http://www.udw.ac.za:80/~steward/
Alternation home page: http://www.udw.ac.za:80/~steward/alternat.html

ISSN 1023-1757

Copyright Reserved: *Alternation*

Alternation

International Journal for the Study of Southern African Literature and Languages

Guest Editor
Rembrandt Klopper

2003

CSSALL
Durban

Preface

It has been a pleasure to serve as the special editor of the cognitive science issues of *Alternation*. To celebrate the first decade of Alternation as an interdisciplinary scholarly journal, and due to the high number of excellent contributions received it was possible to make 2003 the year of cognitive science by dedicating both issues of 2003 (Volumes 10.1 and 10.2) to this discipline which is only now beginning to take its rightful place in South African scholarship.

It was also a privilege to select a cover page image for both celebratory issues. For the first issue, I decided on a painting by Willemien Kruger in a typical cognitive engagement – reading – in the same way that scientists, inventors and discoverers of yesteryear were portrayed surrounded by their sextants, telescopes, microscopes and scale models of whatever formed the focus of their intellectual engagements. Doubtlessly, the various forms of 'reading' in our current, globally and virtually challenging environments, form a central challenge to cognitive engagement.

I would also like to express my gratitude to the artist, for allowing us to use the painting.

Rembrandt Klopper

Introduction

Johannes A. Smit

Ben Kruger develops his argument historically, tracing developments in 'the language of thought' or how scholars have attempted to articulate and explain what the relationship is between language and mind, so to speak. He does this in six sections. In the first, he traces the enlightenment ideal of doing away with all presupposition and the development of pure scientific thought and how this understanding evolved into the postulation of the 'language of thought'. This corresponds to the later view that mind and body stand in a relationship similar to the computer program and the physical computer hardware from within which it is used. It is only possible to comprehend what the programme can do by looking at how it functions as program and not by looking at the hardware. In his next sections, he then exposes the limits of 'functionalism', and then treats other scholarly contributions to this evolving discourse, such as the 'phenomenology of thought', 'embodied thought' and 'language as expression'. Throughout he provides insights from significant philosophers who have engaged this issue. In his conclusion he develops his own position.

As historical linguist, Roger Lass puts his finger on a very central nerve in the sciences - the development of the appropriate discourse (or concepts, or 'metaphors' or 'metalanguage' for short) to accompany science as itself develops and progresses. In order to address this issue at general level, he suggests we deal with it under the general rubric of 'metaphor and knowledge' which to some degree has a different significance in the arts/humanities/social sciences nexus than in the different ones in the natural sciences. More particularly, he focuses this question on the role of genetic metaphor—deriving from the biological sciences and how concepts are analogically transferred into other disciplines—in historical linguistics and

the challenges this discipline has faced since its reception. This concerns the use historical linguistics has made of concepts in the other sciences—'hijacking' them. He also gives an example of such 'hijacked metaphor' and then focuses his main argument on developments within historical linguistics as it relates to the question of the 'units' of analysis scholars have recently developed for this discipline, i.e. especially those derived from the notion of the 'meme'. A central point Lass makes is that because of what we could call the developmental (or, rather 'evolutionary') nature of language, we seem to voice for possibility as discipline in the biological sciences. The question of teasing out the 'ontologies' structure of language in this sense, however, remains still outside our grasp.

Appreciating Lakoff and Langacker's contribution to the change away from the narrowly conceived linguistics—developed to match mathematical precision—and semantics—which was restricted by its taxonomising categories. Gary B. Palmer points out that cognitive studies too needs to be enriched, i.e. with perspectives from cultural linguistics. In his study of Shona noun classifiers, he shows how cultural linguistics could be employed in this task, also referring to the significance ethnographic methods have in the process. His rationale is that 'many lexical domains and grammatical constructions link directly or indirectly to significant cultural models, notably including scenarios and polycentric categories'. In order to understand the grammar and lexicon of a language, one should also understand the cultural models and culturally defined imagery which informs it, and as it derives from the embodied experience and encyclopaedic or world knowledge of that culture. This is the domain in which cultural linguistics works and which he employs in his study. He illustrates how 'cognitive processes of complex category formation' and 'category blending' function within culturally specific models to create the 'polycentric categories' that we know as Shona noun classifiers. Palmer further importantly points to the need for similar research on other classifiers in the Shona system, other Bantu languages and the varied classifier systems in other languages. Important is that it includes the significant element of ethnographic research on scenarios and themes that may impact on linguistic categories.

Ronald Miller focuses his research on the 'inner world' that human beings have in distinction to the outer world or world in which we live. In order to come to a better understanding of this reality, he follows Vygotsky in asking questions about the generation of this world, and explains it as it is not our senses but mediated through a 'social other' as 'intentional agent'. Looking at taxonomies of the senses of self, he points out that, in terms of the understanding of 'mediation', the 'inner self' or 'locus of understanding' is just as crucial as the self as a site from which to perceive and act on the world without. An additional issue at stake in mediation concerns the impact of the action(s) of the mediator on the assumptions of the mediatee, variously described by Gadamer and Ortega. Miller also draws analogical explanations from Heidegger and Hanks. Significant in his analysis is his pointing to the role of pre-understanding as part of our being or selfhood in mediated understanding, how understanding incorporates the mediated situation and how new understanding involves negation. In this sense, and to link up with Miller's title, situations are constituted by answers and the nature of pre-understanding is that it posits unstatable questions to the answers. Understanding then dawns when a mediatee accepts new pre-understandings appropriate to the answers provided. To do so in terms of tasks mastered in an answer-and-question process. Understanding takes place as a mastery where the being of the self incorporates the situation – leading to the 'transformation of the *self* that is achieved by the acquisition of new understanding'. In terms of understanding as the engaging of the playing of a game, the *self* is drawn into the game to re-play it—with the result that play or recreation is a re-affirmation of the *self*. Along similar lines, Miller also explains the dynamics involved in the various ways of appreciating sports and art.

In their article, Andrew Dellis and David Spurrett research what they call the *ontology* of the cognitive agents posited by accounts for the mind as dynamically realised by both body and world. They address this focus by first providing an overview of Dennett's argument in his 'Real Patterns' article, and then Kirsh and Maglio's research on the playing of the computer game, Tetris. That cognitive functioning also importantly exploit non-neural resources. They explain the context in which Dennett's argument occurs, and point to the accumulating evidence for distributed cognition against the

background of more traditional expectations. They also provide a more detailed examination of Kush and Maglio, i.e. the case for the thesis that human players exploit a variety of non-neural resources in the course of play—their 'epistemic action'. The category of physical actions with cognitive benefits in question, and therefore the central focus for the remainder of the article, beginning with a discussion of Van Geiger on the dynamical hypothesis in cognitive science. The question whether the nature hypothesis can be accepted, connects up directly with the debate concerning realist and instrumentalist readings of Dennett. They think that they can make a case for the conclusion that the dynamical structures identified by research like that of Kush and Maglio are indeed real patterns in Dennett's sense (as refined and rationalised by Rossi or at least are candidates for being real patterns.

Similar to Beck and Spurrett, Grant Blair and Stephen Cowley's distributed view of language continues the trend which challenges the perspectives from cognitive internalism, the 'computational theory of mind', or the input-output model. As is the case with labelling, language-activity is based on taking part in human life, a diverse, enculturated world, and not on linguaform (written signs) or the simplistic *tongue—parole* binary. Rather, the common concern is with how language-behaviour links brains, bodies and world. This view of language also means that it espouses an active externalist perspective which entangles microcognitive processes with communication, culture and cognition. This view therefore overthrows the input-output model of mind, while also avoiding 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—where labelling began. They argue. It is therefore a theory of 'language which attempts to provide a new vision of human rationality. It focuses on the examination of human mentality with respect to its dynamics, i.e. cognitive capacities as they articulate with utterance activity to function across the inner and outer world boundaries. It is therefore a theory focused on 'active externalism' which stresses the need to scrutinise what happens 'out in the open'. The argument then looks at the parallels between utterance-activity and actions that underpin performance in computer games, the view that social activity—including 'language—has cognitive properties that connect linguaform with microcognitive activity that

permeates the boundaries of experience or world and brain, and how boundary crossing in this sense contributes to microcognitive aspects of action. It also engages the gap between the external view of linguaform and the use of the world as a cognitive resource, the question as to how people act as their talk crosses agent-boundaries, how label use influences dialogue, events, and how stable and decontextualized aspects of language contribute to the re-membering of experience.

In their article, Gary Mastaler and Gavin Bauer succinctly provide a brief discussion of the definition of 'industrial theatre, put forward their proposal for a 'negotiated dramaturgy' in industrial context, which has to be appreciated as a 'process' and problematizes some models for the 'theatre as communication'. These mainly focus on the Pfister and Doppelfeld models. They then raise the question of how the negotiated dramaturgy is to be positioned in the field of 'performance studies'—i.e. the area of study which investigates the meanings theatre creates and how it articulates meanings 'from both the senders' and the recipients' point of view. Apart from the empirical details concerned with the material substance of the performance, the ultimate question relates to their proposal of a negotiated dramaturgy in industrial theatre context. The authors provide an overview by pointing to the fact that many interactions in industrial context do not engage the real challenges difference pose. Rather than assuming 'standard' techniques which uncritically employ the 'print media, videos, mission statements in languages and cultural codes that may not be mutually comprehended by participants', this proposal of the negotiated industrial theatre may be an alternative to effective 'conversation' and interaction. They argue. The outcome of the article is that it proposes that textbook approaches to internal industrial organisational communication are limited and that negotiated industrial theatre pose an alternate route.

Simon Beck's 'Cognition, Persons, Identity', addresses the non-cognitive critique of the psychological view of identity. Deriving from a materialist view, the views of Peter Singer, Peter van Inwagen and Eric Olson all to various degrees argue that physical or bodily realist forms the essential part of one's identity. In their separate views, they believe that the interests of cognitive science are radically distant from the metaphysical view

However, and supporting the 'sophisticated standard view', Beck argues that this is not the case. The standard psychological view is not inconsistent with materialism and thus with serious cognitive science. Rather, it follows the trend started with Locke and updated by Parfit, namely that what makes someone the same person over time is a matter of facts concerning cognition. Identity derives from overlapping chains which are significantly formed by apparent memories, contrasting beliefs, desires, projects, and emotional attachments. The psychological view, he argues, appears to fit our conceptual scheme better than the biological view. This also means that Locke's notion of personal identity as a 'forensic concept' holds, i.e. that it is of interest and importance because of its role in law and morality. (The question, however, is whether 'law and morality' is universal, and if it is at what price?)

Jacek Brzozowski's 'Anthropocentric Chauvinism' points out that there primarily exists an objectivist, subjectivist, and a dispositional view as to the cognitive explanation of the experience of colour. The objectivist view argues that colours are real properties of objects (either microphysical properties or surface spectral reflectances) in the external world). The objectivist view points to the human experience of colours focusing on red-ness, green-ness, blue-ness, etc.), and argues that colour must be a product of our mind as these phenomenological aspects cannot be found in the perceiver-independent world. The third position, that of dispositionalism or the secondary quality view, attempted at marking out a middle ground between phenomenology and externality. This position's point is that colours are secondary qualities in that they consist in a power or disposition to produce a sensory experience in a perceiver, and that this power is grounded in primary qualities. As Brzozowski points out, this is a relational stance as the identification of colour is dependent on the experiences of the perceiver (or at least on the experiences of what is often referred to as the normal perceiver). The counterpoint of this view is that if there cease to be perceivers, then, so too, there would cease to be colour too. This is what, on the evidence of comparative studies of visual systems across species, he labels a species-specific chauvinist position – a label he also reserves for other objectivist positions – and which should therefore be rejected as an explanation of the ontological status of colour. Ultimately, however, he

accommodates the objectivist stance in his version of the primary quality view. He overviews the dispositionalist view showing its dependence on standard perceivers and standard circumstances briefly explicates the functioning of the visual system highlights the chauvinism that underlies the dispositional account, and then argues that chauvinism should be limited in pragmatic use. He finally shows that by identifying colours as disjunctive microphysical properties that surpass the experiences of them, we can avoid similar criticisms of chauvinism within the proposed ontology.

In her 'No Time like the Present: A Cognitive Approach to Time Differentiation in Discourse', Marie Spreyn explicates perspectivization from 'cognitive rhetoric' – where lines human concepts are the bases of image schemas to account for the symbolic nature of human thought. She shows how the mental spaces approach serves as a background organizing mechanism for the analysis and description of time in everyday discourse as well as in fictional narrative texts. Both kinds of discourse, she shows, operate on a complex configuration of hierarchically related mental spaces. This means that as a sentence in a discourse is processed, the configuration of spaces is simultaneously adapted and based on lexical and grammatical triggers in the sentence. She analogically develops the principles mentioned for time in actual language usage to that of the function of time in story. She does so by consecutively dealing with 'image schemas', 'mental spaces', 'Mental Spaces', 'event frames', 'stories in time', 'a theory of ontology and deixis', and time as a deictic category. She explains that 'time' is here viewed as 'a container in which past, present and future are located up to six smaller containers as cognitive event frames, each with its own internal, causal and modal structures'.

The study and developing of theoretical models for the study of the articulation of text or image representation and recall or memory have been one of the primary foci in the research by educators and psychologists for decades. Schreiber and Ward focus their work on this problematic and review the theoretical framework related to Raymond W. Kulhavy's Choice Retention hypothesis (CR) as well as the research of the past two decades based on this hypothesis. They review how maps improve the recall of associated text and also reviewed of a theoretical model that to some

cognitive-psychological processing of map text information. Existing research provide rough information as to how people process and recall information from maps and text within the conjoint retention hypothesis. They also discuss the current and future directions of the research. They found that structural elements in maps, the order of presentation of maps and texts, individual differences, accuracy and task demands all influence the recall of information and articulate with their model. Both the 'form of representation' and the 'constraints imposed by resource limitation of working memory' impact on the level of information acquisition and recall in the cognitive system. The type of map-text studies reviewed in their contribution allows for the testing of how a pictorial display is likely to be represented and what sorts of tasks a person is able to perform using the representation. 'The information that is retained about the space of a map is dependent upon both the properties of the map itself and on the transformation selected by both the map maker and reader'. More research in this field will make further contributions to the understanding of how maps or organized spatial displays can be usefully employed to increase the recall and retention of text material and, most fundamentally, how we humans process pictorial and verbal information.

If we accept that cognitive sciences to various degrees straddle the disciplines of psychology, philosophy, linguistics, artificial intelligence and neuroscience, the question arises as to the significance of 'consciousness' in this complex. In distinction to physicalism which departs from the perspective that the ultimate constituent of the universe consists of purely physical particles, this focus on 'consciousness' studies the qualitative dimensions of the contents in the universe experiences and enjoys. It is these elements which have given rise to 'life' in some way or another and it is 'life' which is characterized by various senses of the mental phenomenon of 'consciousness'. Given that there has been a rising tide of research in this field internationally, it is unfortunate but undeniable literature reveals that no South African has so far made a contribution in this field which certainly calls the question at least in psychology and philosophy. In order to address this gap in research, the contribution by Michael Mark Feltman provides a critical but constructive response to the claims of the late David Brooks in his 2001 article entitled 'How to Solve the Hard Problem: A Predictable

Inexplicability' published posthumously in the e-journal, *Psyche*. Aiming at delineating a possible course for further contributions to consciousness studies from the South African academy, his own approach is to outline a project similar to that of Brooks but expanded, so that it can articulate with current studies in the field of consciousness studies.

David Spurrett concludes the papers in this issue by engaging the question of the reasons for seeing cognition as being distributed. He first reviews perspectives on cognition and cognitive processes as articulated with the physical elements of the human brain and nervous system, and then provides examples and evidence for seeing cognition as distributed. He concludes with 'a sketch of an argument for the *likelihood* of distributed cognition'. 'Distributed cognition' captures the view that 'cognitive processes' extend 'beyond the brain'.

Religion and Culture
University of Durban-Westville

degree explains the processing of map-text information. Existing research provide much information as to how people process and recall information from maps and text within the conjoint retention hypothesis. They also discuss the current and future directions of the research. They found that structural elements in maps, the order of presentation of maps and texts, individual differences, accuracy and task demands all influence the recall of information and articulate with their model. Both the 'form of representation' and the 'constraints imposed by resource limitation of working memory' impact on the level of information acquisition and recall of the cognitive system. The type of map text stimuli reviewed in their contribution allows for the testing of how a pictorial display is likely to be represented and what sorts of tasks a person is able to perform using the representation. 'The information that is retained about the space of a map is dependent upon both the properties of the map itself and on the transformation selected by both the map maker and reader'. More research in this field will make further contributions to the understanding of how maps or organized spatial displays could be usefully employed to increase the recall and retention of text material and, more fundamentally, how we humans process pictorial and verbal information.

. . .

If we accept that cognitive science to various degrees straddle the disciplines of psychology, philosophy, linguistics, artificial intelligence and neuroscience, the question arises as to the significance of 'consciousness' in this complex. The distinction to physicalism which departs from the perspective that the ultimate constituents of the universe comprise of 'purely physical particles', the focus on 'consciousness' studies the qualities we states some of the elements in the universe experience and enjoy. It is these elements which have given rise to 'life' in some way or another and it is 'life' which is characterised by variable senses of the mental phenomenon of 'consciousness'. Given that there has been a rising tide of research in this field internationally, it is unfortunate that available literature reveals that no South African has so far made a contribution in this field which certainly raise the question at least in psychology and philosophy. In order to address this gap in research, the contribution by Michael Mark Pitman provides a critical yet constructive response to the claims of the late David Brooks in his 2000 article entitled 'How to Solve the Hard Problem: A Predictable

Inexplicability' published posthumously in the e-journal, *Psyche*. Aiming at delineating a possible course for further contributions to consciousness studies from the South African academy, his own approach is to outline a project similar to that of Brooks but expanded, so that it can articulate with current studies in the field of consciousness studies.

David Spurrett concludes the papers in this issue, by engaging the question of the reasons for seeing cognition as being distributed. He first reviews perspectives on cognition and cognitive processes as articulated with the physical elements of the human brain and nervous system, and then provides examples and evidence for seeing cognition as distributed. He concludes with 'a sketch of an argument for the *likelihood* of distributed cognition'. 'Distributed cognition' captures the view that 'cognitive processes' extend beyond the brain'.

Religion and Culture
University of Durban Westville

'Why is There Something and not Nothing?'

Johannes A. Smit

> Although he devotes hours of each day to his new discipline, he finds its first premise, as enunciated in the Communications of Boudewijn Postmodernus: 'Human society has created language in order that we may communicate our thoughts, feelings, and intentions to each other'. His own opinion, which he does not air is that the origins of speech lie in song, and the origins of song in the need to fill out with sound the overlarge and rather empty human soul ('Lucie' in Coetzee 1998:54).

1 Introduction

Alternation celebrates its tenth anniversary in 2003/2004. Established in 1994 as the international journal for the study of southern African Literature and Languages under the auspices of CSSALL (Centre for the Study of Southern African Literature and Languages: cf. Smit & Van Wyk 1998/1999), it aimed at cultivating critical perspectives and interactions from within 'literature' and 'language' studies which constructively feed

[1] In this essay, we highlight not background elements—how we saw and present – in terms of which the CSSALL wished to intervene, the Centre's establishment, its articulation with postcolonial studies, its research projects (both theoretical and applied), and its lecture/learning practices. We tried to also bring to the fore the need to research 'the literary' as to the social realities like aesthetic suffers, labours or happiness – which are normally passed over in silence, or are marginalised or suppressed within science, technology and 'development' discourse.

into developing a discourse which would be relevant to current concerns and challenges'.

Alternation focusses its contribution on this complex through its interdisciplinary approach[2] within the larger ambit of the Humanities. It is feeding into the discourse development we posited as goal. This is one of the significant ways through which the journal contributes to our country's transformation challenges. It transcends the limitations, constrictions, and ossified discursive boundaries and exclusionary formations, as well as the mere transplanting of theories and discursive formations generated elsewhere by and for different communities and for aspirations, into southern African academia[3]. It also productively develops ones which creatively and constructively confront our own realities, critically engage those which are alien to or have alienating and marginalising effects on our contexts, and cultivate, encourage and promote those which are central to the diversity of our societal constitution. As I said in my first 'Introduction' to the journal in 1995, we ought to also publish our own locally generated research for local, continental and contingent consumption. With the productive co-operation of more than 250 scholars having their refereed articles published and 82 other contributors, we published twenty substantial

[2] This related to the various articulations of the 'epistemic' binary and as it can be read both against (and forward beyond) the dominant discursive formations we have inherited from the apartheid and colonial eras. This discursive break simultaneously continues to challenge intellectuals to contribute to the empowering of previously disadvantaged, underdeveloped and dislocated mis-educated sectors of society – especially in research capacity development.

[3] Even though we have published multi-disciplinary contributions – as they asked interdisciplinary thought – as editorial board, we have viewed the interdisciplinary field in terms of questions and problematisations which are not normally asked or brought to the fore in the specific subjects or demarcated branches of learning. As such, they may contribute towards the adding to or broadening of the field, the redefining of its boundaries or its displacement or transcending into a new domain.

[4] Much of southern African academia has been known for this (cf. Boje 1997 for example).

mental issues. They significantly contributed towards these goals. They are only the beginning. Another ten years from now, this discourse and this interdisciplinary field will have been even more radically advanced, and its constructive impact on our societies –especially through our research capacity development engagements—much more enhanced.

In this essay, I provide a few pointers on *Alternation*'s contributions to the development of our common discursive commitment(s)[5]. They were interventions which provide perspectives on the trends and strands in our critically constructive discourse development programme[6]. These strands form part of the critical material we are busy developing for the weaving of a shared new (southern) African socially responsive and responsible literary/language and culture discourse, more broadly speaking, which forms part of the foundations of a new community –the 'inter-nation'.[7]

2 Ten Years of *Alternation*

The ten years of *Alternation* can be divided into two complexes: the first four years and the last six.

2.1 *The First Four Years*

The first few years were characterized by the posing of a number of questions concerning 'alternation' in the study of 'southern African literature and languages'. During this period, contributions mostly treated

[5] These 'pointers' come from an earlier draft of a paper I delivered at the 2000 conference of the International Comparative Literature Association, in Provence titled 'Ten Years at Alternation'.

[6] My views, obviously, are not 'final', but rather a personal reflection on my own appreciation of a few of these 'strands' and 'strands' as they emerged through the refereeing processes, the critical-constructive feedbacks, and at the editing screen.

[7] There are many individual contributions which opened new vistas and perspectives onto southern African literature and language studies. However, for lack of space, I hope to incorporate my view in another publication at a later stage.

topics significant for the programmes offered in and to exist to the rationale for the founding of the Centre for the Study of Southern African Literature and Languages (CSSALL) and its researches. They also reflect contributions to the CSSALL's focus on the development of a Southern African Literary History– also drawing on contributions arriving from its first bi-annual conference in 1995.

1994

Some of the items brought to the fore in the first year concerned the challenge towards the development of a more comprehensive 'Southern African Literary Historiography' (Van Vuuren[8]); the articulation of indigenous critical work 'European Theory' especially problematic models (Van Wyk); orality (Marais; Pienaar; Van Vuuren); discursive perspectives on drama (Mngadi[9]) and novel (Hemson); on travel writing (Primmer); on philosophy (Muller); European, especially French writing on Africa (Sienaert); themes 'modern Africans' write about, especially as these provide access to 'experiences beyond [the] limitations of gender, race and class' (Wade); and the theorizing of narrative analysis from within indigenous culture (Cancelo).

1995

The second year saw the continuation of critical dialogue and research on some of the thematic foci already broached, but also the introduction of some new themes. Theoretically, and as it is relevant and significant for our work, the following issues were addressed: Marxist aesthetics (Barnard); deconstruction (Moran); cultural transformation (Mngadi); the role of a national literary history in SA (Wade); social concerns in Afrikaans prose

[8] Please note that, for reference purposes, I provide only the author. The full details are available in each individual published issue. The full indices for the first four years can be found in Alternation 4.2 (1997). The indices for 1998 – 2003 (also containing all the abstracts/summaries of the contributions of the first ten years) are in preparation.

[9] Cf. Wade's (1994) introductory reflections on the first issue.

(Van Wyk); how to address/ analyse or read history as cultural text (De Kock); ideological analysis in linguistics/ literacy programs (Van Zweel); and elements of change in education (Hoekstra; Samuels and Muthukrishna).

The second volume contained pragmatic (ethnographic) approaches to literature (Killam); narrative analysis of modern Arabic literature (Haresh); Xam narrative (Bonner); orality (Turner); anthropological presentation (Thocoldy/ history writing (Duff); canonized literary texts (Meintjes); historical linguistics (Bailey); pastoral (Molamu); 'language maintenance (Mesthrie); and 'relearning' of the self', as transformative strategy in crisis of the dynamics of 'change' in the sub-continent (Govinden).

1996

Its third year was introduced by the dialogue on 'the prescribed' literature in our classes in South Africa (Chalfoun; Cochrane & Gibson; and Hoffmeyr); and followed by 'radical historical consciousness' in South African drama (Mngadi), a Lacanian analysis of Afrikaans drama (Vermeulen); processes of social and cultural transition in a South African drama (Hockemeyer); the representation of South Africa in Mozambican literature (Barbosa); and challenges in African Philosophy (More).

The second issue of the year was introduced by contributions on African Diaspora literature (Boyce Davies) and black feminist reflections on the politics of the representation of women (Abrahams). Of this issue, two more contributions focused on black women writing and identity, i.e. as present in Afro-American women's writing/ critique (Scott) and two book reviews reflecting on 'the black Diaspora' and the question of the invention of 'cultural identity' in African, African-American, and Caribbean drama (Kadalie). The year also contains contributions on feminist perspectives, four reflections on Jinee Ahosa poetry (Mtuze), Afrikaans women writers and the question of 'identity' (Cornelius), and a feminist perspective on the canon of Afrikaans poetry (Loubser). Moran (at Alverez, Meyer, Chetham, and Mostert), respectively reflect on a critical reception/ challenging of Spivak's particular 'alternative with post' discourse, representations of railway culture in indigenous resistance literature, the problematising of the Boer/ Briton articulation in Du Toit's *Die Kommando kan Skeba* (1880-1880), Eugène Marais' *Dwaalstories* (1924), and neo-liberal strands of critique

Van Wyk, Mesthrie and Smit and De Kadt, again, addressed the problematic issues coming from the representation of poor white realities in early twentieth century Afrikaans literature, and the complex/ contradictory relation with the rise of nationalist 'blueprints', the literary and social significance of 'the young adult novel' in South Africa, and the power relation(s) underpinned by the language(s) of minorities.

1997

The highlights of our fourth year were to publish some significant theoretical reflections deriving from different forms of indigenous critique[1], and a refereed selection of the papers from the OSSALT's first bi-annual conference, held in 1995[2].

As for the first focus, we published Strauss's study of 'indigenous African Gestae', Duggan's analyses of Vida's plays, Chetty's critiques of anthropological and feminist writings around the !Kung, Chetwin's story of Krotoa (Eva) in the Van Riebeeck household in the late seventeenth century, Fitzmorris's researches of the 'wives of Henry Fynn' and their lives 'as authors' but also their 'experiences' in the nineteenth century, and Davidson and Filatova's work on Soviet Perspectives on South African society (1917 and 1960s). A focus we may generally label, 'contrastive rhetoric' is also presented. Singh addresses double-sided analyses of what we may call the work(s) of culture, i.e. the issue of 'cultural entrepreneurship' and its counter-point, the 'culturalisation of politics'. Sole, in an award winning article, provides a meticulous analysis of the production and exigencies of literary 'post-discourse', while Moran weighs in with a critical survey of the post-Marxian representations of current South African literary-theoretical studies—i.e. where they combine a re-presentation of Marxian theory with a deconstructionistic or post structuralist re-articulation. He criticizes this approach

[1] I say this in addition to the theorisings which have already taken place in the *Alternations* of the previous three years.

[2] Another refereed selection was published in the volume, *The Dancing Dead from the Land of Spirits* (1999). A selection of the second biannual conference was published as *Body, Identity, Sub-cultures and Representation: Texts from Africa* (1999).

for its oversimplification of the category of class, the fallacy of replacing the category of class with race, and the further marginalization of the already marginalized. This is complemented by his critique of over-simplified articulations with deconstruction in his 'Derrida and the Political' and Smit's critical analyses of Umberto Eco's oeuvre and its studied appreciation of Joyce from his well-known pragmatic background, contributed two articles. One is on the understanding of his linguistics by reformed versions of 'discourse comprehension', and the other on the rationale(s) for 'contrastive rhetoric'.

The second issue of *Alternation* contained a sample of manuscripts from our first biennial conference. Apart from those already published, we have here contributions on a couple of the discursive perceptions necessary for the comprehension of a Southern African literary history. Smith addresses the historical uses of 'Ethiopia' in ancient, classical, medieval and Renaissance Europe, and what happened to this representation after Europe established contact with 'Ethiopia'. Köppe, Rose, Lloyd, Pridmore, Meihuizen and Gagiano, respectively present studies on W.H.I. Bleek, the !Xam perceptions of literature dating from the early British period in Southern Africa, H.F. Lyon (in the context of early eighteenth century travel writing), Thomas Pringle and A.C. Jordan.

Turning to more recent contributions, we have Davidson and Erkkonen reflecting on the reception and representations of 'the South African War' among the Russian public, Quayson on the South African farm novel, Balfour on novels by Schreiner, Gordimer, and Conrad, Clausen on a significant student protest in the 1950s, Pathirakaran on the social stratification in South African Telugu, Bourgault on reports of occult practices in the Liberia Press in the Samuel Doe (1980–1989), Attwell on the inherent reciprocal 'autobiographical' reflexivity present in 'colonial discourse' and 'South African literature as Colonial Discourse', Staebel contributes a study of utopian writers by Rider Haggard, and Besides, on the political articulations of land, literature and history.

2.3 *The Second Six Years*

Continuing the general *Alternation* approach, the second six years focused both the problematizings as well as the discursive addressing of these into a

volume a piece. I am not elaborating on each of these contributions in the same way I did above. I only provide a brief characterization of each of the contributions. The volumes need to be engaged on their own terms.

1998

* Shane Moran—Significant Theory Development in South African *Alternation* 5,1 (1998). This volume contains studies of a sample of Ngugi's writings, the issue of vernacular languages in the post-colony, and various historical and historicising perspectives on a variety of historical texts.

* Johan van Wyk, Lindelwa Mahonga & Johannes A. Smit—Marginal Literatures, Marginal Figures and Marginal Genres in South African Literature: *Alternation* 5,2 (1998). Including manuscripts initially delivered as papers at the second OSSAL biannual conference, this volume provides a wide overview of critical perspectives on a sample of such literatures, figures and genres in the South African context.

1999

* Pitika Ntuli & Johannes A. Smit - The African Renaissance: *Alternation* 6,1 (1999). The question of the African intellectual's articulation with 'power' and 'truth' introduces this volume. It contains critical studies dealing with indigenous sensibilities and occurrences.

* Johannes A. Smit—Significant theoretical approaches to Southern African Literature and Languages. *Alternation* 6,2 (1999). Literary and critical perspectives addressed, range from Thomas Mofolo and Kati Kxoatsin to Breyten Breytenbach, Alan Paton and Njabulo Ndebele.

2000

* Thengani H. Ngwenya & Jabulani Mkhize—Autobiographical Writing/ South African Life-writing: *Alternation* 7,1 (2000). This volume contains a good sample of critical studies on a number of South African autobiographical publications.

* Inna Filatova, Mandy Goedhals, Stephen Leach and Barry White— Development: Perceptions, Perspectives, Identities: *Alternation* 7.2 (2000). Representing a number of critical perspectives in the order of curriculum restructuring activities, this issue of Alternation provides samples of research from the disciplines of Anthropology, Geography, History and Sociology.

2001

* Judith Lutge Coullie – Identity and SA Literature: *Alternation* 8.1 (2001). This volume contains a cross-section of manuscripts which were refereed and accepted for publication at the time and deals with topics such as colonial women's representations of indigenous realities, curriculum in Post Hard's pioneer years, and the relationship(s) to 'home' in Diaspora circumstances.

* Shane Moran Intellectuals and Knowledge/ The Function of Intellectuals: *Alternation* 8,2 (2001). Focusing on the relationship of intellectuals with regard to knowledge produced but also on the knowledge intellectuals themselves produce, this issue questions the presumed silencing of 'critique', the articulation of 'nationalism' and 'feminism', and 'nation' and 'migration' amongst other issues.

2002

* Catherine Addison – Communication Theory/ Intercultural Communication: *Alternation* 9.1 (2002). The volume contains a number of orally delivered as papers at a Communication Theory conference and provides perspectives on the complexities involved in communication, as they derive from cultural orientations.

* Graham Stewart Information Technology and Research Output Evaluation: Comparative: *Alternation* 9.2 (2002). Due to the significant growth in information systems and technology, this is Alternation's first volume dedicated to different aspects important for the development of literature and language studies in terms of these very significant media.

2003

* Rembrandt Klopper– Cognitive Science: *Alternation* 10.1 & 10,2 (2003). Due to the rising significance of 'cognitive science' these issues provide multi- and interdisciplinary contributions to this field, opening possibilities for developments concerning the study of literature and language.

3 Conclusion

I have attempted to provide a brief sample of some of the contributions to *Alternation* over the last ten years. A large variety of issues and literatures were addressed, and theories developed which importantly contribute towards 'discourse development'. What stands out, I believe, is that a significant number of contributions deals with historical and historicizing issues/ approaches. Such approaches provide new ways of engaging literature, but also open new spaces, ways and possibilities of engaging current literary productions and language studies.

Finally, as token of our appreciation for his important contributions to the founding of Alternation, CSSALL as well as his role in the organising of these biannual conferences, we want to dedicate the two issues of the tenth year of *Alternation*, to Johan van Wyk, who was the Director of CSSALL before his untimely illness. Few before him have come as much to encourage and facilitate inter disciplinary studies in literature and languages in southern Africa."

Religion and Culture
University of Durban-Westville

[9] This is also the context for the title of this paper – the apart from its various metaphysical and post-metaphysical interpretations, it signifies creative endeavour.

References

Coetzee, J.M. 1999. *Disgrace*. London: Vintage.

Smit, Johannes A. & Johan van Wyk 1998/1999. Literary Studies in Post-Apartheid South Africa. Guest Editors: Abebe Zegeye and Robert Kriger. *Critical Arts* pp. 107-127. (Special issue: Cultural Change and Development in South Africa.) Also published as: Abebe Zegeye and Robert Kriger (eds): *Culture in the New South Africa after Apartheid*. Volume 2. Cape Town: Kwela Books & History in Line.

Sole, Kelwyn 1997. South Africa Passes the Posts. *Alternation* 4,1: 116-151.

Ware, Jean-Philippe 1994. Introduction. *Alternation* 1,1: 1-7.

Reflections on the Language of Thought

Piet Kruger

1. From Language of Science to Language of Thought

The view of knowledge that is widely accepted in cognitive science assumes that information is input via the senses from the world 'outside', to be used to construct a representation of the world 'inside' the mind. Thought is equated with the ability to reason, using these representations as symbols within a system of logic.

The origin of this way of conceptualizing thought can be found in the seventeenth century Enlightenment. This movement was characterised by a breaking down of the notion of a meaningful order of being, as represented by the Aristotelian world-view that was current in the middle ages. An important part of this project was the development of a language aimed at the aims of science. The goal was to find true knowledge, free from presuppositions—the 'idols' that mislead, to use Bacon's phrase. To this end, language had to be cleared from all those elements that 'imposed' a pre-existing meaningful order.

A language suitable for the project of rational scientific thought is a language that regards the world as a world of facts, a world of objects that can be carefully individuated in terms of their attributes that can be represented by clear and distinct ideas in the mind. By the proper application of such a language, patterns can be found among the ideas, they can be grouped into classes, the relationships that exist among them can be found, to be expressed in the form of natural laws. The language that was developed was to be a language that is appropriate for collecting and controlling ideas: "It is an instrument of control in gaining knowledge of the world as objective process" (Taylor 1985:226). It is in this context that Hobbes saw language as a kind of calculating device, with sentences acting as 'global sums', with a

meaning that can be computed by adding together individual elements. In a similar vein, Leibniz hoped to develop a purified logical language within which all reasoning can be reduced to algorithmic processes.

This fits in well with a tendency that was visible already in classical times, to equate thought with reason, or to define reason as being the proper application of thought. From the ability to use language in a reasonable way (the notion of *logos*) developed the discipline of logic, aimed at providing the laws of reason, which was to describe how (proper) thought proceeds. The unit of reasoning was taken to be the proposition, a statement that makes an assertion about the nature of an object. Objects are known by their attributes, so a proposition may be seen as the linking statement which makes the classification of an object in terms of its attributes possible. If a representation is regarded as a kind of model of the object that classifies that object according to its attributes, then a representation may be regarded as proposition-like in form, an element in an inner logical language. So from an attempt to develop a use of language that is proper to the scientific enterprise comes a suspicion that this language already exists: it is the language of reason, a logical language that underlies all thought. As Taylor (1995:61) puts it,

> [t]hat is, what were seen as the proper procedures of rational thought were read into the very constitution of the mind, made part of its very structure.

This project was advanced considerably in the late nineteenth and early twentieth century with the development of mathematical logic by Boole and Frege, culminating in the *Principia Mathematica* by Russell and Whitehead, first published in 1910. If one accepts that logic as the rules for the proper application of reason is an embodiment of the 'laws of thought' (which, as it happens, is the title of Boole's 1854 book that initiated the revolution in logic and mathematics that was to lead to modern mathematical logic), then mathematical logic may well be considered as a theory about the underlying structure of thought itself. As it was shown by Turing that mathematical logic can be understood as precisely those processes that can be executed by a mechanism that systematically steps through a problem broken down to its most simple constituents (a so-called Turing machine,

which provides a paradigm for all digital computing devices), it becomes clear why such a mechanism came to be regarded as a possible explanatory principle for the process of thought itself.

The notion of mind as a Turing machine was introduced by Putnam (1960) (who was to express serious reservation about it at a later stage, cf. Putnam 1997), and presented in a radical form by Fodor (1975). According to this view, which still informs much of the debate in cognitive science, mind is related to the body as a program is related to the physical computer hardware on which it is implemented. The program cannot be understood by focussing on its physical realization, but only according to the way that it functions in a coherent way as a program (hence the name 'functionalism' for this point of view), but no mystical non-physical cause need to be presupposed to explain its physical effects. Fodor (1975:59) argues that the notion of inner representations that act as symbols when refer to the outside world, inevitably implies some kind of language of thought:

> ... representation presupposes a medium of representation, and there is no symbolization without symbols. In particular, there is no internal representation without an internal language.

This inner language is a formal language, what is referred to as an *interpreted formal system* (Haugeland 1987), which implies a closed set of symbols with formal rules for their manipulation, defined over them in such a way that the 'meaning' of the symbols are established by how they are used by the system. The thesis that the language of thought must be of this kind is defended by Pylyshyn (1980:113) as follows,

> ... computation is the only worked-out view of *process* that is both compatible with a materialist view of how a process is realised and that attributes the behavior of the process to the operation of rules upon representations.

According to Fodor (1975), the only coherent alternatives to functionalism are Behaviourism and neurophysiological reductionism (reducing all behaviour to consequences of the physical activity of the brain), neither of which can explain how the mind can influence behaviour.

Both tend to regard awareness as an irrelevant epiphenomenon, with the annoying attribute that even as such, its existence cannot be adequately explained. The often possible alternative is of course a full blown Cartesian dualism, with mind as a non-physical substance that somehow has physical effects, a view which is rarely regarded as a serious option these days (but see Popper & Eccles 1977).

2. The Limits of Functionalism

In the traditional information processing view of mind, favoured by cognitive scientists, thought is modelled on reason, and reason is modelled on logic, as explained above. An attempt is made to cut away all aspects of awareness, since these are treated as irrelevant contributions of subjectivity. The model for thought that is proposed shows a remarkable resemblance to the model for the process of doing science that was proposed by logical positivism, as a kind of algorithm for scientific research. As Fodor (1985:4) puts it, "... the higher cognitive processes are notably similar to processes of scientific discovery—indeed, that the latter are the former writ large". The irony is that the model for scientific discovery that was proposed by the logical positivists is widely regarded as a failure. Gardner (1985:65) remarks: "Thus, a model that proved inadequate for the scientific enterprise as a whole still motivates research in circumscribed cognitive domains". This seems to be a curious anomaly. If the model cannot give an adequate description of the reasoning processes in the narrow domain of science (as performed by human scientists), how can it be adequate as a model for human thought in general?

Many of the objections that have been raised against the model are consequences of the model itself. Examples of this are the division of the world into an objective domain of facts and a subjective domain of experiences, the notion of the world as a closed finite set of so-called 'natural kinds' which require exhaustive definitions, and the problem that symbols are intrinsically ambiguous and do not uniquely refer, and seem to need a kind of inner interpreter to establish their meaning (often referred to as a *homunculus*). Another consequence is the whole debate about the nature and role of 'consciousness' as the realm of the subjective and how it can influence the behaviour of people, who are presumed to be living in the

world of facts, and how the processing of inner symbols can give rise to a sense of awareness.

A major implication follows from the kind of world view that is implied, a metaphysical construction that Lakoff (1987) refers to as objectivism. This view is in fact implied in much of materialist philosophy, since it is an artifact of the scientific language (as described above). According to this view, the world is made up out of entities, their attributes, and relationships among them. The entities in the world form directly existing categories on the basis of their objectively existing attributes. From this follows the doctrine of 'natural kinds': the world is made up of a closed set of natural kinds, where each kind represents a category which can be defined in terms of shared essential attributes. Some of the attributes are essential, they are what make something the *kind* of thing that it is, so an underlying *essentialism* is also implied. Lakoff (1987) points out that the basic assumption that the world consist of a finite set of determinable natural kinds is hard to support on an empirical basis. Research on categorization such as that of Rosch (1978) seems to imply that natural categories are far less precise and pragmatic than is required by the kind of logical formalism that is presupposed by the language of thought hypothesis.

There are also other problems that follow from the nature of a formal system. It is not at all clear how reference to anything outside the formal system that is supposed to make up the language of thought is to be established in the first place, given the solipsistic nature of such systems. Symbols are only 'meaningful' in terms of how they are used inside the system, and no reference to anything outside it is possible or relevant, a problem that Fodor (1987) struggles with. Formal systems are also complete: no new symbols can be added to the original set and only new symbol combinations can be formed, which leads Fodor to conclude – like Plato long before him – that all knowledge must be 'innate'.

Another difficulty which seems to undermine the functionalist view is that the theory uses the notion of representations in the mind to explain intentionality, the 'aboutness' of consciousness (according to the view of Brentano & Husserl); how 'propositional attitudes' such as beliefs, hopes etc. can influence behaviour. But the processing of inner symbols is unconscious: a subject is unaware of it. In course of a functionalist explanation, the computational level of mind is separated from the phenomenal mind (cf.

Jackendoff 1987; Varela, Thompson & Rosch 1991:52ff) with no necessary link between them. So, ultimately, it is not clear that the computational account is of any use in explaining conscious awareness or the effect of awareness on behaviour—the phenomenon that it originally intended to explain.

The very notion of a private language is called into question in a famous argument of Wittgenstein ([1953]1988), who asserts that the idea of such a private language—including any formal language that acts as a program inside the brain—is inherently incoherent. The terms in a language get their meaning from how they are used, as they are used within a community of language-users. To use a word correctly entails more than connecting a name to an object, it implies that the user of the word knows a great deal about the inner workings of language. This is dependant on societal norms of usage developed over time among people.

To understand what we mean by the expression 'to think,' we should study what Wittgenstein refers to as the *grammar* of this term: the rules or conditions for its proper use. In other words, we should look for those activities that people perform where the application of this expression is appropriate. We should not be mislead by a superficial resemblance of the expression to processes like 'to walk' which implies a certain use of the body, and use this to infer an invisible process going on inside the brain, or worse, inside an imaginary space referred to as 'the mind.'

3. The Phenomenology of Thought

Scientific thought is inclined to distrust phenomenology, since many of the conclusions of scientific reasoning are counter-intuitive and far removed from the direct apprehension of phenomenal experience. The mistake of Aristotelean science which the scientific method was developed to overcome, is precisely to assume that things are as they appear. Scientific reasoning proceeds by developing abstract models specifying relationships among those variables that are regarded as representing essential characteristics of a phenomenon, expressed in the form of measurable quantities. It is this method that has been reified (or ontologized) to become the explanation of the processes underlying reason, by postulating an abstract model inside the brain, acting as a representation of objects in the world. Yet it is also this model that leads to the dilemma that cannot be

resolved: to explain the existence of abstract representations inside an inner space called the mind (somehow instantiated in the brain) which point to some reality outside the mind and which it is presumed to reflect in an accurate way. This dilemma is often stated as the problem of intentionality: that a content of thought has to be *about* something.

The critique of representationalist metaphysics by Heidegger ([1927]1996) starts from the premise that we have forgotten the nature of the phenomenon (i.e., thought or thinking) that we are trying to explain. Heidegger seems to think that in the case of the problem of the nature of thought, this approach, based on certain *a priori* assumptions linked to the Cartesian quest for absolute certainty, has falsified or at least distorted the actual phenomenon under consideration.

Heidegger argues that in order to get a grip on the problem, it is necessary first to establish the necessary conditions that make it possible for anything like experience or awareness to exist, the conditions for forming anything like representations. Wanting to avoid all connotations of subjectivism, he prefers to speak of the clearing (*Lichtung*) within which things are given, rather than use the word experience (cf. Heidegger [1927]1996:133,135). How is the world formed in our personal existential encounter with it? Why are we aware of a world of things? The fundamental problem of Being is that things can appear or come to light at all.

To get perspective on this mystery, Heidegger begins his phenomenological exploration of Being by asking, what is the fundamental or original way that a human being relates to the world. He points out that before we explicitly identify things, we already stand in a relationship with the world. We already have a fundamental orientation to the world, implicit in our existence even before we identify objects as specific kinds of things, as members of categories, as if in language.

This can be seen in how we go about our daily lives, interacting with the world in skilful ways, without explicitly *thinking* about what we are doing. The world as 'ready-to-hand' (*zuhanden*) is already *there* in a meaningful way before reflection. We engage with the world in skilful ways and we are socialized into practices that contain interpretations that cannot be exhaustively contained in the mental states of individuals. Knowing is in the first place *doing*; the objects that we pick out in an act of reflection are grounded in how we deal with things in an involved coping. To individuate

things as neutral objects of contemplation is possible only against a background or a way of being in the world where things already exist as ready to hand, as tools that we (can) use in a skilful way.

We find that underlying our ability to represent there exists a more primitive level of 'absorbed coping'. This 'coping' is not a mental or cognitive element at all, not a feature of states like beliefs or desire (so-called 'propositional attitudes'), but is a dimension of our 'comportment' (to use a term derived from Dreyfus 1991) and skilful dealings with things. This comportment is not neutral—we deal with things as given in one way rather than the other, things are grasped in a certain way relative to a stance or orientation. The coping incorporates understanding, relative to a fundamental involvement with what is given.

Humans inhabit a world, find a home in it, display a 'caring' for it, the world is encountered as something to value or despise, in terms of a mood to which it (the world) is attuned. This absorption in the world is a fundamental feature of our being in the world, it is also social, it occurs in the company of others. The precondition of there being a world is to be in a totality of significance within which things already have a meaning, in relation to each other and relative to one's personal existence within a concrete situation (*Dasein*).

This insight of Heidegger's undermines the representationalist view in a fundamental way, because while it is possible to distinguish a picture of something from the thing depicted, it is not possible to separate our act of dealing with something from the thing that is being dealt with. You cannot deal with something, be involved in it, if it is absent. In this regard Taylor (1995:12) writes,

> The notion that our understanding of the world is grounded in our dealings with it is equivalent to the thesis that this understanding is not ultimately based on representations at all, in the sense of depictions that are separately identifiable from what they are of.

Reference occurs against a background within which a person is already related to the world in a certain way, which is experienced as inherently meaningful. However, this background against which objects show themselves cannot be uncovered as a set of descriptive terms, such

as the 'frames' used in artificial intelligence systems, nor can it be postulated as the effect of unconscious formal processes in the mind, such as, for example, the subcognitive processes referred to by Fodor (1990). Underlying our representations of the world are not further representations, or atoms that can be put together to form representations. Rather, the ability to represent is a product of our relationship with the world, of our existence as active agents in the world; a product not of constituents of representations but of the lived whole which is to be represented.

The error of the metaphysical tradition was to have paid heed only to a derivative mode of intentionality: a subject's representation of objects. It is the prejudice of the scientific viewpoint to believe that one can reduce any process to a clear explicit description that reflects the 'true' state of affairs, which consequently leads to the argument that if technical knowledge does not seem to be of this form, then some such process must be hidden somewhere beneath it. For Heidegger this is not a solution but an imposition of metaphysics (i.e., assumptions about how the world works, based on an age-old philosophical tradition). Heidegger wants us to stick to the facts, the facticity of our personal existence as it is given (*Dasein*)—and rejects what he regards as a contrived explanation in terms of something that is invisible (the mind, the mental realm).

In a sense, his conclusion—that we cannot explain things up in terms of an 'inner realm' that is language-like in form—is similar to that of Wittgenstein. Wittgenstein also stresses that the act of naming, of establishing a relationship between something and the word we use to pick it out, draws on an implicit background (of societal norms and attitudes, a form of life—*Lebensform*) without which it would not make sense, which is a point of agreement with Heidegger. Dasein's world is defined by the related purposes of a certain way of life shared with other people.

4. Embodied Thought

The scientific mind-set is understandably uncomfortable with all this talk about an aconceptual or preconceptual ground out of which meaning emerges. The whole motive behind the development of a scientific language is to make all concepts clear and 'well defined'. For this reason, Heidegger is often accused of an unnecessary obfuscation.

To gain some perspective, it might at this point be worthwhile to reconsider the way that a living organism is related to its environment, and what the likely role of Mentalese is. One of the more central characteristics of cognitivism is a tendency to assume physiology in general and living beings in particular. This is a consequence of the doctrine that thought is rational which proceeds by calculation, and of calculation, it is believed, can be adequately modelled by a formal machine, no physical instantiation being necessary initially.

While classical cognitivism tends to completely ignore the physical nature of the brain, a different point of view has more recently developed, owing its current fortune and the birth of a major source of confusion, to connectionism. Connectionism has been used under attack from [illegible] for which the mind can be modelled by so called neural nets. Fodor has even [illegible] as defined by Dreyfus (1979) shows of [illegible] which is very simplified indeed. According to his view, the information in a cognitive system is distributed [illegible] among single processing units, each interconnected in a parallel network. In such a system, processing consists of spreading activation through the network, weighting the units and strengthening or inhibiting [illegible] of [illegible] connections. According to Fodor's definition of what cognitivism is, [illegible] does not [illegible] and [illegible] is [illegible] and [illegible] does not [illegible] and [illegible] is [illegible] and [illegible] does not [illegible] and [illegible] is [illegible] and

scaffolding in the form of artifacts. For example, using a pencil and paper to solve a mathematical problem (Clark 1998, 68 ff.; and Hutchins' 1995).

This insight has led to a rejection of the view that cognition is the product of an isolated, passive mind processing information that is fundamentally static, but is the product of a living body in interaction with a physical environment. A brain evolved to control the body of an animal that has to move through its ever-changing environment that both sustains it and threatens its continued survival. As Stewart and Gapenne (1999:126) put it,

> The story of brains is inseparably bound to that of nerve cells, because networks of nerve cells provide both the 'hardware' and the 'computational power' needed to run a brain. It is also inseparably bound to the development of senses, because initially the main reason for having nerve cells was to interpret and act upon the signals generated by sense organs. Nerve cells are also implicated in movement, so all these aspects of living creatures—senses, locomotion, brains, and intelligence—are part of a single package.

Cognition, according to this point of view, is a product of experiences that come from having a body with sensory and motor capabilities that are primarily involved with a physical environment via a continuous interactive process.

The influence of this line of thought can be seen in attempts to place it on a more solid scientific footing by linking the notion of neural nets with dynamic systems theory, which studies the way that two or more complex systems jointly evolve over time (Van Gelder 1998; Clark 1999). Brain, body and environment are to be regarded as a mutually specifying evolving system, and the observed activity of the brain is to be interpreted as a reflection of its ongoing interaction with an ever-changing environment. Along these lines, Globus (1995) argues that brain function can be described in connectionist terms, but only if the notion of the process as a form of computation is rejected. Complex neural nets are in a constant state of flux:

> In living nets, everything is continually fluctuating, connection weights, transfer functions, parameters, input, and even the very

> connectivity, without leaving any traces of the fluctuation (Globus 1995:64)

There is a continuous process of self-organization which regulates the structuring of the organism's interactions with the world. *"The outside is not re-presented inside but participates on the inside as part one constituent in a self-organizing process"* (Globus 1995:63; cf. 65).

Describing neural nets in dynamic terms means the process can be discussed in terms of trajectories in a phase space, which are prone to deterministic chaos. The trajectory of states is an attempt to find equilibrium, by moving away from certain states ('repellors') and towards other ones ('attractors'), where the system attempts to settle or 'relax'. This can be used to describe a situation where both the surround and the body are in flux, and the network attunement is also in flux, so that a trajectory evolves out of their mutual participation and settlements.

The idea that living creatures should be studied on their own terms, as systems that struggle to maintain their structural integrity relative to an environment, is presented in a radical form by Maturana and Varela (1980, 1998), who maintain that cognition can be understood as a direct consequence of this, in this they dispense with the notion of information altogether. It is worth describing their view in some depth for the very different perspective it gives on the nature of cognition. Cognition is to be regarded as an implication of life itself, not the activity of a mechanism whose main purpose is to reflect entities and their relations as 'individuals' in terms of existing pregiven properties, as if an exercise in matching something with an inner template.

Maturana and Varela introduce the term *autopoiesis* to describe the self-organisation of a living system that operates within the domain of natural selection. Autopoiesis is most evident in the life of a single cell. Such a living system is to be regarded as autonomous, i.e., all of the operational features of an autopoietic system are active solely within the bounds of the system.

Effects from outside the system are to be regarded as perturbations: 'triggers' which necessitate a continuous reorganisation of the system in an attempt to maintain its structural integrity, but such a reorganisation will always occur in terms of the existing structure of the system at this point in

time. This will bring about a structural coupling between cells and their environment, a coherence between what happens to the cell and the physical environment that it finds itself in, driven (or selected for) by the requirements of natural selection—it will only survive in an environment that enables it to survive. Through this, ecologies of cells can come to work together, and so more complex organisms will develop. This is referred to as a first order (structural) coupling. The kinds of organisms that can develop and survive in a physical environment will be specified by neural drift—how the particular system develops under the constraints of its internal structure and the demands of the environment, under the pressure of evolution.

The interactions of more complex organisms with their environment are to be understood in the same way. It attempts to maintain the organisation of a system relative to environmental events. In mobile animals this can consist of coherent sensory and motor interaction with the world. Nervous systems developed in complex animals to enable coordination of sensory and motor functions among organs (ensembles of cells) separated by distance, as part of the process of maintaining structural integrity.

Among environmental features will be other animals (usually of the same species), and the pressures of evolution often dictate that cooperation with these will be beneficial. This is referred to as a second order structural coupling. To make this possible, forms of communication become necessary. An example is *stigmergy*, the communication by means of chemical traces, used by social insects such as ants. Other animals use sound, expressing their present state and detecting the state of other animals in the group, so to coordinate their behaviour.

The interaction of animals that use signalling systems to bring about a level of mutual influence, thereby bringing about a social system of coherent behaviour that is mutually reinforcing, is therefore intrinsically *lingual* or language-like. It is out of this that human language developed as a culmination but also gives rise to some unique (dare one say, unforeseen) consequences. To gain more perspective on this, it will be necessary to take a closer look at the phenomenon of language.

5. Language as Expression

In the *Logical Investigations*, Wittgenstein ([1953]1958, 2e, paragraph 1) criticises a view of the way that language functions which he attributes to

Augustine. In the passage referred to by Wittgenstein (from the *Confessions* I.VIII.13), Augustine imagines himself to be an infant, learning to use language by observing the way that people use sounds along with their gestures and bodily postures to point to objects in the world. On another occasion Augustine ([427]1958:35) writes, 'Nor is there any other reason for signifying, or for giving signs, except for bringing forth and transferring to another mind the action of the mind in the person who makes the sign' which makes it clear that he regarded language as a signalling system for communicating pre-existing thought from one person to another. Under the influence of the Enlightenment this theory was given a particular nominalist flavour (i.e., words are regarded as nothing but names), and words came to be regarded as arbitrary signs which refer to objects, which can be learned by the systematic association of sounds with objects, and which is used to communicate thought.

This forms the basis of what Taylor (1985:218 ff.) refers to as a *designative* theory of meaning, the notion that the meaning of a word is the thing (or, in more sophisticated form, the factual state of affairs) that it points to. It is the result of the Enlightenment project to develop a language fit for expressing factual statements, the language of science referred to before—or, rather, this view of language and the Enlightenment stress on reason and the intelligibility of all things are different manifestations of the same cultural developments. This in turn leads to the modern tendency to stress the formal (syntactic) aspect of language, where, for instance, syntax is regarded as a branch of recursive function theory, in the tradition of Chomsky, while semantics is to deal with meaning in terms of reference and truth, in the tradition of Frege.

The theory that words are signs that refer to objects purely by way of association (as exemplified by the view of Condillac) was criticised in the eighteenth century by Herder ([1772]1986), on the basis that it presupposes language to explain language. The mystery of language, according to Herder, lies not in the power to designate things by signs, but in understanding how such an ability could exist in the first place.

Herder was a representative of the eighteenth century Romantic movement, a movement that originated out of an explicit rejection of the values and conceptual frame of reference of the Enlightenment. The Romantics rejected the notion of a disengaged reason which contemplates

the harmony of nature from a safe distance, as was implied in the Enlightenment view. Rather, it was argued, one should throw oneself into the stream of life, partake in it, so to make the power of nature manifest in one's creative endeavours. Hence, the stress on literature and art that was so typical of the Romantic period.

A human being is regarded as a being who can give form to the brute forces of nature out of his own existence as a creature of nature. A personal life is the manifestation of an inner power (or will) which struggles to realize its own form in the face of the brute presence of nature but which also finds its own origins in nature. A human life gets its meaning from its power of expression. Expression was not the realization of pre-existing mentalistic forms as was held in classical times. Rather, it was the expression of a self, and could even arise in so far as a self became manifest within the creative activity in a subject actively involved in giving meaning to his or her life.

It is from this Romantic point of view that Herder develops his view of language. According to him, it is necessary to already be in possession of the power of reflection (*Besonnenheit*) to understand how a word can stand for a thing. This power of reflection is however inseparable from language itself; it is prior to language only in a logical sense (it must be presupposed to make language intelligible), but not in a contingent or causal sense. It is not something that was there before the first word was acquired. It is in learning to use words that this ability becomes manifest or is incarnated. Reflection or reflective awareness requires a medium, it exists only within the process within which it becomes manifest. And this medium is language—understood in the broad sense of the symbolic forms of expression, including cultural and artistic activities in general. So the power of reflection can only exist for a being who posses language, and conversely, language is a consequence of the power to reflect, which is, in essence, the expression of the self.

This Romantic vision is linked to the views of Wittgenstein by Lucie (1992) in an explicitly Romantic interpretation. To Wittgenstein, the human spirit expresses itself in the observance of a shared culture. Life begins in doing, in action, and not in reason.

> It is the ability of a group of human beings to adopt shared responses and to develop common judgements (as to what counts as the 'same') which brings about the formation of concepts (Lucie 1992:198).

Language is the medium through which shared norms and communal attitudes become manifest and through which the creative power of a human life comes to expression.

Intellect is not a precondition for the existence of language, thought does not exist before language, contrary to the traditional view as exemplified by Augustine. Rather, to quote Lucie (1992:198) again,

> The adaptation of certain shared responses and common judgements (by means, for example, of linguistic practices) is the emergence of what is recognized as intellect.

Culture is the refining of natural behaviour into socially mediated practices. And it is because these practices exist that the possibility exist both to formulate rules and to follow them, so it cannot be that the rules underlie the behaviour.

What Wittgenstein seems to do is to revive the Romantic tradition of language as a creative act of expression, but without the implicit subjectivism of the Romantic view. Rather than thought—as more manifest in language—being the product of self-expression by a subject, it is the rules and practices of the concrete way that is carried in the 'language games' of that community and that becomes manifest through (and only through) the use of language. There is nothing behind this, no deeper level of analysis, which will lead us to the 'essences' of meaning.

There is however a danger in putting so much stress on language interpreted as a collection of cultural practices that has its locus in the community, that it is sometimes forgotten that it also has a source which is anchored in the personal, concrete existence of a particular person. Some of the more enthusiastic variations of social constructionism and post-modernism seem to err in this direction, giving the impression of disembodied discourse going on without a clear link to life or any physical world at all. Some perspective may be had by turning again to Heidegger. Heidegger develops a notion of thinking that owes much to the Romantic

tradition, but like Wittgenstein he wants to avoid the trap of subjectification. Thought interpreted as 'a creative act of a subject' a product of the will (e.g., the view of Nietzsche).

In our practical involvement with the world there is a 'space' or a 'field of contact' where the world may be said to show itself to us, or where we encounter the world. It is this that Heidegger refers to as the clearing (*Lichtung*). The metaphysical tradition, was originally to ontologize this. This is what Plato does when he finds the clearing, the encounter of human with the world, in the 'place' where the forms are to be found. Plato's attempt to ontologize the field of encounter between human and world, paves the way for the subject-object dichotomy of Descartes and Kant. Heidegger, in contrast to the metaphysical tradition, refers to the clearing as the 'place' where the world comes to presence, where it shows itself to us. The presence of life in the living organism is not a given, but is a result of the organism's ongoing engagement with its environment. This engagement in life is expressed by a particular living creature as the mode of being or being in the world in question, for example the form of life in the weed.

The second aspect involves being a member of a community that have developed ways of dealing with things and situations over a long span of generations, which is then enshrined in norms and practices. In short, being a member of a culture.

The clearing therefore refers to the cultural logic of embodied living and skilful coping-in-the-world, and the medium within which world is present itself to language. The world which is encountered in our concrete involvement with it, is expressed, articulated, given form, by being placed in a frame of reference that has been presented by culture—yet, which also is the process within which culture comes to be. It is here where experience becomes meaningful, can become a target of contemplation to be raised into a world of concepts and objects that are abstracted out of the skills and practices and traditions within which they have their place. This makes it possible to see

things as instances of a kind, so that descriptions of things in terms of attributes become possible—and therefore a designative theory of language.

By conceptualizing language and thought of which it is the expression in this way, some of the insights of the Romantic movement can be retained. Language provides the power to shape things, to alter our perception of things, to enable ever more subtle distinctions that were not possible before. Out of this is formed a world filled with meaningful objects. As Maturana and Varela (1998:209f) put it,

> In the flow of recurrent social interactions, language appears when the operations in a linguistic domain result in coordinations of actions about actions that pertain to the linguistic domain *itself*. As language arises, objects also arise as linguistic instruments of linguistic distinctions that obscure the notions they accord, too

Abstraction becomes possible when we attend to the distinctions that language makes—which imply regarding language as its own object. For example, Lakoff and Johnson (1980) see conceptualization as the consequence of the way metaphors and metonyms can be used to project concrete involvement (embodied experience) into ever more distract realms. Language is not passive it is not just a reflection of the world but plays an active role in shaping it. This in turn influences the stance of humans to the world, which has real effects on their consequent ways of acting. The world comes to be experienced as not just a world of things and events out as a world of values—the importance of how things *ought* to be taken giving shape to practices and norms. Taylor (1995:11) writes,

> We express our emotions, establish our relations, and actualize our values in our body language, style, and rhetoric. But we can also articulate all of that in poetry, novels, dance, music; and we can also bring all of that to descriptive art criticism, where we name the feelings, relations, values, and describe and argue about them.

What is important to keep in mind is that this linguistic encounter with the world is not something we *do*, or which occurs within an inner subjective realm. Heidegger's whole conception of the clearing was

developed to elucidate this point. Thinking is a product both of the way in which a living person encounters a thing and how it comes to be articulated and shaped in language, but it comes to him as an encounter, not as an act. According to the perspective of Heidegger ([1954]1968), what genuine thought requires is to allow the thing thought about to present itself to us (letting it lie before us), to make it appear to us in its way of being in an attitude of humility and thankfulness. We must open ourselves to what is already given (take it to our heart), within its conceptual history (thinking is recollecting). The image is one of the thoughtful encounter as a situation which forces the articulation of an answer. From out of the clearing we are called upon to think.

Words are not a gift associated with meaning but come to us already imbued with meaning. Language is what happens when embodied existence meets a culture. Conversely, culture is carried in language. Language is the application of a culture to embodied experience, the 'mode which directs the human way of existing' (Mertens & Viviers 1998) suggests the word 'linguaculture' as the human variant of the orpbellaan found among social animals). Thinking is the expression of this existence as it is carried in the medium of language, and thinking is authentic where there is a willingness to let the world show itself in its full contextuality.

Reflective awareness occurs about when language becomes the object of its own activity, when the way that language is *about* things is noticed, and with it the separation of expression and what is being expressed. Thereafter it becomes possible to divide the world into objects that are not myself, and a self which contemplates the objects. Pylkkö (1998: 34f) puts this as follows:

> This doesn't mean that the world or being is 'created' or 'constituted' by a subject and its concepts because both subjects and concepts arise from the aconceptual experience as a result of a special kind of articulation process the end product of which, namely the so called subject, 'reflects' the objective part of the experience. The objective part of experience is called 'reality' or the 'world of things'.

6. Conclusion

The information processing account of cognition which is popular in cognitivism displays many of the outer features of a science. In theory it exhibits the basic stance of scientific reasoning that the road to truth is via the application of disengaged reason, focussing on factual information with all residuals of value or 'relevance' removed, so to limit the contaminating effects of prejudice and affect. Furthermore, it displays the common trait in popular scientific accounts of phenomena that the system to be studied should be broken down to its basic elements, and the structure or relationships among these should be established. The latter requirement leads to stress being placed on mathematical descriptions, as is typical of scientific discourse, and which plays such a large role in formal accounts of cognitive processes.

There is however another feature of the scientific approach which seems to be absent. This is the requirement that a scientific account should be grounded in empirical research. The explanation given should be true to the phenomenon that it purports to explain. Here the merits of what the cognitivism model makes feasible, and the actual way that people learn to deal with their day-to-day lives in the world is tenuous. To say the least. There are also more logical problems, referred to before, such as the split between subject and object, the existence of and role of reflective thinking, etc.; that are resolved in ways that remain unsatisfactory. While the classical model may start with a naturalistic impulse – to explain naturally occurring cognitive processes – it ends up with a notion of 'mind' which continues to defy explanation in spite of an endless array of books and papers purporting to explain consciousness.

All of these problems are a consequence of accepting a model of thought as reason, which was developed over centuries of philosophical reflection aimed at solving quite another problem. The traditional quest of epistemology was to establish certainty, to find absolute truth. In this process, the discovery that words name kinds of things led Plato to the theory of forms or ideas. These ideas were subsequently somehow translated to 'inside' the human soul by Descartes, a process within which the link of these forms with the ultimate essences of things (the source of their status as truth) became increasingly tenuous. The link was ultimately severed in the work of Kant, while the notion of formal processing of

sensory input was established. This in turn lead to the soul being purified of all transcendent and religious baggage to be redefined as the modern conception of the mind. It is at this point that the modern notion of the mind as a mechanism for reflecting objects and their relationships came to be established. While Kant understood only too well that a ground beyond the contents given in reflection is needed to explain conceptual thinking, this insight was in time lost or ignored. In this process, the formal model of a procedure for determining absolute truth came to be ontologized, to be presented as a description of the actual physical process that underlies all thinking.

In all this, little consideration was given to how actual living people deal with their physical existence, or how the structures of 'natural' language are related to this. This is why Heidegger accuses the classical model of being the imposition of metaphysics on the phenomenon of thought, or why Lakoff (1987) sees it as a product of *a priori* reasoning – what he calls philosophical speculation – with no basis in research. What is needed is a phenomenological analysis, such as that conducted by Heidegger, and also in his own fashion by Wittgenstein, to clarify how the phenomena of language and thought (or 'cognition' in the basic sense of 'coming to know') present themselves, and how they are linked. Then the scientific work of explaining how such phenomena could come to exist within the natural order of things becomes possible, based on empirically observable features of the phenomenon.

What Heidegger and Wittgenstein and others working in the respective traditions which they established show, is what kind of phenomena it is that needs to be considered in an attempt to explain cognitive capabilities—how the phenomenon of coming to know presents itself. They show that the discourse typical in cognitive science (if it is to be regarded as an attempt to deal with animal or human cognition) does not cut deep enough. An outer aspect of thinking—the existence of words and concepts and the relationships among them—is presented as the very essence of cognition. Yet what is ignored or simply not perceived is that symbols can only be meaningful against a background of possible interpretation within which they get their form.

Consideration of this contextual frame of reference leads to the insight that thought exists at the interface of embodied existence and culture,

and needs the medium of language for its expression. You do not think with your brain as is the commonly-heard notion, rather thought is a product of a brain in a body in an environment inhabited by social beings who work together in forming a common culture, which is a way of life, a way of dealing with what is given in embodied existence.

Embodiment supplies the personal aspect of thought but also guarantees a relationship with reality because it is in the first place concerned with physical existence, with what needs to be done in interacting with what is encountered in the activities necessitated by life. It is the product of a nervous system, coordinating sensory and motor systems, in a body that had to develop ways of subsisting and surviving, in a changing and only partly predictable physical environment over evolutionary time.

Culture supplies the transpersonal aspect of thought – a frame of reference not developed by an individual, but in the history of a community over many generations, into which a person is inculcated since birth. It is the frame of reference within which a personal embodied encounter with the world comes to expression or articulation, in the form of language. This brings about also the dimension of the incompleteness of knowledge. Its fact that it is relative, always a social construction indicating a point of view. The objective and relative aspects of knowledge, however, represents a duality that can only be separated *post hoc* within reflection. In its source it is one.

The cognitive model presented by Maturana and Varela (1980; 1998) as described above, represents an attempt to place these notions on an empirical footing – to 'naturalize' it, to use a term of Pylkkö's (1998). What they attempt to show is that nothing is involved that comes from anywhere beyond the natural world. 'Cognition' is a process that flows logically from life itself.

A living system is a self-organizing process surviving in a physical environment which both sustains it and threatens it, reaching ever higher (or more complex) levels of organization and wider system-in-relationships with other living creatures. Nervous systems develop along with other relevant systems in the body to coordinate the activity of a mobile creature which has to react to its environment to ensure its survival. These developments are guided by the pressure of structural drift (the way structure changes over generations) brought about by evolution. Consequently, the nervous system will develop to enable behaviour that is congruent with the requirements of

this particular animal' in its species-specific history with this particular environment. It is this which guarantees a link to what we call physical reality, since inappropriate behaviour will entail survival.

Nothing of this, however, requires anything like specific inner representations of objects or their attributes in a world outside that can serve as symbolic descriptions of entities or classes of entities. Such descriptions only become possible due to interactions with the environment along with the learning dimension, might also be by a shared culture; they cannot just be *there* at once as pre-existing facts to be waiting for them to originate. If the brain is to be thought of as a computer at all, it should be required to do analog computation. Due to the development of coordination of sensory and motor functions during early infancy, like establishing some of the processes necessary for life, it can not be an analogy of another outside (like Descartes was, relative to the relations of this particular animal) captured within its physical structure. Moreover, this structure should be understood as it exists of and for its own sake, not as the representation of something else.

Out of the cooperation among social animals, which can display logical language-like characteristics, language developed among humans with a certain level of complexity, enabled by a specialised nervous system, this cooperation developed into ways of living, into a culture. Language as expression is the acting out of the state of one's way of being (the relationship with one's environment and how it affects one) at a given moment, placing it into a frame of reference that was communally developed within the history of a culture. So language is not the communication of pre-existing thought, but is itself the form in which thought becomes manifest.

Department of Psychology
School of Applied Social Science
Faculty of Humanities and Social Sciences
University of South Africa

References

Augustine [427] 1958. *On Christian Doctrine*. Robinson, DW jr. (trans). Indianapolis: Bobbs-Merrill Publishing Company.

Augustine [400] 1991. *Confessions*. Chadwick, H (trans). New York: Oxford University Press.

Bechtel W 1997. Embodied Connectionism. In Johnson, DM & C Erneling (eds): *The Future of the Cognitive Revolution*. New York: Oxford University Press.

Clark, A 1999. *Being There. Putting Brain, Body, and World Together Again*. Cambridge, Mass.: MIT Press.

Dreyfus, HL 1991. *Being-in-the-world: A Commentary on Heidegger's* Being and Time. *Division I*. Cambridge, Mass.: MIT press.

Fodor, JA 1975. *The Language of Thought*. New York: Thomas Y Crowell.

Fodor, JA 1985. Précis of 'The Modularity of Mind'. *The Behavioral and Brain Sciences* 8:1 - 42.

Fodor, JA 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, Mass.: MIT Press.

Fodor, JA & ZW Pylyshyn 1988. Connectionism and Cognitive Architecture: A Critical Analysis. *Cognition* 28:1 - 71.

French, RM 1990. Subcognition and the Limits of the Turing Test. *Mind* 99: 53 - 65.

Gardner, H 1985. *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.

Globus, G 1995. *The Postmodern Brain*. Amsterdam: John Benjamins Publishing Company.

Haugeland, J 1987. *Artificial Intelligence: The Very Idea*. Cambridge, Mass.: Bradford.

Hebb, DO 1949. *The Organization of Behaviour*. New York: Wiley.

Heidegger, M [1927] 1996. *Being and Time*. Stambaugh, J (trans). Albany: State University of New York Press.

Heidegger, M [1954] 1968. *What is Called Thinking?* Gray J Glenn (trans). New York: Harper & Row.

Herder, JG [1772, 1986. Essay on the Origin of Language. In Rousseau, J-J & JG Herder. *Two Essays on the Origin of Language*. Moran, JH & A Gode (trans). Chicago: University of Chicago Press.

Jackendoff, R 1987. *Consciousness and the Computational Mind*. Cambridge, Mass.: MIT Press.

Lakoff, G 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.

Lakoff, G. 1987. Philosophical Speculation and Cognitive Science. *Philosophical Psychology* 2,1:55 - 76.

Lakoff, G. & M Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Lorie, Y. 1992. Culture as a Human Form of Life: A Romantic Reading of Wittgenstein. *International philosophical Quarterly* XXXII, 2:193 - 208.

Maturana, H. R. 1978. *Biology of Cognition. A Phenomenological programme of work*. Westport, Connecticut: Greenwood Press.

Maturana, H. R. & F. Varela. 1980. *Autopoiesis and Cognition. The Realization of the Living*. Dordrecht: D. Reidel Publishing Company.

Quinton, A. M. et al. writing 1965. *The Idea of Knowledge: The Stirling Lectures*. Cambridge: Cambridge University Press, second edition. Boston: Shambala.

Popper, K. R. & J. C. Eccles. 1977. *The self and its Brain*. Berlin: Springer International.

Putnam, H. 1960. Minds and Machines. In S. Hook (ed.) *Dimensions of Mind: A Symposium*. New York: Collier Books.

Putnam, H. 1977. Cognitive Issues: Cognitive Science or Science Fiction? In [illegible] (eds) *The Future of the Cognitive Revolution*. New York: Oxford University Press.

Reichle, F. 1968. *The Incorporated Mind. Heidegerrian Theories in Holistic Narratives*. Amsterdam: John Benjamins Publishing Company.

Pylyshyn, Z.W. 1980. Computation and Cognition: Issues in the Foundations of Cognitive Science. *The Behavioral and Brain Sciences* 3:111 - 169.

Rosch, E. 1978. Principles of Categorization. In Rosch, E. & B.B. Lloyd (eds). *Cognition and Categorization*. Hillsdale, N.J.: Lawrence Erlbaum.

Taylor, C. 1985. *Human Agency and Language*. Philosophical papers, Vol. 1. Cambridge: Cambridge University Press.

Taylor, C. 1995. *Philosophical Arguments*. Cambridge, Mass: Harvard University Press.

Van Gelder, T. 1998. The Dynamical Hypothesis in Cognitive Science. *Behavioral and Brain Sciences* 21:615 - 665.

Varela, F.J. E. Thompson & E. Rosch. 1999. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, Mass.: MIT Press.

Wittgenstein, L. [1953] 1986. *Philosophical Investigations*. Anscombe, GEM (trans). Oxford: Basil Blackwell.

Genetic Metaphor in Historical Linguistics

Roger Lass

1. Metaphor and Knowledge

We talk in metaphor, usually without knowing it. We grasp ideas, see the point of an argument, our past is behind us, our future ahead. In scientific disciplines too our speech is loaded with metaphor. Sometimes it is obvious and designed: evolutionary arms races, selfish genes, the Red Queen scenario, hawks and doves. Often it is not so much designed as virtually natural, growing out of the way we perceive the processes or objects we describe. Metaphoric usage may be nearly overdrawn: macrophages roam through the body looking for foreign objects to devour; other immune cells take in antigens, snip out peptides, present them on their surface so that others can recognize them, cells are recruited for immune defense, neoplasms are detected by immune surveillance, DNA strands are unzipped.

This is merely obvious, though it is worth noting if only as a reminder of how saturated our most objective disciplines can be with the poetry that underpins ordinary human speech and thought. Many people mistakenly think that metaphors are 'mere', not to be taken seriously, just 'figures of speech'. On the contrary, metaphor is often a potent, field-defining device: a good metaphor can open new mental spaces, suggest heuristics, and even do really useful (virtually) empirical work. So de Preez (1991:69) says: 'a metaphor is a conjecture about the nature of reality: a significant metaphor is an insight into unexpected properties of the object of an inquiry'. Metaphors delimit conceptual spaces, particularly so that a higher-order metaphor generates what has been called an 'M-language' (Paton 1992): the metaphor itself calls up natural associations, which can be used to flesh out its applications. Consider for instance the evolutionary

arms race (parasite/host or predator/prey). Once the M-domain 'Military Competition' is invoked, associations arise, like outflanking the enemy, going round its defences, recruiting soldiers to the battle, etc. (For discussion and references see Lass 1997: 65f.6-7)

The use of metaphor in supposedly empirical disciplines however can lead to two classes of epistemological problems:

(a) Domain-mismatch: how 'real' is a given metaphor, and what might happen to you if you started believing it?

(b) Domain-transfer (my main concern here): What happens when metaphors that work reasonably well in one field get transferred to another, as they so often do?

Quine is a relevant sleeper. That good heuristic or other metaphors that seem to work well in one field do have a solid physical or coherence well-defined substrate can be hijacked (another metaphor of course) by others and used. Hijacking can produce the illusion of explanation or what is metaphorical pseudo-explanation. Linguists for instance constantly use metaphors from other disciplines, not to mention everyday life: rules are recursive, substitutions occur, features have hierarchical slots, boundaries which can be crossed, rules apply iteratively, clitics have hosts, parasitic vowels get into clusters, slots need to be filled (or the reverse), linguistic family trees branch, mothers tongues get lost, daughters who now each other's stories. This is mainly useful and entertaining: more seriously, the question is whether it does anything else. The fundamental issue is whether the terms of the metaphor actually have referents, or at least can point to some ontologically specified domain.

2. Historical Linguistics as a Metaphor-Hijacker

Historical linguistics has at various times—now more than ever—borrowed from geology. But why should it want to do this? The motivation belongs perhaps as much to the sociology of science as to the discipline itself: at certain theoretically stressful times self-descriptions and praxis are driven by what are perceived to be prestigious field-external models. At least since the

mid nineteenth century linguistics has borrowed terminology and conceptual domains from other sciences (in the loose, non-Anglophone sense of 'Wissenschaften'). The richest source has been biology, at points when it seemed to be the 'hardest' science that linguistics could reasonably hang itself onto[1]. A good-sized subset of linguists seems always wanted to be natural scientists. Some actually were: the nineteenth century August Schleicher was a botanist, and the contemporary William Labov was an industrial chemist. One major scar to this development in the nineteenth century was Darwin: linguistics in the 1850s was primarily a historical discipline, and Darwin was probably the first theoretician who succeeded in turning one of the classical historical sciences into something approaching synchronic natural science by sophisticated (if informal) modelling[2].

Subject-envy is not uncharacteristic of relatively new disciplines, especially ones whose ontological substrate is ill-defined. Linguistics has tried at various times to make itself look like biology, physics, mathematics, cognitive science, or all of them. And when in the 1930s and 1940s it looked like becoming about the best thing of its kind around, less secure subjects like anthropology began to borrow some of its concepts, imagery, and terminology. Linguistics is certainly the source of the fourth notion of an 'emic' level of organization, or what we might call 'structuralism' in general. The spirit of Saussure brooded over the face of the waters for a long time, and social scientists like Lévi-Strauss and Piaget were certainly among those who felt the (cold or warm) breath of linguistics.

But why should we linguists keep doing it, since others are apparently so enchanted by us that they have borrowed our lexicon and imagery? The answer I suppose is that we are not (or not yet) an empirical or hard science: we seem to rest uneasily somewhere among the 'human sciences' or 'Geisteswissenschaften', if with a widespread yearning for a different status. Even the universities that hire us are not always sure where we go: before we reached our theoretical puberty we used to get attached to

[1] See for instance Percival (1987); Davies (1987); and Wells (1987) on the 'organic/organismic' metaphorical domain in linguistics.

[2] Darwin himself was a great fan of metaphor: his work is full of good ones taken from elsewhere like the 'tree of life', and original ones like the 'tangled bank' and ecological 'wedges' (Darwin 1859).

anthropology. Over in language departments. (In two universities I have taught at, linguistics is in the Arts Faculty; in others it is among the social sciences; at my present university it is in the Faculty of Humanities (along with Arts and Social Sciences!); but it always seems a bit of a misfit wherever it is. My own view is that linguistics is not entirely suited to cohabit intimately with either 'arts' or 'social sciences' *versus situm* (at least as the boundaries are currently drawn). Since for me interpretation and concern is human cultural products, I rather think we are (or ought to be) a biological science, as the 'social sciences' ought to: culture at the end is inseparable from and based on biology (Dooley & Gonzales 1992; Wilson 1997:105–106). But most humanists would regard this as historical reductionism, and biologists would laugh at our pretensions. The problem has been nicely summed up by E. O. Wilson (1997:111–112):

> The principal theoretical difficulties of the social sciences are two in number. First, in the study of culture they are no 'natural kinds', basic atomic equivalents to genes, cells, and organisms that can form the base of permutational operations in analysis. The lack of natural kinds guarantees the second difficulty, 'atomic isolation'. Each major discipline –anthropology, sociology ... and so forth– has been required to develop its own conceptual base and language.

In the context of an interdisciplinary discussion involving geneticists, linguists and archaeologists, I think we can be honest about our envy. But the situation is complicated. One thing that has both helped and hindered us is that biologists want to talk about language and similar things as well, and we have been given (without asking) some new concepts that may be more trouble than they are worth.

3. An Example of Hijacked Metaphor

The problem Wilson raises, which upsets me enough to make me write this article, can be stated another way. Historical linguistics, unlike genetics or evolutionary theory, presently lacks a physical or quasi-physical basis for construing the objects it assumes to be passed down through time via variation and selection. We would have loved to have linguistic genes,

nucleotides, 3rd codons, base-pairing, even pretechnical as messy as polygenes, pleiotropy or incomplete penetrance. (I'll suggest later that maybe we might, in a speculative reconstruct of our domain.) At least, we would like to have ways of imaging things like this naturally, in a way that makes us less 'eccentric' and more 'mainstream' with respect to other sciences.

This brings me, finally, to my subtext: the use of hijacked generic metaphor in discussing characteristic 'genetic' preoccupations like descent, modification, copying errors in transmission, propagation of errors, lineages, sister taxa, cladogenesis, and assignment of lower-level entities to classes. Let us begin by seeing what happens when we as historical linguists appropriate, in what is in the conscious working level a morphological discipline, terms that represent (more or less) physical objects in another subject.

It is a curious fact that the first connection between genetics or (at least historical genetics—and historical linguistics was established before either discipline actually existed. In the eighteenth century, the first important formulations of genealogy in terms of 'lineages defined by shared errors' was developed by textual critics, in the course of work aimed at establishing archetypes or originals of major classical and biblical texts (Cameron 1987). Cladistics, as it were, was invented and used in an intelligent and sophisticated way for a century before evolutionary theory arose. And of course, until the rediscovery of Mendel's famous paper at the turn of this century, and the invention of the idea of 'mutation', even geneticists didn't have much idea what 'copying errors' might be. Genes themselves, in the early days when both they and a mutational theory of evolutionary change were first accepted, were not much more 'real' (or non-metaphorical) than viruses were to Koch or Pasteur.[2] There had to be something there to do what was being done, but it was another half century before we had a clear notion of what it was. And however much we don't

[2] In fact well into this century viruses were defined negatively, invisible pathogenic residues left behind in ceramic filters when everything visible like cells was filtered out. Hard (sensory) evidence for viruses was only available in the 1940s when the electron microscope provided wavelengths short enough to visualise them.

know about it still, at least there is something physical there, those helixes coil and separate and form templates for the attraction of complementary nucleotides, triplets code for proteins, DNA sequences make RNA that makes proteins by stringing together amino acids.

Now that genetics is in most ways a much more sophisticated operation than Historical/Historical linguistics, we are borrowing like mad from geneticists, and indeed the more generics I read the more I wish I was a geneticist, so I could avoid the intractable epistemological problems that face me as a historian of linguistic lineages. Or at least I had a really clear sense of precisely how I use genetic metaphor (assuming it is really metaphor and not something more precise) in my work.

The problem is this. Genetics, especially in combination with the so-called 'New Synthesis', provides a very attractive conceptual framework for describing linguistic change and language diversification, and for phylogenetic taxonomy. But unlike the geneticists, we lack a material basis for our model-building. Alternatively, maybe we have one, but it is too far away from what we are actually modelling to be of much use. (I make the tendentious assumption that somewhere at the bottom of all we do as linguists is not directly manifest in our practical work, lies a set of Darwinian processes in the brain (Dennett 1992; Calvin 1996; Churchland 1996—more on this below.)

A characteristic example of how we would use genetic imagery in historical linguistics might clarify things. In the past our terminology has been different from that of historical genetics (which is I think a fair construal of phylogenetic systematics), by and large, even when we are engaged in similar activities like establishing filiations, family trees, etc. I am going to use a mix of old linguistic and new genetic terminology to illustrate both the kinds of things we do that impinge on a cross disciplinary relation, and the problems this raises.

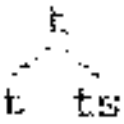
Take for instance the typical treatment of what we call 'correspondences', our primary tool in reconstruction and filiation. As a simple case, every time a native Germanic word begins in English or Dutch with a *t-*, it begins in German with a *z-*, phonetically [t] vs. [ts]. This correspondence is regular and replicable, as well as generally semantically transparent enough for us to recognize both the English and Dutch *t*-items and the German *z*-items as clearly 'descendants of a common ancestor'. (The

arguments for ruling out convergence are good, but not relevant at this point). Some typical cases:

English	*Dutch*	*German*
tongue	tong	Zunge
mouth	mond	Zaun
ten	tien	zehn
tip	tip	Zipfel
to	te	zu
two	twee	zwei
twelve	twaalf	zwölf
twenty	twintig	zwanzig

This looks like a clear example of a language family in which two members (English, Dutch) constitute a subgroup, with German as an outsider. The relationship is patent, as is the subgrouping. Dutch is 'closer' to English (in this respect) than German. In standard linguistic comparative method we would ask ourselves which represents the older or 'ancestral' state, *t* or *z*? For various reasons not at issue here (including the presence of *t*-forms in other closely related languages belonging to a different subgroup, e.g. Swedish *tunga, två, tand* = *tongue, two, tooth*), we would reconstruct the *t*-group as representing the proximate ancestor or proto-language, and call German innovatory in this respect.

Looking at this through the eyes of a different metalanguage (to produce a rather poor metaphor), we could say that *t* and *z* are character-states in a character-field, and English and Dutch (and Swedish) constitute a lineage showing one character state, and German another. Accepting that English and Dutch are not homoplasious, our question is whether they are a synapomorphy or a symplesiomorphy, or reversing the coin, whether German *z* is an apomorphy or a primitive character. Cladistically, whichever alternative we choose, the result is the same: there is an event of cladogenesis, there are two sublineages which 'diverge' at the point where we might say (taking the usual view) that '*t* becomes *z* in German'. Regardless of which metalanguage we happen to use, the conceptual basis is the same in biosystematics and historical linguistics. It is some version or another of a bifurcation like



So from the point of view of character inheritance (and there is no reason, mathematical or other, why a /t/ isn't just as much of a character as an eye), historical glotto-classic cytologists who are not antihistorical 'reformed dialects') and historical linguists are doing the same things: studying lineages in time, and looking for clues in the internal makeup of character profiles of lineage members that can date or order events with some reliability, establish ancestries, closeness of relation, nodes and groups of taxa.

4. Beyond Systematics: 'Units of Inheritance' in Linguistics

Until recently, the use of genetic metaphor and imagery in linguistics was largely restricted to stemmatic/cladistic matters. Up to a certain fineness of resolution this poses no problems: no matter what the subject, the characters, atoms, elements, or whatever can be taken, if necessary, 'agnostically' as relative points or nodes in trees, and the usual catchall terms like 'clade', 'taxon', 'primaeval', 'derived' can be used interestingly and uncontroversially.

But since the 1970s things have changed somewhat: a new field of discourse and epistemological display was opened up for exploration. Earlier, the problem of 'units' in the social sciences, including linguistics, was not taken seriously except at the synchronic or 'structural' level. Linguists didn't generally seem to ask questions like 'just *what* is it that is passed from generation to generation, that children acquire, that changes?'[4] Or if such questions were asked (as they were after Chomsky's various definitions of language as a 'mental organ', a computational module, etc.), the answers were not given or thought about in terms of anything

[4] Some of us did, actually, but the answers were not 'hard' in a biological sense, and bypassed the whole 'inheritance' question. See for instance the rather unsatisfactory discussion in Lass (1997:chs.1,6-7).

'heritable' in a proper sense. In the standard Chomskyan model, the child uses its internalized Universal Grammar module to deduce from linguistic input what language it is being exposed to (the 'child as little linguist' model), and 'acquires' it[5]. But strictly speaking, nothing material 'passes' or 'is transmitted' between generations; each performs computations on the output of the previous one, and somehow establishes the results of these computations in its left brain.

The first serious attempt at an answer to the 'units' question came in the middle days of early sociobiology, in the 'culturgens' of Lumsden and Wilson (1981). These were conceptual or cultural units (structure and/or substrate not specified) that construed the 'heritable material' of cultural evolution (or the cultural side of gene/culture coevolution in slightly later terms). But the soil this notion fell on was stony: in the 1970s human sociobiology was not very Politically Correct. There were also problems with Lumsden and Wilson's maths, the reviewing community was nearly uniformly hostile, and maybe the *Zeitgeist* was just wrong, or the presentation, and in any case the undertaking was tainted by association with Wilson in the influential left-wing circles dominated by figures like Gould and Lewontin. For some reason (largely careless and tendentious reading, I think), the early attempts at theoretical sociobiology got caught up in the 'nature' vs. 'nurture' debate, confused with vulgar genetic determinism and generally misunderstood[6]. So at least an opportunity was missed.

But a proposal in the later 1970s took off rather better, and has led to the growth of a new field or subfield of combined rational inquiry and flakiness. This was Richard Dawkins' invention in 1976 of the *meme*. Would he have done it if he had known how much trouble it would cause? Of course being a biologist he could simply drop it in the path of the humanists and social scientists, and let them get on with the exegesis and their own

[5] For a standard treatment of this view, but with a rare biological sophistication, see Pinker and Bloom (1992).

[6] There was a good deal of paranoia about the apparent right-wing totalitarian tendencies of sociobiology, partly because people who ought to have known better, like Stephen Jay Gould, saw in its genetic focus the (unintended) shadow of eugenics. See Gould's rather obtuse review of Lumsden and Wilson 1983, reprinted in Gould (1988:107-123).

proper work, using the concept if it seemed appropriate – which is more or less what happened.

Anyone interested in organism/culture relations, and as literate and cultivated as Dawkins was bound to see the possibility of heuristic analogy, and to want something more or less parallel to the gene which could play the part in cultural inheritance, structure and change that the gene does in the biological realm, and further would embody as many of the standard Darwinian-replicator system properties as possible[1]. That is, a cultural 'message-making system' should display at least:

- The presence of replicators.
- Relative fidelity of replication, plus the possibility of copying error.
- Selectional bias or 'sifting' in the course of historical transmission.
- The possibility of going to stabilization in a population.

And, to add Dawkins' special personal take on genes and their relation to the organisms they occur in and are blueprints for, the possibility of 'selfishness'. Twenty-odd years on it is still not entirely clear what memes are, where they live, or what kind of internal structure they might have.

I suppose that if memes are like genes, they are rather more like the broader definition in *The extended phenotype* (1982): units of whatever size or constituency that produce coherent phenotypic effects. The typical meme may be more like a polygene on the one hand, or a highly pleiotropic single gene on the other. But this is too early in the story to get into that.

One characterization, that of John Hyde Bonner (1980), is so loose and insightful as to be really tempting. He manages to use the concept very elegantly without ever telling us precisely what it is. First he notes uncontroversially that 'Dawkins has not attempted a rigorous definition of a meme' (though we would all agree that he gives enough examples for an ostensive definition in terms of 'things like'). Then Bonner (1980:17) provides his own: 'any bit or collection of bits of information passed by behavioural means from one individual to another'. This somewhat skimpy definition provides him with a beautiful framework for comparing behaviours and

[1] This would, if successful, unify both 'mental' and 'physical' phenomena under the rubric of 'universal Darwinism' (see Dawkins 1983).

elements of what we may call 'culture' from nearly all organizational grades in the animal kingdom, and constructing something of a historical narrative.

Memes have all the correct Darwinian properties, and another, normally reserved for pathogens, which my sense is taken as their real secret: 'inferiousness'. Dawkins sees memes (especially ones he doesn't care for, like religion) as 'parasitising' minds, like cultural tapeworms. This would appear to make them rather more like whole organisms with their own 'interests' than genes (which though having interests, are generally not thought of as parasitic). But whatever, we do have a problem of specificity: if memes are nothing more than just 'bits of information', the concept is too vague to be of much theoretical use.

The lack of specificity is indeed troubling. At least the following have been considered, in one place or another, as reasonably prototypical memes: lines of verse, proverbs, novels, tunes, chord-progressions, symphonies, phonemes, syntactic constructions, political systems, kinship systems, modes of agricultural production, insect taboos, religions ... or just all the potential inhabitants, of any size or degree of structural complexity, that might be found in Popper's equally vague 'World 3'. Popper of course carefully avoids most of the hard epistemological problems by claiming first that World 3 objects are 'ideal', and second that they 'exist, but exist nowhere' (Popper & Eccles 1977:480). I don't think this is good enough. Perhaps we should look at a recent proposal about what memes might 'really' be, how they could be given a physical basis, and might be replicated, inherited and changed.

Nikolaus Ritt (1995) has made a brave if problematical attempt to specify, within a theory that allows for both memes and biological realism, what he calls 'linguistic genes'. This appears to be the first proposal for a micro level characterisation of what a meme might look like, how and where it might be organised, and what its replications would consist of. That is, he presents a first approximation of a physicalist and reductionist linguistic version of memetics, with a quite biological looking substrate and *modus operandi*.

His primary example of a linguistic replicator is a speech-sound or phoneme. He suggests that successful acquisition by a child of a speech sound of its native language ought to be considered from a radically nontraditional point of view. That is, rather than looking at the child as an

'agent', using as wired-in computational language module to process input (though of course this is part of the process). Ritt takes the viewpoint of the acquiree (the meme) rather than the acquirer. Since children as potential speakers can't control their language acquisition, it might be reasonable to look at the acquired items, fed to children by the speech community, as parasitising them and replicating in their brains, and see the acquisition process from *their* point of view. Thus the language learner becomes a Dawkinsian vehicle for the propagation of linguistic 'genes' = memes.

'Language itself' then (whatever this really means) may not be physical, but it has a physical substrate. It is cognitive, which in a non-dualist epistemology means neurological. Ritt appears to take it for granted (as I do) that all storage and activity in the brain are in some sense material or physical — on grounds of parsimony if nothing else. Occam's Razor alone (even if there wasn't a wealth of better philosophical and empirical argument) would militate against Cartesian dualism or Ghosts in the Machine[8].

In this provocative if preliminary paper[9], Ritt tells a story of the 'neural Darwinism' type, in which memes compete in the cerebral arena for space, and those that are ultimately selected win out, and are established in the speaker as part of his 'linguistic competence'. Ritt suggests as the substrate for this process something like a Hebbian cell-assembly (Hebb 1949), a neuronal cluster or unit of high connectivity and specificity, with projections to and from all centres involved in speech. Thus the phoneme */t/* would in the adult speaker be stored rather like any other small-scale (accessible, 'declarative') memory.

So in the mature user a */t/* or any other unit is hypothetically some kind of small discrete assembly. This is not at all out of line with recent theorising about the nature of storage units, e.g. Calvin's diffuse network of neocortical pyramidal hexagons as the basic units of the 'cerebral code' (Calvin 1996). We assume then, following this speculative but sensible line, that in the mature speaker each such assembly has afferent projections (e.g.

[8] For some of the better arguments see Dennett (1991); Damasio (1994); Cairns-Smith (1996); and Carruthers (1996).

[9] For further commentary on Ritt's paper, from other points of view, see Lass (1996).

from the auditory cortex) and efferent ones (to the motor cortex)[10]. There are then two sides to this assembly's behaviour:

A. *Productive*. Each time the motor cortex triggers this assembly, it 'fires' or activates or is represented, much as in Calvin's view memories are activated in the course of identifying an object. Thus when I say a */t/* (this is my interpretation, not Ritt's), it is within the model the linguistic 'equivalent' (a metaphor) of an instance of gene-expression'.

B. *Receptive*. Each time the auditory cortex perceives and mediates a token of this unit coming in from the environment, it is identified in the first instance as such a token via afferents to (presumably) the medial temporal lobe, and from there further identified as a 'code unit', and processed by the more 'dedicated' language centres in the posterior temporal and temporo-parietal regions ('Wernicke's area').

But how would such memetic assemblies get established in the first place, so they can serve as substrates for these hypothetical 'genes' or units? Ritt suggests differential reinforcement of synaptic connections through stimulus-repetition and some kind of accompanying 'reinforcement'. This the mechanisms behind the establishment and use of the stored meme would be entirely consonant with neurobiological principles. Consistent perception or activation in the motor mode strengthens specific synaptic connections (presumably in the first instance by Long Term Potentiation of neurons in the hippocampus, Squire 1987); and a stable collection of communicating and functionally integrated synapses forms the internal representation of a meme. This we can get cross-cortical 'cloning' of copies of the meme, which replicates itself within the speaker in competition with all other similar memes (e.g. speaker inputs 'from other individuals' or output variants

[10] Ritt does not go into much neurological detail; speculations about parts of the brain active in the processes he suggests, and mechanisms, are my addition, designed to flesh out the picture a bit more.

[11] This would presumably occur through efferents from the 'language centres' like 'Broca's area', presumably with retrieval of posterior stored material via the arcuate fasciculus.

the community). And to the learner, it is exposure to the meme that sets up the memory and retrieval circuits necessary to its later incorporation and regular production. So the acquisition story becomes Darwinian (in either the Dennett or Calvin sense): replication of one assembly among many leads to synaptic strengthening, and other near-competitors are selected against, so that eventually the internal wiring comes to look more or less as it did in the parent, or the speech community, from which the meme propagates or attempts to.

The language system of the community then is a 'memeome' (Kitts's term), which attempts to replicate itself, meme by meme, in the brains of learners. Successful language acquisition is the end-product of replication of a set of memes (not as successful both of a normal fission as of replication of the zygote's genome). In the case of language acquisition, the child is 'rewarded' with communicative competence by mastering or internally replicating—or better shifting the internal replication of—the memeome. And of course there can be miscopying as well as variation, or infidelity 'alleles' can win out, and new lineages can be established by propagation of these variant copies, which will compete with the old ones. Thus language change is enabled by essentially the same mechanisms as acquisition—though not necessarily at even very likely at the same ontogenetic stage. That is, after acquisition later peer groups can be loci for locally more prestigious variants, which would then presumably undergo the same kind of competition with established variants. In Kitts's (1995:53) words: 'speakers figure exclusively in the environments in which the replication of linguistic elements takes place'.

This is at the moment somewhat simplistic, underdetailed, and speculative, but no less valuable for that. Its great virtue is that it suggests a way in which we might construct research programmes incorporating and testing biological metaphor. This kind of thinking could just make our use of other folks' metaphors precise enough so we can find out if there is any genuine ontological or structural congruence. And that would be nice to know, whether the answer is positive or negative.

Professor Emeritus of Linguistics
Honorary University Research Associate
Department of English
University of Cape Town

References

Barkow, JH, L Cosmides & J Tooby (eds). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford: Oxford University Press.

Beckett, DS (ed) 1983. *Evolution from Molecules to Men*. Cambridge: Cambridge University Press.

Bonner, JH 1980. *The Evolution of Culture in Animals*. Princeton: Princeton University Press.

Cairns-Smith, JG 1996. *Evolving the Mind: On the Nature of Matter and the Origin of Consciousness*. Cambridge: Cambridge University Press.

Calvin, WH 1996. *The Cerebral Code: Thinking a Thought in the Mosaics of the Mind*. Cambridge, MA: MIT Press.

Cameron, HD 1987. The Upside-down Cladogram: Problems in Manuscript Affiliation. In Hoenigswald, HM & LF Wiener (eds): *Biological Metaphor and Cladistic Classification. An Interdisciplinary Approach*. London: Frances Pinter.

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

Damasio, A 1994. *Descartes' Error. Emotion, Reason, and the Human Brain*. New York: Putnam.

Darwin, C 1859. *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Davies, AM 1987. 'Organic' and 'Organism' in Franz Bopp. In Hoenigswald, HM & LF Wiener. *Biological Metaphor and Cladistic Classification. An Interdisciplinary Approach*. London: Frances Pinter.

Dawkins, R 1976. *The Selfish Gene*. Oxford: Oxford University Press.

Dawkins, R 1982. *The Extended Phenotype*. Oxford: Oxford University Press.

Dawkins, R 1983. Universal Darwinism. In Bendall, DS (ed): *Evolution from Molecules to Men*. Cambridge: Cambridge University Press.

Dennett, DC 1991. *Consciousness Explained*. London: Viking.

Edelman, GM 1992. *Bright Air, Brilliant Fire: On the Matter of the Mind*. London: Allen Lane.

Edelman, GM 1987. *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.

Edelman, GM & G Tononi 2000. *A Universe of Consciousness: How Matter Becomes Imagination*. London: Allen Lane.

Ehrlich, P 1991. *A Science of Mind: The Quest for Psychological Reality*. London: Academic Press.

Gould, SJ 1988. *An Urchin in the Storm. Essays about Books and Ideas*. London: Collins.

Hebb, DO 1949. *The Organization of Behavior*. New York: Wiley.
Hoenigswald, HM & LF Wiener 1987. *Biological Metaphor and Cladistic Classification. an Interdisciplinary Approach*. London: Frances Pinter.
Lass, R 1990. Of Noses and Mouses: On the Trail of the Wild Replicator. *Vienna English Working Papers* 3.1:3-1.
Lass, R 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
Lumsden, CJ & EO Wilson 1981. *Genes, Mind, and Culture*. Cambridge, MA: Harvard University Press.
Lumsden, CJ & EO Wilson 1983. *Promethean Fire: Reflections on the Origin of Mind*. Cambridge, MA:
Paton, RC 1992. Towards a Metaphorical Biology. *Biology and Philosophy* 7.279-94.
Pennati, PK 1987. Nonsyntactic Diversity in the Study of Languages before the Advent of Comparative Grammar. in Hoenigswald, HM & LF Wiener (eds): *Biological Metaphor and Cladistic Classification. An Interdisciplinary Approach*. London: Frances Pinter.
Pinker, S & P Bloom 1992. Natural Language and Natural Selection. In Barkow, JH, L Cosmides, & J Tooby (eds): *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford: Oxford University Press.
Popper, KR & J Eccles 1981. *The Self and its Brain*. Berlin: Springer.
Ruhl, N Language Change as Evolution: Looking for Linguistic 'Genes'. *Vienna English Working Papers* 4,1:43-56.
Squire, LR 1998. Memory and Brain Systems. In Rose, S (ed): *From Brains to Consciousness? Essays on the New Sciences of the Mind*. London: Penguin Books
Tooby, J & L Cosmides 1992. The Psychological Foundations of Culture. In Barkow, JH, L Cosmides, & J Tooby (eds): *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford: Oxford University Press.
Wells, RS 1987. The Life and Growth of Language: Metaphors in Biology and Linguistics. in Hoenigswald, HM & LF Wiener (eds): *Biological Metaphor and Cladistic Classification. an Interdisciplinary Approach*. London: Frances Pinter.
Wilson, EO 1997. *In Search of Nature*. London: Penguin Books.

Cultural Linguistics and Shona Noun Classifiers

Gary B. Palmer

1. Introduction[1]

In the 1960s and 1970s, most linguists and linguistic anthropologists studied grammar as an innately configured, abstract realm having an almost mathematical precision. In the realm of semantics all categories were taxonomic, with category membership based on the possession of certain necessary and sufficient features. This logician's image of grammar and meaning divorced from everyday life encountered a dramatic challenge in 1980, when George Lakoff and Mark Johnson published *Metaphors We Live By*. Lakoff followed it with *Women, Fire and Dangerous Things* in 1987, and in the same year Ronald Langacker published *Foundations of Cognitive Grammar* I. With these landmark publications, the hermetic seal of idealist grammar was broken and the scientific study of semantics began to look outward to general cognitive processes, encounters with the physical world, communication, and culture. A paradigm change was underway. The new semantics was a semantics of life.

The virtue of the new approach was that it found the source of semantic categories in embodied experience and encyclopedic or world knowledge. This means that linguistic meaning was seen as emergent from physical experiences and as acquired from other people in the course of infant nurturance, growing up among peers and parents, and living in society as an adult. Culture and history could now factor into the semantics of lexemes and grammatical constructions, where in prior theorizing they could only influence

[1] This paper is excerpted and revised from 'When Does Cognitive Linguistics Become Cultural? Case Studies in Tagalog Voice and Shona Noun Classifiers' (in Luchjenbroers)

language performance. For example, Lakoff has argued that metaphorical idioms involve cultural knowledge in the form of conventional images and that units in radial semantic categories are structured by experiential domains, which may be culture-specific (Lakoff 1987:95, 1999:69)[1]. Langacker (1999:13,16), too, has recently reaffirmed that 'language is an essential instantiation and component of culture, whose reflection in linguistic structure is pervasive and quite significant'.

I have used such observations as a starting point for *cultural linguistics*, an approach which foregrounds cultural schemata and cultural models in explanations of grammar and semantic patterns (Palmer 1996). In this respect it contrasts with the typical practice of cognitive linguists, who, in spite of their recognition of the importance of culture, typically foreground universal cognitive processes such as figure-ground relations, spatial schemas, force dynamics, prototype categories, and Lakoff's famous *Idealized Cognitive Models*, leaving cultural dimensions of language somewhere in the background, or at least relabeled as such. Cultural linguistics offers a shift in emphasis. Though it draws on the theory of cognitive linguistics for many essential analytical concepts, it explicitly extends cognitive linguistics into cultural domains and it treats cultural categories as potential semantic categories.

Specifically, I am claiming that many grammatical phenomena are best understood as governed by cultural schemata rather than universal innate or emergent cognitive schemata. The sources of such cultural schemata include mythology, such as the Australian Dyirbal myth of the sun and moon, which Lakoff used to explain membership in Dyirbal noun classes (Lakoff 1987). They also include social structure, repetitive domestic and subsistence activities, salient rituals, and a host of other cultural phenomena. For instance, they include such activities as the pulverizing of maize or meatie with a mortar and pestle, an activity practiced throughout Africa, mainly by women and girls. The daily routine of lifting and dropping the pestle and hearing the thump, time after time, must surely entrench the scenario and embody the schemas of lifting, of the falling pestle, and the crushing, punctuating, reverberating thumps, felt in the hands and feet as well as heard. The emergent categories must also register the femaleness of pounding grain. The experience of pulverizing is culturally structured in at least two ways: first, by

[1] Lakoff actually used the phrase 'characterized by', rather than 'structured by'.

the assignment of tasks by gender and age; second by the technology of the mortar and pestle, which are cultural artifacts. If such basic embodied cultural experiences structure semantic categories, then we should expect to see their expression in grammar. The example reveals how essential it is that linguists do ethnography or at least read it systematically as a source of semantic categories. Linguists can not rely solely upon their own non-native intuitions about the semantics of complex domains (Myhill 1995). The purpose of this paper is to apply the approach of cultural linguistics to the analysis of the noun classifier system of Shona, a Bantu language of Zimbabwe, with special attention to class 3/4.

2. General Theory of Cognitive and Cultural Linguistics

Unlike postmodern approaches to cultural theory, which posit no fixed points of reference or stable meanings, cultural linguistics sees grammar as an entrenched system of meaning and form. Following Langacker's (1987; 1991a & b; 1999) theory of cognitive linguistics, the minimal units of grammar are verbal symbols, each of which represents a linkage of two kinds of units, one phonological, the other semantic. Semantic units are characterized relative to semantic domains (Langacker 1987:63). Since these may include any concept or knowledge system, linguistic semantics is encyclopedic and therefore cultural in scope. When a class of linguistic expressions is seen as relative to one or more semantic domains of relatively extensive scope with complex category structures and rich details, then cognitive linguistics becomes decidedly cultural. It is this difference in emphasis and elaboration of the cultural dimension, not an underlying difference in theory, which justifies the new label of *cultural linguistics*. The label also differentiates the approach from that of contemporary linguistic anthropology, which is typically discourse-oriented and heavily invested in pragmatism, often displaying scant interest in cultural categories or cognitive processes. In my view, culture and cognition are not separate entities, just two views on the process whereby people with minds, which are embedded in physical bodies situated in social and physical environments, communicate, learn, think, and pursue social goals. Similarly, Edwin Hutchins (1996:354) proposed an integrated view of human cognition, 'in which a major component of culture is a cognitive process ... and cognition is a cultural process'.

Certain types of cultural models merit special attention from linguistic anthropologists and culturally oriented linguists. These are *scenarios* and *polycentric categories*. In the case study of noun classifiers in Shona, I demonstrate that a better understanding of classifiers can be achieved by analyzing each singular/plural classifier pair as a polycentric category. Scenarios are important elements of the polycentric category. The polycentric category is a synthesis of Langacker's (1987) complex category with Lakoff's (1987) radial category. Unlike the radial category, which has a single central prototype, a polycentric category has multiple central categories connected by conceptual metonymies. In the next section I will elaborate on the concepts of scenario and polycentric category. Then, in the following sections, I will apply them to the case studies.

2.1 Scenarios

Scenarios are schematic cultural models of action[3]. Cultural linguistics is based on the premise that grammar is relative to imagery that derives from cultural models. Cultural models are cognitive entities, but they are often more richly elaborated and further removed from basic physical and cognitive experience than the spatial-mechanical schemas and figure-ground relations typically investigated within cognitive linguistics. Examples of cultural models include the conventional knowledge systems governing kinship, ways of preparing food, navigation, rituals, myths, ceremonies, games, and speech events such as conversations. Imagery arises from construing models at different levels of abstraction or specificity, from different points of view, or at different stages in a process[4], and from admitting various features of models within the scope of attention (Langacker 1987, Lakoff 1987; Palmer 1996).

Cultural models include some, but perhaps not all, of what Lakoff (1987:113-14) termed *Idealized Cognitive Models*, in which he included propositional, image-schematic, metaphoric, and metonymic models. With respect to metaphoric and metonymic models, it seems more accurate to speak of metaphoric relations between models or parts of models, or to say that models comprise functional relations, which provide the imagery for verbal metonymy. Universal image-schemas derived solely from the common experience of inhabiting a human body would not in themselves be cultural models. However, universal image schemas may be incorporated into cultural models, and in fact most physical experience reflects not only universal constraints, but also cultural modifications or culturally specific uses of tools, dwellings, and habitats. Embodied universal categories may simultaneously belong to cultural domains.

It is more precise to recognize the elements of convention and social construction by referring to some kinds of linguistically significant models as cultural, while conceding that all cultural models are also cognitive. Most ICMs are cultural products, and the same may be said for *domains of experience* (Lakoff 1987). Thus, it seems more appropriate and accurate to refer to an approach which examines such cultural constraints on language as *cultural linguistics*. By using the term, we make it obvious that existing ethnographic studies contain a wealth of information of potential immediate use to linguistic theory.

Relatively abstract or decontextualized images are called schemas or image-schemas. Those involving actions and sequences of actions are scenarios. The scenario concept is particularly important in cultural linguistics because the term directs attention to the imagery of social action and discourse, which has largely been overlooked by cognitive linguistics, particularly in the study of non-Indo-European languages. The reason for its neglect may lie in the fact that scenarios are strongly influenced by history and socio-cultural context and therefore relatively independent of more basic cognitive processes of attention, accessibility or saliency of information, and basic concept formation which many linguists regard as the strongest determinants of grammar. It is true that Langacker (1987:63) extended as possible semantic domains 'the conception of a social relationship' and 'the speech situation', but at the very least, one can say that social scenarios have not been clearly delineated as a type of imagery having linguistic significance to the same extent as, for example, spatial imagery. And yet, humans probably direct as much verbal attention to orienting in society as they do in space, if not more. Not all of this social orientation can be reduced to metaphors of force and space. The approach pursued here resembles that of Anna Wierzbicka in that her cultural scripts are something like scenarios

[3] Lakoff (1987) treated a scenario as a kind of Idealized Cognitive Model (p. 75) and equivalent to a script (p. 284). He regarded it as metaphorically structured by a SOURCE-PATH-GOAL schema in the time domain (p. 285) and having a 'purpose structure, which specifies the purposes of people in the scenario' (p. 286). My usage is more general.

[4] The construal of schematic processes at different stages has been termed *image schema transformation* (Lakoff 1987:440-444; 1988:144-149).

(Wierzbicka 1996; 1997; Palmer 2000). However, unlike Wierzbicka, I do not reduce scenarios to statements composed of a small set of semantic primes. I take scenarios to be gestalts or constructions built up from lower level scenarios and event-schemas.

2.2 *Polycentric Categories*

Cognitive linguistics presents us with at least two types of complex categories. The first is Langacker's, which he characterizes simply as a *complex category* (Langacker 1987:371; see also Palmer 1996:96-97). It begins with a prototype and a variant. Since these necessarily have something in common, there is also a schema, which is elaborated by both the prototype and the variant (Figure 1). Langacker's complex category appears to have no place for conceptual metonymy.

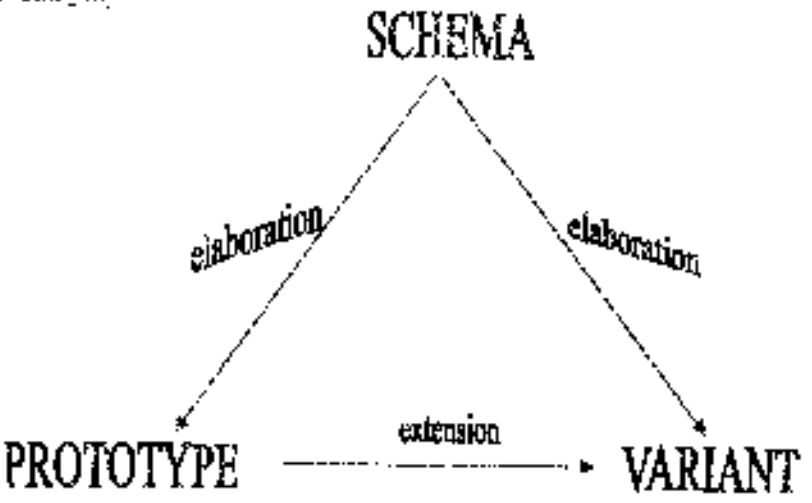


Figure 1: Complex category as envisioned by Langacker (1987).

Another kind of complex category is the radial category as described by Lakoff (1987). A radial category has a central subcategory and non-central extensions or variants. This is very much like Langacker's model, except that Lakoff does not include the schemas which can be abstracted from each extension of the prototype to a variant. In his discussion of Dyirbal noun classes, Lakoff (1987:95) also states that 'complex categories are structured by chaining: central members are linked to other members, which are linked to

other members, and so on'. Some of the links which he describes as conceptual metonymies (the sun causes sunburn); others are by similarity (sunburn is like the sting of the hairy mary grub), or variant to prototype (the sun is a mythical woman). Rather vaguely, he asserted that *Experiential Domains* and *Idealized Cognitive Models* can 'characterize links in category chains' (Lakoff 1987:95). A bit of cultural theory seeps in as well: '*Experiential Domains* ... are basic domains of experience, which may be **culture-specific**' (bold face added).

I hold that such linguistically significant experiential domains are in most instances actually cultural scenarios that have been given high salience by virtue of occurring in myth, ritual, crisis, social structure, or even the daily drudgery of domestic life. The functional links within domains are what we regard as conceptual metonymies. In a further suggestion of the importance of

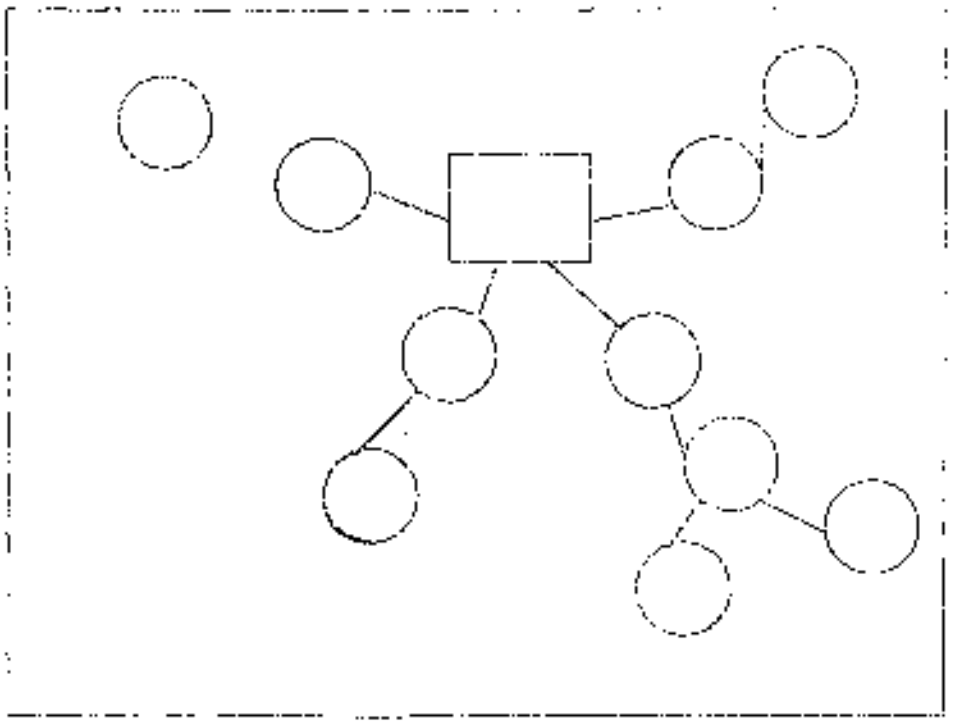


Figure 2: Radial category balan as envisioned by Lakoff (1987)

conceptual metonymy over schematization. Lakoff (1987:9n) asserted that 'specific knowledge (for example knowledge of mythology) overrides general knowledge'. We are left with a picture of a category that has a central prototype from which radiate a number of chains based on similarity and conceptual metonymy (Figure 2).

Lakoff used this concept to develop a theory of Dyirbal noun classifiers. Three of the four classifiers were characterized as radial categories (*bayi, balan, balam*). The fourth (*bala*) was characterized as an 'everything else' category. Noun classifiers represent a common and important kind of grammatical category, which was once thought to be arbitrary organized. Lakoff (1987) demonstrated that a class may have hundreds of members that share no common features of meaning. In my opinion, this important advance in the theory of linguistic categories depended crucially on understanding the governing role of cultural scenarios.

Tom Myhre (1995) took issue with Lakoff's (1987) analysis of Dyirbal noun classifiers, accusing him of imposing a Western world view on the Dyirbal system because it proposed human males and females as prototypes for the classes *bayi* and *balan*. Myhre proposed instead that the linguist should seek to discover which concepts have particular relevance for the Dyirbal and use those as the basis for the analysis. He proposed that the four classes of *bala, balan, bayi*, and *balam* could each be defined by combinations of values on the dimensions of *potency* and *harmony*, which have special relevance in Dyirbal culture and society. Thus, Myhre's critique appears to be an argument for an explanation that is more cultural than cognitive, but based on parameters or features, rather than on scenarios or cultural models.

My analysis of classifiers is like Myhre's in two respects: First, I am arguing that the important criteria for classification are concepts that are culturally salient. Second, I am arguing that one finds no single prototype at the center of a typical noun class. But unlike Myhre, I do not try to explain the category by replacing the prototype with one or two abstracted dimensions. (Similar approaches have been attempted in Bantu studies by Contini-Morava 1994; Spitulnik 1987; 1989 with unsatisfactory results, as discussed by Palmer & Antal 1999 and Palmer & Woodman 1999).

The third type of complex category is the *polycentric category* as proposed by Palmer and Woodman (1999).

The third type of complex category is the *polycentric category* as proposed by Palmer and Woodman (1999).

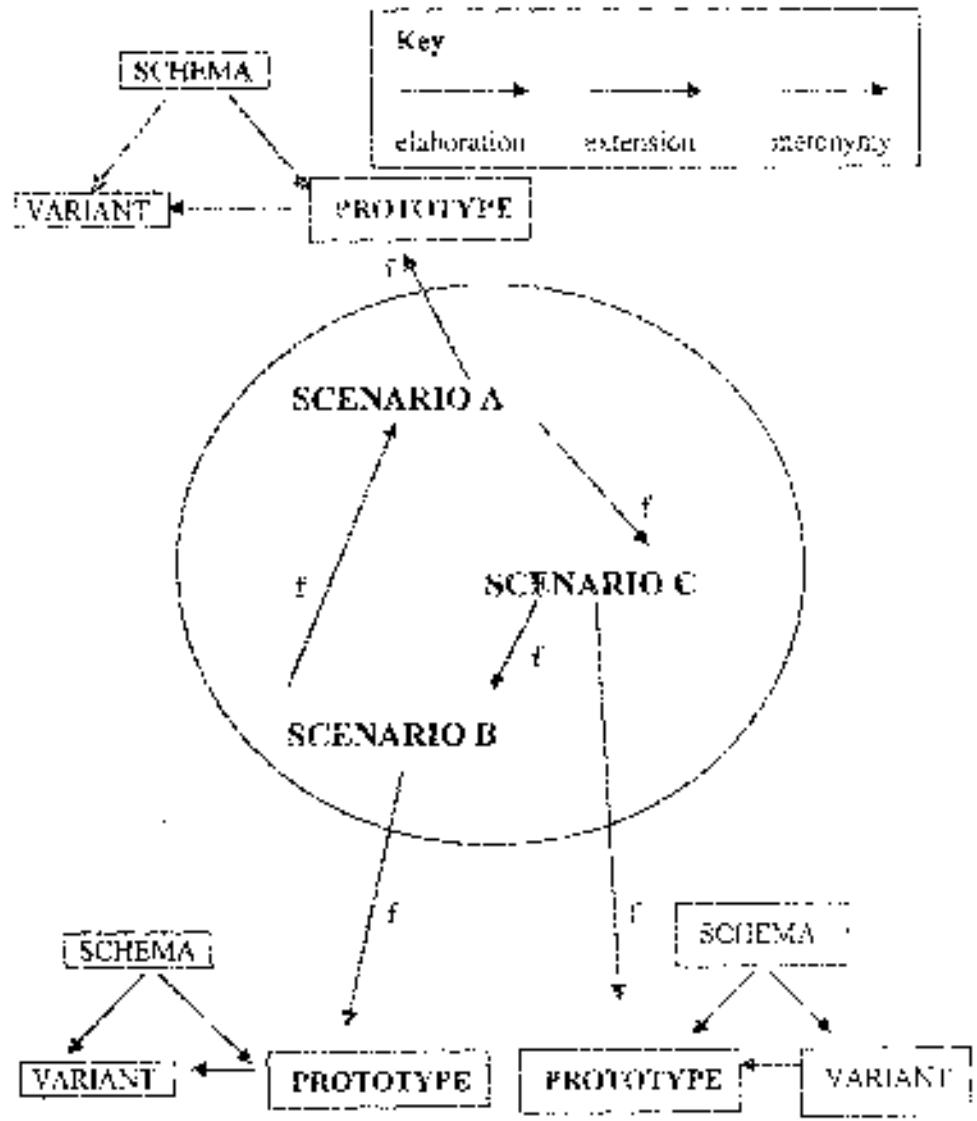


Figure 3: Schematic of polycentric category as proposed by Palmer and Woodman (1999)

A polysemous category has multiple central categories, each of which may be modeled on a prototype derived from the scenario (Figure 1). I show only scenarios in the central region of Figure 3. I treat the central categories as a flat and complex structure rather than as parameters which must have contrasting values across categories. Second, I would not rule out the possibility of a level of interest that would apply across classes to clusters of category members. In central categories are related to one another and to more peripheral categories and instances either by function (contiguity, conceptual correlation), by similarity (prototype to variant, metaphor), or by schematization (reference to more abstract). I call these complexes *polysemic categories*. They consist in part of complex categories as defined by Langacker (1987:371) and of radial categories as defined by Lakoff (1987). Some of Langacker's kinds of polysemic categories are all embedded in cultural scenarios (or in the sets of cultural models one draws on in an act) both cognitive and cultural.

3. Shona Noun Classifiers as Polysemic Categories

Many languages have gender classifiers that segregate nouns. There are, for example, the genders of Guarani and Romance languages, the numeral classifiers of Chinese, Japanese, and Maya, the verbal classifiers of Navajo, and the 20 or more classifiers of the Bantu languages (if one counts the plurals). Other languages have substantive affixes that can function as classifiers. These would include, for example, the anatomical suffixes of Tarascan and Southern Tiwa (Quoted in Allen) (Friedrich 1970:394-395; Palmer 1996:60,145-146)[7]. Even phonemes can function as classifiers in sub-Saharan languages (Bernárdez n.d.).

For decades linguists have struggled to make semantic sense of classifiers. Most commonly they have concluded that the assignment of lexemes to classes is arbitrary or that the classes center on such basic physical qualities as shape, texture, number, and animacy. While there is some explanatory value in the physical prototype approach, it has ultimately proved to be limited, leaving unexplained such interesting phenomena as the presence in some Bantu languages of the human term *chief* in the same class

[7] The figure of 20 for the Bantu classes includes singular and plural forms. If these are not counted separately, the figure would be ten. Classes 1 and 2 for (CP), for example, *lala* is the singular and plural of the class that includes most terms for humans.

as wild animals (Guthrie's 9/10) (Creider 1975; Guthrie 1967). Another approach was needed.

As early as 1959, the famous paleontologist Louis S.B. Leakey proposed in his Kikuyu lexicon book that the noun classes are ranked on a hierarchy of spiritual value. For example, humans appear in Leakey's class 1 (Guthrie's 1/2), the highest in spiritual value; class II (Guthrie's 3/4) is for "second class spirits"; and class III (Guthrie's 9/10) is for all other living creatures. Regarding Guthrie's class 5/6, Leakey (1959:13) asserted that "every single word in this class is an object which is used, or has been used until recently, in connection with religion, magic or ritual or some other form of ceremonial". To my knowledge, Leakey's proposal was never followed up.

The year 1987 saw a breakthrough in the understanding of classifiers. The key to their explanation was most widely publicized by George Lakoff in the book that drew its title *Women, Fire, and Dangerous Things* from a noun class of the Dyirbal language of Queensland. Lakoff was actually reshaping a middle-level theory proposed by Dixon (1982). Lakoff held that each noun class had a central member and that other members were linked to the central member by category chaining. The basis of the chaining was a common domain of experience, which was culture-specific. The Dyirbal classifier *balan* (one of four) marks a category whose central member is human females. In Dyirbal mythology, the sun was a woman. Other members of the class were birds (mythical females) and plants and animals who either appeared in the myth or were seen as somehow similar to fire (they were hot or may had stingers). Fire belongs to the class because it belongs to the same domain of experience as the sun. Thus, with some exceptions, category membership seems neatly explained by this approach. Problems with the approach have been raised by Mylne (1995), whose critique was discussed previously.

In the same year, Debra Spitulnik (1987) published a study of Chewa (Bantu) classifiers[8]. Her approach leaned heavily on highly abstract schemas, which she called "central notional values", but she also proposed that some nouns belong in their classes by virtue of cultural associations. "The [ChiBemba] noun *mfumu* 'chief' occurs in the class dominated by nouns for wild animals (Cl. 9/10) because of the cultural association of the chief with the animal world" (Spitulnik 1987:110) [e.a.]. She did emphasize the cultural approach, because in her view, grammatical factors compete for control over the classifiers. At about the same time, Ellen Contini-Morava proposed in a

[8] See also, Spitulnik (1989).

paper made available on the internet that the Swahili (Bantu) noun classes were dominated by 'superschemas' that were linked by schematicity and extension to spatial, supernatural, and psychological features and schemas (Contini-Morava 1994).

To sum up these approaches to understanding classifiers, Leakey described classification by spiritual hierarchy, Dixon and Lakoff showed clear mythical motivations for Dyirbal classifiers, Spitulnik presented a plausible cultural explanation for the apparently anomalous classification of Chewa chiefs, and Contini-Morava saw supernatural schemas underlying Swahili classes. These observations suggest that it might be worthwhile to apply a cultural approach to the Bantu classifiers with special attention to the supernatural and to apply the approach more systematically than had been previously attempted. That is what I and students Bonibeu Neal Arin, Claudia Woodman, and Russell Rader have begun to do for the Shona language of Zimbabwe. But before discussing those findings, I will present a brief description of the classifier system involved:

> Bantu noun classifiers are defined by characteristic prefixes on the nouns and concordial affixes on adjectives, verbs, and deictics. The classes are usually designated by numbers from 1 to 22. In classes 1 to 13, odd numbers are singulars, even numbers are plurals. Thus, for Shona singular class 1, *mu-*, the plural is class 2, *va-*, and for singular class 3, *mu-*, the plural is class 4, *mi-*. Of the first 15 classes identified by Guthrie (1967), the only ones to which he attributed clear semantic correlates are 1/2 (persons) and 9/10 (animals). He observed that parts of the body appeared more frequently in 3/4 and 5/6, but otherwise found no definite correlations of meanings to classes. Fortune (1955) observed that 'class 3 contains nouns indicating trees, parts of the body, atmospheric phenomena, things characterized by length, *and miscellaneous*' [e.s.]. The only atmospheric phenomena that he listed are *m'bado* 'breeze, wet weather' and possibly *m'ke* 'air, soul' and *mandi* 'cold.' (Palmer and Woodman 1999).

Palmer (1996) and Palmer and Arin (1999) proposed that the semantics of classifiers in Shona and other Bantu systems are governed by salient ritual scenarios that are more culturally specific and richer than the stereotypes and features proposed by Spitulnik (1987; 1989) and Contini-Morava (1994). After reading all available ethnographies of Shona culture and society, they

identified nine specific and two general scenarios that might govern the distribution of Shona noun classes. Scenarios 1, 2, 10, and 11 are listed below. The numbers of these scenarios do not correspond to the numbers used by Bantuists to identify the noun classes.

1. *The spirits of ancestral chiefs live in the bodies of lions (mweneni).*
2. *The chiefly ancestral spirits (mhondoro) reign over both the things of the wild and human affairs. They are the protectors of the land and the wild animals.*
10. *There is a scenario of protection in which the central participants are dominating protectors, protected ones, and the agents of domination.*
11. *There is ritual danger, resulting mainly from foreign ancestors with grievances or from contact with the paraphernalia of mediums.*

Palmer and Arin (1999) proposed that Guthrie's class 9/10 is governed by scenario 10 (which also subsumes scenarios 1 and 2), and that Guthrie's 5/6 might be governed by scenario 11. Subsequent research by Rader (1998) suggests that 5/6 is more directly governed by the imagery and mythology of fertility'. Palmer and Woodman (1999) examined Guthrie's class 3/4, finding that its central members involve an important domestic scenario and an ethno-ecological model as well as mythical and ritual scenarios. Central physical items in this class are those used in ritual and domestic activities. There is a network of salient categories and chains of extension, which justify using the term 'central' for the salient categories. We concluded that a noun class is more than a radial category centering on a prototypical member or a single domain of experience. It is more like a network of radial categories based on a cross-section of the cosmos, including physical experience, domestic scenarios, ritual scenarios, and world view. We proposed that a classifier organized like this be termed a *polycentric category*.

Shona noun class 3/4 grammaticizes and lexicalizes four scenarios and one ethno-ecological model which are salient themes of Shona culture. Scenario 3 was among the 11 previously defined. Three new ones include two new ritual scenarios (12, 14) and a domestic scenario (13). Item 15 is an ethno-ecological model.

[7] In spite of the earlier date of publication, Rader's paper followed Palmer and Arin (1999).

3.The spirits of ancestral chiefs bring rain, thunder, and lightning.
12.People pray to the ancestors.
13.Grain is pounded daily with a mortar and pestle.
14.Doctors cure with herbal medicines that are ground in a mortar and pestle.
15.Trees, shrubs, and herbs are associated with coolness, moisture, and medicine.

The conceptual elements provided by these models find lexical expression in many of the members of Shona class 3/4. Those lexemes in the class that do not predicate any of the major elements in the five models are semantically linked in various ways as described in Table 3. The more inclusive cognitive model of a noun class that emerges from inspection of the semantics of the lexical members and their associative links to the ethnographic models is what I refer to as a polycentric category. The general structure of such a category is summarized in Table 2 and diagrammed in Figure 4.

Table 1 The structure of a *polycentric* category: Shona class 3/4[5]

(1) *Multiple Central Models*: A class may be governed by one, two, or more salient cultural models and/or scenarios that are different from those governing other classes. The central models of Shona class 3/4 are:

The spirits of ancestral chiefs bring rain, thunder, and lightning.
People pray to the ancestors.
Grain is pounded daily with a mortar and pestle.
Doctors cure with herbal medicines that are ground in a mortar and pestle.
Trees, shrubs, and herbs are associated with coolness, moisture, and medicine.

[5] This table is based on the framework presented in Palmer and Woodman (1999). Principle (7) from that listing has been subsumed into principle (6). All examples are from Hannan (1984).

(2) *Multiple Prototypes*: A central model may be sufficiently complex to offer more than one prototype concept. For example, trees provide large poles and sticks, shrubs provide small poles and sticks. All provide medicinal leaves and fruits. The term for tree, *muti*, also means 'medicine.' Any of these items may therefore then serve as prototypes as explained.

The scenario of pounding grain with the pestle and mortar presents pounding, grinding, crushing, and grain as salient elements from which abstractions and extensions can be derived. The grain itself assumes the form of piles of grain, piles of finely ground meal, and scattered grains. These provide additional prototypes for spatial distribution of dry granular or powdery solids.

The ancestral scenarios of curing and rain-making offer component scenarios of propitiation of ancestors and grinding and giving of medicines. They also offer physical models of cool liquids. Lexemes for all these elements appear in Shona class 3/4. Examples: *mukwo* 'pestle', *mun* 'pestle', *muni* 'pestle', *mun* 'tree, medicine', *mudzvuku* 'tall, straight object (e.g. tree; skyscraper)', *mudzunura*, *mutsvuta* 'meal ground in *duri* (mortar)', *muchaka* 'meal from green mealies', *muchinjwa* 'mealie meal ground by engine-driven grinding mill', *muhwa* 'young, green mealie', *mudedi* 'green mealies', *muguri* 'mealie cob (with the grains on it)', *munyuchu* 'mealie-rice', *mubukirwa* 'green maize cob', *mudakunanzwi* 'sweet-tasting liquid', *mukamba* 'porridge made with milk as the liquid' *mujunju* 'any liquid thinner than it should be', *muchenga muchenga* 'abundance of grain', *muchenganherera* 'general rain <-*chenga*', *munyikamwe* 'springtime (beginning of rainy season)', *mutsvasatire* 'gusty rain', *muzhandwa* 'crops, animals or people struck down in large numbers, <-*zhanda*; act of crushing (e.g. as heavy object does when it falls)', *muchato* 'sound of footsteps, hoofbeats, etc.'.

(3) *Chaining of Central Models by Metonymy:* The themes that provide the backbone of a class are closely related, not by similarity, but by function or metonymy. For example, the pestle, a kind of stick or pole, provides the conceptual link from the originating model of trees, shrubs, and herbs to the scenario of pounding grain with a pestle: Medicines for curing are made from plant leaves and bark. One cures with herbal medicines, but also by appeal to ancestors who bring the rain associated with cool, moist forests and good plant cover. Examples*mukwerera* 'ceremony to pray for rain', *munamato* 'prayer (act of praying; words of prayer)', *musuuro* 'small pot of beer offered to husband to notify him that beer has been prepared and is now ready; amount of any prepared food or drink brought to head of family so that he may say the polite words of welcome to a guest, opening words of prayer to *midzimu [ancestor]*', *mukwerera* 'ceremony to pray for rain'.

(4) *Radial Categories:* Non-central items are linked and chained to central members by metonymy and metaphor. For example, witchcraft, which appears in this class, is a kind of pounding and crushing. Examples: *muchandwa* 'crops, animals or people struck down in large numbers [as by sickness]; act of crushing (e.g. as heavy object does when it falls)', *mupfuka* 'trampled grain or grass, peaceful place, case of witchcraft, fee for such a case', *muchapo* 'paddle, medicine for killing witches', *mushinhiriro* 'spell; act of bewitching'.

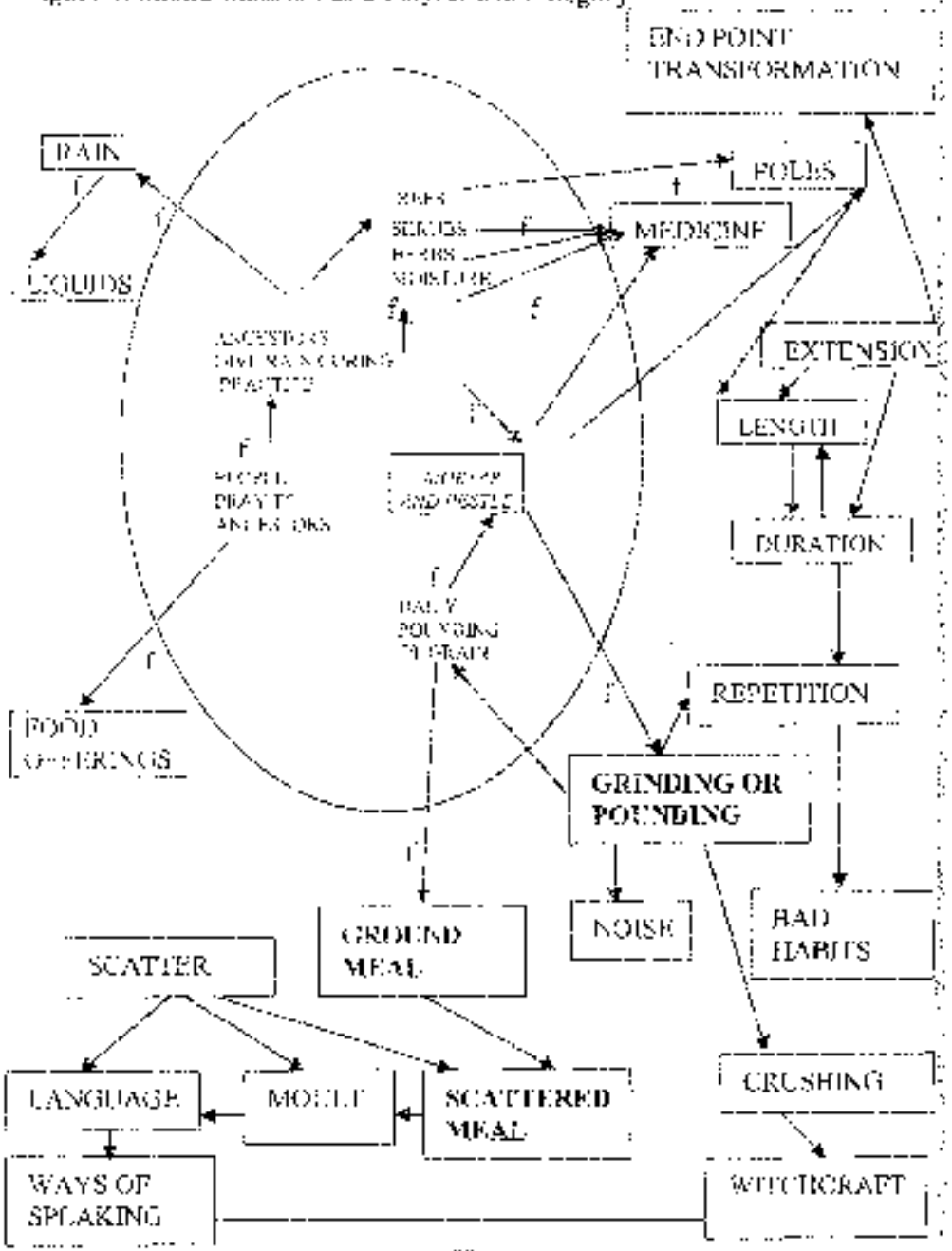
(5) *Primary Schematization:* Spatial and temporal schemas may be abstracted from any substantive concept. The pole or stick provides the abstraction of a solid cylinder or extended solid object. From pounding of the pestle it is an easy step to repetition, and to duration of time. Examples: *mudhodhodha* 'long object (e.g. low building, letter to someone); cursive writing', *mugashanya-gashanya* 'repetition of an action without interruption', *muchimbo* 'index finger, <-*chimba*',

mudhidhi 'penis (polite expr)', *mutambwi* 'time since', *mukanwa* 'period of time (gen the present)', *mukare* 'era, period of history'.

(6) *Secondary Schematization and Extension:* Spatial schemas are subject to various abstractions and extensions. The end-point transformation of an extended spatial object or time is a common extension, yielding ends of paths, beginnings, last times, and worn out objects. Examples: *muvambo* 'commencement, action of beginning', *mutangiro* 'beginning, way of beginning', *mugumo* 'end (of action, extent, etc)', *mufika* 'tapered end of axe or hoe blade', *mujumeeuno* 'last time, occasion, etc.', *mudemo* 'useless, worn-out axe'.

(7) *Extension of Concepts to Human Behavior:* The schema of repetition is extended to repetitive behaviors, mostly bad habits and propensities. Spatial and physical are extended. For example, in Shona, theft is a narrow passage between two objects. Language is a metaphorical scattering, the feathers of a moulting bird. Examples: *nuro* 'way of stealing', *mukoro* 'narrow passage between two objects; pass; act of stealing something in order to sell it, object stolen in order to be sold, act of stealing', *mutauro* 'language, discussion of a misdemeanour gen leading to legal case' < *ma* 'speak, moil', *muhwerekeiero* 'way of speaking', *mukafamwera* 'foolish, thoughtless way of speaking', *mukunya* 'peremptory, emphatic way of speaking', *muririro* 'call; characteristic cry or way of speaking'

Figure 4: **Shona Class 3/4 as a Polycentric Category**



A polycentric category has more complexity than a radial category, but it does not seem to display unnatural or excessive complexity for the semantic system of a natural spoken language. One expects people to have salient ideas based on rituals and daily domestic tasks, and it is essential for them to orient their environmental surroundings. It is natural to identify clusters of nouns that are functionally related and to regard them as a cultural unit. It is natural to abstract schemas from the elements of those models and to discover similarities and metaphors across conceptual domains. And it is natural to recursively apply such thought processes to the derived categories. Finally, it is natural for a lexeme to be polysemous within the sub-domains of a polycentric category. When such a complex is grammaticized, the categories are culture-specific and based on models that can be discovered by the procedures of ethnography, but also structured by mental processes that have been long described in the literature of cognitive linguistics.

This approach explains the numerous instances of noun words that appear to satisfy the criteria for more than one class but characteristically appear in only one class. The archetypal example in Bantu studies is the classification of chiefs with wild animals, rather than with humans (Creider 1975). Many terms do in fact satisfy the criteria for multiple classes, but they are judged by their speakers to fit one better than another. Each class has multiple criteria, and these may be activated by the context of a discourse. The selection and classification of a term is the product of multiple competing and synergistic activations. In Bantu, some nominal roots have more than one common classification. It is likely that some classifications are well entrenched, while others are more subject to reassignment.

This approach raises a question of boundaries. What are the boundaries between classes, if any? If every class has multiple criteria and nominal participants are sufficiently complex in their semantics to satisfy multiple criteria then classes will necessarily compete for members in an ecology of classification. In fact, there are no clear boundaries between classes. The overriding criterion is cultural salience, which varies with situations, but how can cultural salience be evaluated by the linguist? How can one predict which classifiers will be used with Bantu nominal roots? Currently, conclusions regarding the motivations for particular classifications are largely a matter of interpretation based on familiarity with the culture gained through participant observation or reading of ethnographies. One could devise tests that would manipulate the salience of criteria and observe the assignments of nominal participants to categories, but such tests may not

reproduce the motivations presented by naturally occurring discourse. Nevertheless, in the event that such tests are undertaken, two hypotheses are suggested:

(1) Reassignments will be more likely to occur where a domain that is inherent in both the semantics of the nominal root and in an alternative classifier is sufficiently evoked by the discourse situation.

(2) It will be more difficult to elicit reassignments to more entrenched category members, where entrenchment is independently measured by frequency of usage or stable assignment in natural discourse.

We must ask also how one can evaluate this analysis in comparison to other possibilities. Are there other analyses that would be just as convincing? Can our analysis predict which nouns will be classified together? Then, are a number of possible criteria which could be used to evaluate competing analyses. They do not entirely solve the problem of arriving at an analysis that is both replicable by others and true to native-speaker thinking because they remain subject to judgement and interpretation, but if taken seriously, I think they are better than having no criteria. The criteria are as follows:

(1) An analysis should be based upon thorough and comprehensive ethnography with attention to salient cultural scenarios.

(2) Given an adequate description of the cultural scenarios, an analysis should be plausible, that is, it should consist of obvious connections. Non-obvious connections may be adduced only where they are supported by native speaker attestations.

(3) A plausible analysis that is supported by native speaker attestation and reasoning is to be preferred over one that is not supported.

(4) A plausible analysis which explains the largest number of terms in a class is to be preferred.

(5) A plausible analysis of a classifier which excludes terms normally found in other classes is to be preferred, though even in a correct analysis many terms

will not be excluded, only preferred more strongly by their canonical classifier.

Finally, we must ask whether the cultural approach with polycentric categories can predict the emergence and structure of classifier systems cross-linguistically. The theory predicts that some kind of classifier system can emerge wherever there are salient and stable cultural practices and institutions. These are the necessary conditions. Certainly, many of the languages around the world have classifier systems, though some are hardly recognized as such. For example, the anatomical suffixes of the Salish languages are usually not regarded as constituting classifier systems, yet they function in much the same way as they take on abstract values of shape (Palmer 1996). Also marginal to our not on of noun classifiers are the click classifiers of the Khoisan and the verbal classifiers of Apache, but they have similar functions (Basso 1990, Berndtsen n.d.). One might even regard a finite paradigm of honorifics, as in Japanese or Korean, as a classifier system in the social domain. The approach does not currently specify the conditions that are *sufficient* to motivate the emergence of classifiers. Further cross-linguistic studies along these lines are needed.

4. Conclusions

Many lexical domains and grammatical constructions look directly or indirectly to significant cultural models, usually including scenarios and polycentric categories. Understanding the grammar and lexicon of a language requires grasp of cultural models and culturally defined imagery. The most appropriate term for this approach is *cultural linguistics*.

The perspective of cultural linguistics shows obvious utility compared to a more narrowly cognitive approach in its application to the problem of Bantu noun classifiers, where the use of ethnographic methods to identify salient cultural models and scenarios is a necessary step in the research. In this application, it was possible to show how cognitive processes of complex category formation and category chaining operate within culturally specific models to create the polycentric categories that we know as Shona noun classifiers. The polycentric category introduced by Palmer and Woodman (1999) has multiple central scenarios and prototypes, from which radiate category chains and complex categories as defined, respectively, by Taxonf

(1987) and Langacker (1987). More research is needed on other classifiers in the Shona system and more work of this kind is needed on other Bantu languages. More research of this kind is also needed on the many and varied classifier systems in other languages. Such work should be undertaken concurrently with ethnographic research on salient cultural scenarios and themes that may influence linguistic categories.

Department of Anthropology & Ethnic Studies
College of Liberal Arts
University of Nevada at Las Vegas

References

Basso, K 1990. *Western Apache Language and Culture: Essays in Linguistic Anthropology*. Tucson: University of Arizona Press.

Bernández, Enrique ed. Categorization through Phonetic Symbolism. Radial Categories Based on the Clicks in the *San* Languages. Unpublished ms. in possession of the author.

Contini-Morava, Ellen. 1994. Noun Classification in Swahili. Publications of the Institute for Advanced Technology in the Humanities, University of Virginia. Research Reports, Second series.

Cseider, Chet 1975. The Semantic System of Noun Classes in Proto-Bantu. *Anthropological Linguistics* 17:227-358.

Dixon, RMW 1982. *Where have all the Adjectives Gone?* Berlin: Walter de Gruyter.

Fortune, G. 1955. *An Analytical Grammar of SHONA*. London: Longmans, Green and Company.

Friedrich, Paul 1979. *Language, Context, and the Imagination: Essays by Paul Friedrich*. Stanford: Stanford University Press.

Guthrie, Malcolm 1967. *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages*. Amersham, England: Gregg Press, LTD.

Hannan, M. 1984. *Standard Shona Dictionary*. Revised. Harare, Zimbabwe: The College Press.

Hutchins, Edwin 1996. *Cognition in the Wild*. MIT Press.

Lakoff, George 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Lakoff, George 1988. Cognitive Semantics. In Umberto Eco, Marisa Santambrogio, and Patrizia Violi (eds): *Meaning and Mental Representations*. Bloomington and Indianapolis: Indiana University Press.

Lakoff, George & Mark Johnson 1980. *Metaphors We Live By*. Chicago/London: University of Chicago Press.

Lakoff, George 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.

Langacker, Ronald 1987. *Foundations of Cognitive Linguistics, Vol 1: Theoretical Prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald 1990. Subjectification. *Cognitive Linguistics* 1: 5-38.

Langacker, Ronald 1991a. *Foundations of Cognitive Linguistics, Vol 2: Descriptive Application*. Stanford: Stanford University Press.

Langacker, Ronald 1991b. *Concept, Image, and Symbol*. Berlin/New York: Mouton de Gruyter.

Langacker, Ronald 1999. Assessing the Cognitive Linguistic Enterprise. In *Cognitive Linguistics: Foundations, Scope, and Methodology*. Berlin and New York: Mouton de Gruyter.

Langacker, Ronald 2001. Discourse in Cognitive Grammar. *Cognitive Linguistics* 12,2:143-188.

Leakey, Louis SB 1959. *First Lessons in Kikuyu*. Nairobi: The Eagle Press.

Mylne, Tom 1995. Grammatical Category and World View: Western Colonization of the Dyirbal Language. *Cognitive Linguistics* 6.4:379-404.

Palmer, Gary 1996. *Toward a Theory of Cultural Linguistics*. Austin: University of Texas Press.

Palmer, Gary 1997. *Semantics: Primes and Universals*. New York: Oxford University Press. *Journal of Linguistic Anthropology* 10,2:279-284.

Palmer, Gary 2000. Review of Anna Wierzbicka, *Understanding Cultures Through their Key Words: English, Russian, Polish, German, and Japanese*. New York: Oxford University Press.

Palmer, Gary & Dorothea Neal Ard 1999. The Domain of Ancestral Spirits in Bantu Noun Classification. In Hiraga, Masako Chris Sinha & Sherman Wilcox (eds): *Cultural, Typological and Psychological Issues: Selected Papers of the Bi-annual ICLA Meeting in Albuquerque, July, 1995*. Amsterdam: John Benjamins.

Palmer, Gary & Claudia Woodman 1999. Ontological Classifiers as Polycentric Categories, as Seen in Shona Class 5 Nouns. In Martin Pütz & Marjolijn Verspoor (eds): *Explorations in Linguistic Relativity*. Amsterdam and Philadelphia: John Benjamins.

Palmer, Gary B. When Does Cognitive Linguistics Become Cultural? Case Studies in Tagalog Voice and Shona Noun Classifiers. In Luchjenbroers, A. (ed): *Cognitive Linguistics Investigations across Languages, Fields, and Philosophical Boundaries*. Amsterdam: John Benjamins.

Rader, Russell 1998. Life and Land-ownership: The Autochthonous Nature of Shona Noun Class 5 and 6. *California Anthropologist* 25:8-17.

Spitulnik, Debra A. 1987. *Semantic Superstructuring and Infrastructuring: Nonpast Class Struggle in Chiltenbo*. Bloomington, Indiana: Indiana University Linguistics Club.

Spitulnik, Debra A. 1989. Levels of Semantic Structuring in Bantu Noun Classification. In Bahuc, R. & P. Newman (eds): *Current Approaches to African Linguistics* 6. Dordrecht, The Netherlands: Foris.

Wierzbicka, Anna 1996. *Semantics: Primes and Universals*. New York: Oxford University Press.

Wierzbicka, Anna 1997. *Understanding Cultures Through Their Key Words: English, Russian, Polish, German, and Japanese*. New York: Oxford University Press.

To Be is an Answer; Not to Be is the Question: Action, Understanding, and Human Mediation

Ronald Miller

The first lesson we learn from Vygotsky is that psychological functions, in their mature state, present as knots, or tapestries, can only be understood if they are undone and the generative processes of their formation revealed. Vygotsky's distinctive contribution to our understanding is grounded in his concept of human mediation. The Vygotskian paradigm is often cast in opposition to or in contrast to that of Piaget and is often captured, if not caricatured, in the statement that '[t]he path from object to child and from child to object passes through another person' (Vygotsky 1978:30). The other person is a 'social other' who serves as a cultural guide and inducts the developing child into the surrounding social milieu. The emphasis Vygotsky places on the role of this mediator of the child's actions, beliefs, and understandings, is evident in his claim that is virtually synonymous with his name:

> Every function in the child's cultural development appear twice: first on the social level, and later, on the individual level; first *between* people (*interpsychological*), and then *inside* the child (*intrapsychological*) (Vygotsky 1978:57).

But to understand the psychological implications of that distinctive form of human action that Vygotsky called mediation, we need to penetrate into its functional structure, or way of working. The social other who guides, shapes, informs, and regulates the child's activities is an intentional agent and one possible point of entry is to begin by attempting to unravel the concept of

agency in order to understand the conditions of its genesis

A theory of mediation entails an account not only of the mediator as a purposeful or intentional agent but also of the recipient of mediation, the *mediatee*, for whom mediation is experienced as a happening that is not of his/her own making. In the same way that Bakhtin (1981:293) argues that 'The word in language is half someone else's' so mediated actions are half someone or something else's. Wertsch (1991:32,38) refers to 'The agent of mediated action' and to 'agency as mediated' and, in this mediated sense, agency represents a partnership in which ownership of action is distributed unevenly between the actors. But in using the term 'mediated' as a modifier of agent or agency, Wertsch seems to miss or omit the psychological significance of mediation. Mediated actions are not performed by *agents* and to the extent that action is mediated, agency is undermined. Wertsch's (1991:12) argument that 'mediational means shapes the action in essential ways' obscures the point that it is not the means as such (tools, language) but the effects produced by the means, the happenings brought about in conjunction with an enabling or constraining intermediary that impact on the mediatee[1]. The fact that tools (material and symbolic) mediate between actions and their effects' consequences and that they prescribe or constrain the kinds of actions that can be performed, places the mediatee in the role of an actor whose actions are regulated from without. In this sense, agency is the antithesis of mediation whose central psychological feature entails a surrender of agency by the mediatee and submission or subordination to the *other*.

The concepts of action and agency conflate doings and happenings in the person of an agent who is both the subject of doings and the object of happenings. This serves to obscure the formative psychological processes that generate the condition of agency. By treating the same actor who performs, and for whom the performance is enacted, as an *agent*, we lose sight of the psychological processes that make the condition of agency possible. An action is something that not only is done, a doing, but is also something that happens. A cup of coffee may *happen* to me through the agency of considerate others or as a result of my own actions. We typically distinguish between these two conditions by using the grammatical terms *I* and *me* and these terms can be

[1] This is the point of the distinction Newman and Holzman (1993:37-41) make between tool-for-result and tool-and-result.

mapped onto the psychological categories of *ego* and *self*. Harré (1998:179) contends that although the concept of self is a useful fiction, it has 'generated an ocean of metaphysics' and that '[b]y paying attention to the forms of expression of the sense of self we have condensed this ocean into a drop a grammar'. The terms *ego* and *self* do not negate or dilute Harré's assistance that the basic ontological parameters in the human world are persons but whether they amount to mere grammatical fictions is more contentious[2]. William James (1962-1891) commented that,

> the total self of me, being as it were duplex, partly known and partly knower, partly object and partly subject, must have two aspects discriminated in it, of which for shortness we may call one the *Me* and the other the *I*.

These categories correspond to Vygotsky's (1987:250) distinction between 'perceptive consciousness' and 'intellectual consciousness', what James (1962:26) called 'knowledge of Acquaintance' and 'knowledge About', and also to the contrast Ingold (1986:112) draws between a practical or presentational awareness *of* doing, and a conceptual or representational consciousness that saturates happenings with an understanding of what they are *about*. This distinction between an *awareness of* and a *consciousness about*[3] reflects two different aspects or qualities of our experience and suggests a bi-polar or, to use James' expression, 'duplex' conception of our subjectivity that can be captured in the terms *ego* and *self*. The *ego* is the aware subject of purposeful actions that begin with reasons and end with the realization of goals. In contrast or as a complement to the *ego*, the *self* is the

[2] The term 'psychological categories' is used to indicate that *ego* and *self* do not refer to entities but to kinds of events. As theoretical constructs they are analogous (but not necessarily equivalent) to the concepts of phenotype and genotype. As Flanagan (1994:205) points out 'The idea that the self is a fiction is compatible with its being real and its playing a functional role in an individual's psychological economy and social life'.

[3] For a comprehensive critique of Harré, see Archer (2000:86-111) and for an alternative approach see Dennett (1993:412-430).

[4] See Dennett's (1997:25-73) comments about intentionality and what he calls the intentional stance.

conscious predicative aspect of subjectivity', the locus or centre of understanding without which reasons are reduced to motives and actions to behaviours. In some respects, this conception of ego is similar to what Harré (1998) identifies as self1 in his triplex of selves that together represent aspects of what we mean (or should mean) by the term persons. According to Harré (1998:3), self1 (in contrast to self2 and self3) is a locus or 'site from which a person perceives the world and a place from which to act'. But Harré's self2 (a person's unique set of attributes) and self3 (the impression a person makes on others) that are anchored to self1 (rather like a camera mounted on a mobile tripod) seem to neglect or omit a crucial part of what it means to be a person. In discussing what he calls 'the most essential difference between man and animal', Ortega (1957:141ff) alerts us to an aspect of our subjectivity that seems anterior to Harré's triplet of selves.

> Observe that this marvelous faculty that man possesses of temporarily freeing himself from his slavery to things implies two very different powers: one is his ability to ignore the world for a greater or lesser time without fatal risk; the other is his having somewhere to take his stand, to be, when he has virtually left the world ... But the world is the whole of exteriority, the absolute *without*, which can have no other *without* beyond itself. The only possible without to this *without* is, precisely, a *within*, an *intus*, the inwardness of man, his *self*, which is principally made up of ideas.

This conception of *self*, not as a site from which to perceive and act on the world without but as an inner site or locus of understanding, is crucial in the context of a theory of mediation. All forms of mediation consist of directives for the mediatee to act in particular ways but the actions produced by the mediatee, as a consequence of other-regulation, have their origin in the

[1] Discussing inner speech, Vygotsky (1987:248,243) points out that there is a 'preponderance of sense over meaning' and that predication is its natural form: 'psychologically, it consists of predicates only'. Another feature of inner speech is that it lacks what Bakhtin (1986:99) calls 'addressivity'. Rather than an utterance that is spoken, inner speech has the quality of a voice that is heard. When we talk to ourselves, we listen rather than speak. We hear ourselves thinking and in this sense inner speech is more a happening than a doing.

understanding that constitutes the *self* of the mediator and it is this understanding that grounds the mediator's intentions (or her reasons for acting in a particular way). In a novel situation in which new understanding must be acquired, the mediatee is of course aware of the actions she/he performs but not of what the actions are about. For example, consider a situation in which a child who has never before encountered a pair of scissors is shown how to use them by indicating to the child where to insert the thumb and finger and how to place the blades in relation to the cutting surface. In this *mediated scenario* the obviousness of scissors, their cutting function, is then revealed as a happening or performance that emerges out of the child's actions. But, for the child as mediatee, the understanding that the scissors is about a way of cutting is not initially constitutive of that action. In the sense of other-regulation, mediation then involves the interception of the *self* of the mediatee, that is the pre-understandings that constitute the consciousness of the mediatee, and the substituting of the *self* (or understanding) of the mediator. This ontological notion of *pre-understanding* is central to an analysis of mediated learning. Gadamer used the terms *prejudice* and *tradition* to express the taken-for-granted quality of the (pre-)understanding that we impose on our experience.

> Long before we understand ourselves through the process of self-examination, we understand ourselves in a self-evident way in our family, society and state in which we live. The focus of subjectivity is a distorting mirror. The self-awareness of the individual is only a flickering in the closed circuits of historical life. That is why the prejudices of the individual, far more than his judgments, constitute the historical reality of his being (Gadamer 1975:245).

Ortega (1960:49) expressed much the same idea using the terms *habituality* and *blindness*.

> These instances of blindness vary from one period to another but they are never absent ... Every idea is thought, every picture painted, out of certain assumptions or conventions which are so basic, so firmly fixed for the one who thought the idea or who painted the picture that he neither pays heed to them, nor, for that matter, introduces them into his picture or his idea; nor do we find them there in any guise except as pre-supposed and left, as it were, at one side. This is why we

sometimes fail to understand an idea or a picture; we lack the clue to the enigma, the key to the secret convention.

In a mediated learning[4] situation, or to use Vygotsky's catch-phrase of the zone of proximal development, the mediator and mediatee do not share (by definition) a common understanding of the situation. It is for this reason that mediated learning cannot be modelled on dialogue or the dynamics that govern a conversation between interlocutors. Mediation and dialogue are *different* kinds of communication that serve different purposes. In dialogue, meanings or understandings are exchanged whereas mediation is concerned with the learning and teaching of new understanding in situations where prior understanding is inadequate. As Volochinov (1973:102) points out: 'Any true understanding is dialogic in nature. Understanding is to utterance as one line of dialogue is to the next'. But in situations of mis-understanding or misunderstanding dialogue breaks down and a different form of communication is required to restore what Bakhtin (1981:275) calls the 'primordial dialogism of discourse'.

Drawing on the idea of discourse as 'language-event', Ricoeur (1983:197-221) proposes that text can serve as a model for the interpretation of the kinds of 'meaningful actions' that we encounter in the human sciences. At the core of Ricoeur's model is his concept of *appropriation* that is the counterpart of the distanciation that is inherent in the nature of text that mediates between author and reader. The critical point that Ricoeur[5] makes is that 'appropriation does not imply any direct congeniality of one soul with another' and that '[n]othing is less inter-subjective or dialogical than the

[4] The term 'mediated learning' compounds two problems. The acid test for a theory of learning is that it must address the problem of novelty and avoid the learning (Meno) paradox such that existing understanding cannot be the basis for new understanding. (The common and persistently fashionable wisdom that learning should proceed from the known/familiar to the unknown/unfamiliar seems to fall foul of this paradox.) The bottom line for a theory of mediation is that it must avoid an infinite regress of teachers' teachers. (As an explanatory concept, imitation, in its various guises including 'scaffolding', seems to fall headlong into this trap.)

[5] In stark contrast to Ricoeur's model of the text, Harré (1998: 45) argues that 'A useful model for skilled action, that is action that is intentional and normatively constrained, is conversation'.

encounter with a text' (Ricoeur 1983:191). In a novel situation, the task for the mediatee is to appropriate new understanding that dislodges existing pre-understanding whereas for the mediator, the task is to entrench her understanding by disseminating it through what is essentially a monologue, in the sense that instruction does not require or invite an exchange of ideas. The effect of mediation, then, is to reduce an agent to an actor whose other-regulated actions are now experienced as happenings by the *self*. Ricoeur uses the terms 'relinquishment' and 'letting go' to capture what he calls a 'fundamental moment of appropriation' in which 'the ego divests itself of itself' (Ricoeur 1983:191). In this sense, the acquisition of understanding that scissors are about cutting is not something that the actor *does* with scissors but something that *happens* to the actor (or that scissors do to the actor) in the course of acting in conjunction with the scissors. In performing a set of actions, the child or novice is *aware* of the situation, what she is doing, what the scissors are doing, and her part in that doing. But this awareness of doing is a different experience from that construed by a prior understanding of the aboutness of scissors[6]. Once that understanding is acquired, the awareness of the scissor-object is transformed into a consciousness about scissor-as-cutting-object.

The quality of experience that we call consciousness is a property of the *self* whose mode of being is understanding[7]. This is not to suggest that the *self* is the complement of the Cartesian Ego that presides over our acquired understandings that increase with experience. On the contrary, it is our understanding that constitutes our sense of self or selfhood, and about which we are conscious. To be conscious of happenings is to understand and understanding is not something we do but something that happens to us.

[6] The aboutness and awareness of scissors corresponds to the distinction Heidegger (1980:102-114) makes between ready-to-hand and present-to-hand.

[7] Although Heidegger is not addressed explicitly in the text, the concept of self that is elaborated is not far removed from that of *Dasein*. Although not by design, *Being and Time* seems to provide a common thread that is woven into the ideas of a number of the authors whose ideas have informed the text. Ortega claims provenance for some of Heidegger's ideas; Gadamer and Ricoeur both acknowledge their debt to Heidegger; Clarke's book is deliberately entitled *Being There*; and Harré acknowledges the more or less equivalence between Heidegger's thrown into a life-world and Wittgenstein's forms of life that provide the springboard for his analysis.

however, all understanding is mediated by action without which there are no happenings[9]. Agency, then, is a form of internal mediation in which the *self* or the understandings that ground reasons for action, is also the *self* at whom actions are ultimately directed. This circle of understanding preserves the identity or being of the *self* and can only be disturbed by actions that are not given by self-understanding, the consciousness of the *self*, but by actions that are other regulated. The sources of this other regulation are either biological mechanisms that regulate the course and direction of action (and here space permits only a nod in the direction of Piaget's explication of the logic or psycho-logic of action); a *self* or consciousness that is *other* to the constraints and enabling conditions imposed by the otherness of a novel situation, of which the mediated is a constituent part. This conception of a situation as an occasion or enticement to action is different from the notion of a situation as a setting that is independent of the learner and into which the learner may step or be thrust. For example, Wertsch (1991:15) states that 'I use the term sociocultural because I want to understand how mental action is situated in cultural, historical, and institutional settings'. By posing the problem in this way, Wertsch seems to imply that mental actions and settings are separate or independent entities such that one can be situated in the other[10]. He argues that Vygotsky 'did little to spell out how specific historical, cultural, and institutional settings are tied to various forms of mediated action' and that by focusing 'on small group interaction, especially the interaction of the adult-child dyad' he failed to 'deal with broader sociocultural issues' (Wertsch 1991:46). Again, Wertsch's argument reflects the view that mediated action and broader sociocultural issues are separate entities that need to be tied together. The point that Wertsch seems to miss is that the adult-child dyad (teacher-learner) is precisely the place where culture, history, tradition, and the institutions of social life transform actions into the 'mediated actions'

[9] The tension in this dialectic of doings and happenings is captured in Dennett's (1993:418) account of a *narrative self* that is constituted by the stories we tell about who we are. 'Our tales are spun, but for the most part we don't spin them; they spin us. Our human consciousness, and our narrative selfhood, is their product, not their source'.

[10] In contrast, Clark (1997:7,33) refers to 'embodied and environmentally embedded agents—beings that move and act upon their world' such that 'Brain, body, world, and artifact are discovered locked together in the most complex of conspiracies'.

that constitute the broader sociocultural issues and settings. An alternative, if not opposite, interpretation of a sociocultural approach is that to the extent that culture, history, and institutions are settings at all, they are 'located' in mental actions in the sense that they are constitutive of such actions.

The constraints that are inherent in a situation are revealed when a situation is transformed into a task whose solution requires the transformation of prior or pre-understanding into new understanding. To understand a situation is to understand how it works; and how it works is what we do when confronted with new and novel objects and events. What we do, or can do, is determined by the constraints and enablements imposed by the situation: the limits set by the horizons of the inner world (*Welt*) of our understanding and outer world (*Umwelt*) of objects and events. Both Ortega and Gadamer employ the concept of 'horizon' in relation to the understanding of a situation.

> Environment, then, is the patent or semipatent world around us. But in addition to this, beyond our horizon and our environment the world at any particular moment contains a latent immensity made up of pure compresences: an immensity that, in each situation of ours, is a hidden, eclipsed immensity, concealed by our environment and enveloping it .... Yet in this state of latency and eclipse, they act on our life as habituality .... The horizon is the dividing line between the part of the world that is patent and the part of it that is latent (Ortega 1957:67).

> Understanding of the past, then, undoubtedly requires an historical horizon. But it is not the case that we acquire this horizon by placing ourselves within a historical situation. Rather, we must always already have a horizon in order to be able to place ourselves within a situation. For what do we mean by 'placing ourselves' in a situation? Certainly not just disregarding ourselves. This is necessary, of course, in that we must imagine the other situation. But into this other situation we must also bring ourselves. Only this fulfils the meaning of 'placing ourselves'. If we place ourselves in the position of someone else, for example, then we shall understand him, i.e. become aware of the otherness, the indissoluble individuality of the other person, by placing ourselves in his position (Gadamer 1975:271).

On the one hand, we are constrained by our prior conceptions or pre-understandings of and about the world. On the other hand, the objects and events we experience also place constraints on what can be done to and with them in terms of their properties. A scissor is constrained in the way it works. Unlike a knife, its blades are concealed and will only cut when the blades are exposed. But a pair of scissors can be used as a dagger or a punch, by a person for whom its cutting function is opaque. However, if its cutting function remains concealed to its user, then it is profoundly misunderstood; and to the extent that it is used as a dagger (in the absence of understanding its cutting function), this imposed pre-understanding prevents or inhibits the potential understanding that is inherent in the design of the tool: the aboutness that it designates in its working. In this sense, a scissor is ambiguous in its functional structure: its working or design.

Like all tools (including, Vygotsky's symbolic tools that encompass speech and language in general), a scissor has the potential to expand or open up new horizons, to use Gadamer's (1975:269) metaphor or, in Ricoeur's (1983:182-191) terms, to reveal a new world: a world of cutting, of manufacture, of capital accumulation, and so on. But for this to happen, the scissor must pose a question and, thereby, present itself as a task whose functional structure or design[1] requires to be understood. The design of a situation is revealed in the working of the structure in which the constraints (and enablements) of the inner and outer world together serve to regulate the actions that transform a situation or experience into a task. This intersection of the inner and outer worlds, or what Gadamer calls the *fusion of horizons*, constitutes the situation. For this to happen, the situation must pose a question or the experience produce a disjunction, in the sense that habitual ways of acting on and with the objects in question are obstructed. If a scissor is used as a stabbing instrument, the questions it poses are why two daggers should be joined together through a central pivot and why the handle should consist of two large protruding hollow discs that interfere with its effective use as a weapon. A task, then, requires the surrender of the pre-understandings that obstruct or inhibit those actions whose doing would serve to reveal the design

[1] The distinction between structure and functional structure or design is captured by Bakhtin (1990:267) who distinguishes between *architectonics* or structure of the aesthetic object and its composition. "The structure of a work, understood as a structure that actualizes the aesthetic object, we shall call the *composition* of a work".

that lies concealed in the structure of the situation. In this sense, the demands of a task, the questions posed by a situation, are the counterpart of the constraints of a situation.

The difficulty with the concept of situation is that it does not stand still. A situation is the beginning and end of a task in the sense that the situation, like the horizon, moves with us as we engage in the task of understanding and persists until understanding is achieved. Because understanding signals the end of a task, the achievement of understanding is also the dissolution of the situation. A situation that does not pose a question or require a solution ceases to be a situation in the sense of an object of understanding. Tying our shoelaces is not a task that we engage in but an activity that is performed routinely without effort or thought. This is not the case for a young child for whom tying shoelaces is a complex task; a problem occasioned by shoes and their untied laces and a question whose answer is given by tied shoelaces. We experience a set of events as a task when there is a need to understand; when sense and reference become unstuck; when the inner and outer horizons of consciousness are dislocated and resist the harmonious coupling of horizons within which pre-understandings are contained. These prior understandings, or prejudices to use Gadamer's term, are what we carry with us, not as baggage that can be discarded, or lost and later recovered, but as part of our being or selfhood. When understanding is achieved, even after a period of intense effort, it has a quality of being self evident. New understanding does not change the set of objects or events that constitute a task, such as the configurations of blocks, beads, and water levels characteristic of Piagetian conservation tasks, but transforms the *self* by incorporating the situation into its being. Younger and older children understand the 'conservation' situation differently but in saying this we are saying that they are conscious of different situations. For the young child, the events are about different quantities of liquid whereas for the older child they are about the same amounts. In both cases, the relation between experience and situation is one of understanding and, in this sense, what distinguishes the children is their being. The element of difference in the two situations does not reside in the events of the external world but in the internal world that constitutes the consciousness of the children. Understanding is not a skill that improves with practice, or a bit of information that may be forgotten or added to a store of similar bits of information. The *self* that is transformed is not a bloated version of its former *self*. New understanding demands a negation and not an elaboration or refinement of previous understanding. This inner world

of understanding is the product of negating transformations of the *self* whose being is prior to any situation or task. If understanding is considered an incomplete work, then for any given situation, the *self* is in a potential relation of pre-understanding *vis-à-vis* new situations.

The task inherent in every situation that begs a question is to overcome pre-understandings that attach inappropriate questions to the answers that constitute the situation. That understanding is nourished by prior or pre-understanding is the foundation of the hermeneutic circle. The problem for any analysis of understanding, how it is possible and how it is achieved, is to penetrate the circle by finding a point of entry. Mediation as a form of other regulation is a means of penetrating the circle of understanding by furnishing the mediatee not with ready-made understanding but with alternative pre-understandings that facilitate rather than obstruct engagement with the task at hand. The 'otherness' that defines all mediation is not merely another person, analogous to an interlocutor in a dialogue but a horizon or world view: a perspective of a world constituted by situations that are alien to the mediatee. To enter another world, that of the mediator, the mediatee must experience the situations that constitute the life of the mediator by addressing the questions whose answers construct the situations that the mediatee encounters, situations whose 'otherness' is not the individual other person but the world(s) of the other.

> Understanding is not to be thought of so much as an action of one's subjectivity, but as the placing of oneself within a process of tradition, in which past and present are constantly fused ... we do not try to recapture the author's attitude of mind but, if this is the terminology we are to use, we try to recapture the perspective within which he has formed his views (Gadamer 1975:258f).

From the perspective of the *self*, new understanding is experienced as a revelation, a happening that is brought about by actions whose source or generative power derives from a consciousness that is other. But the otherness of the consciousness that flows through mediation is not necessarily the intentionality that drives the actions of the other as mediator. The mediator's reasons for regulating or directing the actions of the mediatee, or the purpose of the mediation, is not the goal of understanding of the mediatee. As Ricoeur points out, understanding of a text does not, and should not, require an

understanding of the author's intentions behind a text but of a world that is revealed in front of the text.

> Therefore what we want to understand is not something hidden behind the text, but something disclosed in front of it ... Understanding has less than ever to do with the author and his situation. It wants to grasp the proposed worlds opened up by the references of the text. To understand a text is to follow its movement from sense to reference, from what it says, to what it talks about (Ricoeur 1983:218).

To ask a question is already to impose understanding on a situation. We do not ask, cannot ask, questions about matters we do not understand. The task of education or understanding, in Vygotsky's (1978:89) sense of learning that runs ahead of development, is to trace the path that leads from answers *backwards* to questions. In this sense, we can distinguish between training and education. Training is a process of providing answers to questions whereas education requires that the actions of the mediatee or learner must reveal what the situation is about; the question to which the situation is an answer. In moving from answer to question, situations are transformed into tasks, the essential characteristic of which is the surrender of agency or submission to the design inherent in the situation. The task demands that the actor produce a performance, the happening of which is the revealing of the design, in which the actor plays a part, or as Gadamer (1975:95f) suggests in respect of play, in which the actor is played with or constrained to perform in ways prescribed by the situation of the game.

> ... ad playing is a being played. The attraction of a game, the fascination it exerts, consists precisely in the fact that the game tends to master the players .... Whoever 'tries' is in fact the one who is tried. The real subject of the game ... is not the player, but instead the game itself. The game is what holds the player in its spell, draws him into play, and keeps him there.

The essential difference between the activities that constitute play and those of a task, is that play does not pose a question or culminate in the transformation of the consciousness of the *self*. In play, the situation of the game is re-created and the design of the game, it's playing, is an expression of being in which the *self* marks time.

Thus the child gives itself a task in playing with the ball, and such tasks are playful ones, because 'the purpose of the game is not really the solution of the task, but the ordering and shaping of the movement of the game itself' (Gadamer 1975:91).

The situation of the game, or the experience of playing, is not negated by play as is the case of a task. Play is never accomplished in the way that the mastering of a task marks its termination. In play, becoming is suspended as creation is overtaken by re-creation. Unlike a task, whose resolution is its ending, in play we return again and again to the game whose playfulness consists in the infinite variety of possible answers to the same question. 'The movement which is play has no goal which brings it to an end; rather it renews itself in constant repeat on' (Gadamer 1975:93).

A task is not a certain kind of play and play is not a frivolous kind of task. To the extent that a game presents a situation that requires understanding, a potential player or novice will engage in the task of becoming a player. Once this is achieved, however, the actions required by the game are performed in the same repetitive and routine manner as any other actions. Whereas a task culminates in the negation of the experience of a situation, such that the situation is incorporated into the being of the *self*, in play, the *self* is drawn into the game, a situation whose design is the re-creation of experience with each playing of the game. Recreation, then, is a re-affirmation of the *self* in contrast to the transformation of the *self* that is achieved by the acquisition of new understanding. In engaging in a task, the performance that is enacted entails a negation of the pre-understanding that the learner brings to the task, such that the new understanding dislodges the old and to the extent that the new takes hold and is grasped, it serves to displace the old. To treat a situation as a game is to retain a degree of detachment from the situation by substituting for the task, routine actions that sustain the situation. To play the system, to name the game, to go through the motions, are all ways of maintaining the identity and being of the *self* as against engagement in a task whose realization is the negation of the *self* as a moment in its transformation from one state of being to another and in the transformation of the situation from subject to object of understanding.

Unlike the achievement of understanding, accomplishment in play depends on the improvement through practice of the routine actions that constitute the game. Clearly, in Piagetian terms, the achievement of conservation by concrete operational children cannot depend on the practice and perfection of non-conservation. It is in this sense that skills must be

distinguished from understanding. Whether we speak of the artisan, sportsperson, actor, or writer, it is not understanding that is made perfect by practice but actions whose performance not only depends on prior understandings but is generated by them. Horses and athletes can and are trained by trainers who cannot themselves achieve the level of skill they are attempting to improve in their trainees. In contrast, teaching becomes redundant when the teacher's level of understanding is matched by the student. This is why the best students outgrow their teachers or convert them to trainers who can coach them to improve their techniques but not their understanding. In this sense, the ends of education justify its means whereas the means of training justify its ends. Practice, in the sense of training, may strive for perfection but it is the antithesis of negation; a celebration of habit and stability in the face of innovation and transformation. The pleasure we derive from a game is that it presents a challenge to our skill, or ego, but not to our being or *self*, and the satisfaction we derive from improving our performance in play is self-satisfaction. The effort and energy that is expended on the activities in which we engage for relaxation, serve to affirm the *self* by entrenching its way of being.

It is part of play that the movement is not only without goal or purpose but also without effort. It happens, as it were, by itself. The ease of play, which naturally does not mean that there is any real absence of effort, for phenomenologically refers only to absence of strain, is experienced subjectively as relaxation. The structure of play absorbs the player into itself, and thus takes him from the burden of the initiative, which constitutes the actual strain of existence. This is also seen in the spontaneous tendency to repetition that emerges in the player and in the constant self-renewal of play, which influences its form (Gadamer 1975:94)

The excitement that surrounds sporting contests, the compulsion to pursue 'who-dun-it' stories, and the addiction induced by Soap Operas, lies in the suspense of the answer that is given at the end of the game, in the concluding chapter of the book, or the next episode in an infinite series of television banality. But no mystery attaches to the questions, all of which are given in advance by the genre of which they are an expression. Even when we do not participate as players we are drawn into the game as spectators. In this situation, of which we are a part, we may take delight in the skill of the players or be disappointed by their ineptitude. But in the event, the game, each time it is played, is a reaffirmation of the inner world of our being, of the way in which we understand the world and our place in that world. Entertainment

is essentially a situation in which our mode of participation is that of spectator. The spectacle that unfolds does not pose a problem or impose itself as a task. On the contrary, entertainment is predicated on prior understanding and dedicated to its preservation. This is why entertainment is often contrasted with education and, perhaps more significantly, why learning should not be confused, or confuse itself, with entertainment.

It is as spectators and not as players that we experience the world of art in all its expressive forms. This does not mean that as spectators we stand apart or detached from works of art. Every work of art is a potential situation, an open invitation, not only to participate in a world whose horizons are contained within the work, but to enliven that world. As a product of work, art is an answer the understanding of which does not lie in an explanation of its production but in the questions that its expressive power evokes. Whereas games are performed primarily for the player, art is performed primarily for the spectator. Gymnasts perform for themselves as do dancers at a party, unlike ballet dancers whose performances are directed at an audience. In this sense, it is the spectators who play or are played with by the artistic performance. For the performers, the performance is work as is the creation of the artistic work whose destiny is to find expression in performance. Artistic expression is part task part play embodying both creation and re-creation. In its re-creative aspect, art most resembles play when viewed from the perspective of the spectator or audience. It is also in its re-creation that art endures and lives as part of tradition and it is this power of endurance that we attribute to great works of art. But what is being done or happening when we attend performances of classic works with which we are thoroughly familiar? Wherein lies the excitement and pleasure in experiencing again and again a Beethoven symphony, a production of Hamlet, a Van Gogh landscape?

Every performance has a festive air and is a celebration that culminates in applause, itself an expression of joy and appreciation. Clearly, the audience's applause is most directly addressed to the performers' skill. But in applauding the performers we celebrate more than the performance. Artistic works are not re-produced in their performance but re-created and we celebrate their creation in our applause. But in applauding the work and genius of its creator, we again reaffirm the tradition of which it is a part and which is part of us and, in so doing, set the stage for its next performance. But the act of creation that we celebrate with each performance of a work of art, in contrast to the act of production, is not the playing of a game. It is the transformation of our being that occurs when understanding shifts the horizon

of our inner world. In its creative aspect, art shows itself as an enchanted task that playfully teases our understanding, enticing the *self* to surrender to its charm. Once understood, our aesthetic experience cannot be undone or set aside and a special effort, that requires understanding, is needed to appreciate unfamiliar forms of artistic expression in the same spontaneous, self-evident, and intuitive way that we are drawn to the familiar. Aesthetic intuition is the expression of a tradition of art that culminates in the consciousness that confronts new forms of art as pre-understanding.

In the same way that the truth or meaning of a score is revealed in its design, an artistic work reveals its meaning in its performance which is an expression of a truth that endures and is perpetuated with each performance. But the truth or meaning of a work that is realized in performance, does not reside in the work as a part, aspect, or attribute of its constitution. Truth and meaning are properties of consciousness, attributes of understanding that ground the inner world of the *self*. Understanding is the stuff of consciousness, or to use the more dynamic idiom of James, consciousness is the stream through which understanding flows and constitutes the inner world of our being: a world that lives in us as distinct from the world in which we live. To be in that world is an answer; not to be is the question.

School of Psychology
University of Natal

References

Archer, MS 2000. *Being Human. The Problem of Agency*. Cambridge: Cambridge University Press.

Bakhtin, MM 1981. *The Dialogic Imagination: Four Essays by M.M. Bakhtin*. Holquist, Michael (ed). Emerson, Caryl & Michael Holquist (trans). Austin: University of Texas Press.

Bakhtin, MM 1986. *Speech Genres and Other Late Essays*. Emerson, Caryl & Michael Holquist (eds). McGee VW (trans). Austin: University of Texas Press.

Bakhtin, MM 1990. *Art and Answerability: Early Philosophical Essays by M.M. Bakhtin*. Holquist, Michael & Vadim Liapunov (ed). Liapunov, Vadim (trans). Austin: University of Texas Press.

Clark, A 2001. *Being There: Putting Brain, Body, and World Together again*. Cambridge: MIT Press.

Dennett, DC 1993. *Consciousness Explained*. London: Penguin.
Dennett, DC 1997. *Kinds of Minds: Towards an Understanding of Consciousness*. London: Phoenix.
Flanagan, O 1994. *Consciousness Reconsidered*. Cambridge: MIT Press.
Gadamer, H-G. 1975. *Truth and Method*. New York: Crossroad.
Harré, R 1998. *The Singular Self*. London: Sage Publications Ltd.
Heidegger, M 1980. *Being and Time*. Oxford: Basil Blackwell.
Ingold, T 1986. *Evolution and Social Life*. Cambridge: Cambridge University Press.
James, W 1962. *Psychology: Briefer Course*. New York: Collier Books.
Newman, F & L Holzman 1993. *Lev Vygotsky: Revolutionary Scientist*. New York: Routledge.
Ortega Y Gasset, J 1957. *Man and People*. New York: Norton.
Ortega Y Gasset, J 1960. *What is Philosophy?* New York: Norton.
Ricoeur, P 1983. *Hermeneutics and the Human Sciences*. Cambridge: Cambridge University Press.
Voloshinov, VN 1973. *Marxism and the Philosophy of Language*. Matejka, L. & IR Titunik (trans). New York: Seminar Press.
Vygotsky, LS 1979. *Mind in Society*. Cambridge: Harvard University Press.
Vygotsky, LS 1987. *Thought and Language*. Cambridge: MIT Press.
Wertsch, JV 1991. *Voices of the Mind: A Sociocultural Approach to Mediated Action*. London: Harvester.

Real Patterns and Distributed Cognition[1]

Andrew Detlis and David Spurrett

Introduction

Cognitive science has always been multi-disciplinary[2], growing as it did on of a recognition that cognition as a general object of study demanded expertise from a variety of fields including psychology, linguistics, computer science and philosophy. More recently, of course, biology has come to play an increasingly important role in this disciplinary mixture[3]. For all the

[1] This paper grew out of research conducted by Detlis during the first half of 2001, and by Spurrett in 2000 and 2001. Parts of Spurrett's contribution here were presented at the April 2001 conference in philosophy of science at the Inter University Centre, Dubrovnik, Croatia, his attendance at which was partly funded by the NRF, whose assistance is hereby acknowledged.

[2] Bechtel, Abrahamsen and Graham open their account of 'The Life of Cognitive Science' in Bechtel and Graham (1998) with the following definition: 'Cognitive science is the multidisciplinary scientific study of cognition and its role in intelligent agency. It examines what cognition is, what it does, and how it works. We don't entirely agree with this definition (see the following note) but do endorse the prominence given to the multidisciplinary character of the field.

[3] In the account by Bechtel, Abrahamsen and Graham (1998:94) relatively little emphasis is given to biology, which appears in none of their diagrams of the various disciplinary contributions to the field. Nonetheless the extensive references to biological studies in, *inter alia*, works by Brooks (e.g. 1991a; 1991b; 1997), Clark (especially 1997), and Thelen and Smith (1994) speak for themselves.

advantages brought by such a melting pot, some issues and questions can fall between the interdisciplinary cracks, and it is one of these that we want to consider here. Working scientists are not typically concerned with metaphysical questions[3], but their efforts nonetheless often raise such questions, or suggest new ways of their being framed. Metaphysicians on the other hand are not always as attentive to current science as would be desirable. The particular issue which concerns us lies at one of these intersections of empirical and metaphysical questions, and draws chiefly on two lines of thinking.

The first is Dennett's (1991) most metaphysical piece of work to date, 'Real Patterns'[5]. This paper is, in part, Dennett's response to demands that he take some definite position with respect to the question whether he is a realist or an instrumentalist about descriptions generated from his 'intentional stance'. His comments on the reality of patterns are supposed to meet this challenge. A recent critical engagement with that paper by Ross (2000) argues that the Dennettian position on a range of issues is best cashed out by taking the argument of 'Real Patterns' (with significant modifications) as precisely a statement of fundamental ontology, that good Dennettians should hold the world to be 'patterns all the way down' (Ross 2000:160).

The second line of thinking is a particular body of research pointing towards the view that cognition is, and that it *can* be, 'distributed' in the sense of exploiting non-neural resources[6]. In a series of papers Kirsh and

[3] There are cases where such questions become central, as has been persuasively argued by, for example Koyré (1957) and others in connection with the astronomical revolution in the sixteenth and seventeenth centuries.

[5] 'Real Patterns' is reprinted in Dennett (1998). All page references to 'Real Patterns' in this paper are to the pagination of the 1998 version.

[6] There is another sense of 'distributed' typically associated with connectionist architectures, where calling a computational process distributed indicates that the ways in which information is processed by the system in question are unlike those in classical computational systems. For more detail and applications see Rumelhart and MacLelland (1986), and for an excellent overview and introduction see Clark (1989). We take it for granted in the present paper that cognitive processing in brains is distributed in broadly connectionist ways.

Maglio (1991; 1993, 1994; 1997)[7] studied the behaviour of human players of the computer game Tetris and compared it with predictions based on a classical and non-distributed model of optimal play. On the basis of differences between observed human play and the classical predictions, and supported by a number of supplementary experiments they make a compelling case for the thesis that human players exploit a variety of non-neural resources in the course of play, and coin the term 'epistemic action' to name the category of physical actions with cognitive benefits in question.

A short way of putting the question which concerns us, then, is to say that we want to know whether the structures identified by Kirsh and Maglio are real patterns in Dennett's sense. Before explaining in more detail how we see the issues, and how we want to approach that question, though, it will be worth saying a few things about why it is important.

One reason is, bluntly, to 'keep the riff-raff out'. The hard won recognition that cognition is distributed beyond the brain and into the body and world is all too amenable to appropriation by vague, nebulous and unscientific forms of thinking. Part of the way to deal with this, in the interests of a general commitment to naturalism for which we won't argue on this occasion, is to get the metaphysics right. Feyerabend (1987, 1993), who had few metaphysical commitments but vociferously defended the perceived rights of what he called 'traditions' (which could include metaphysical positions) notoriously sought to defend astrology from scientific attack (see especially Feyerabend 1978:91-96)[8]. More recently Dupré (1993:10,263) who is, in some ways sympathetic to Feyerabend has found himself struggling to come up with a good reason (he does admit to having

[7] Although see also Maglio *et al.* (1999) which extends the research to consider the game of Scrabble.

[8] A supporter of Feyerabend might object that his major purpose was to poke holes in the scientistic arrogance visible in the 1975 anti-astrology 'Statement of 186 leading scientists' which appeared in *The Humanist*. Quite so, but in the course of doing so he does draw on a range of empirical evidence for the effects of solar and other celestial activity on terrestrial life as though such evidence supports astrology *per se*. In later life Feyerabend did, admittedly, express dissatisfaction with his 1978 book (see Feyerabend 1995:147).

'prejudices') for rejecting astrology – a fact we would argue is largely explicable by reference to its commitment to a highly disordered metaphysics. Talk of mind, or at least cognition, being spread wider than the limits of the brain can too easily sound close to – for example, Huxley's (1954) peyote-induced talk of 'Mind at Large'. If this seems implausible or even alarmist, consider *The Embodied Mind* by Varela, Thompson, and Rosch (1991) who take their often subtle and interesting reflections on embodied, embedded agents anchored in some good empirical work, to count as evidence against realist and objectivist views of body mind and world. Instead they favour a focus on Buddhist 'transformative analysis' and meditative practices in order to account for 'the basic circularity of our condition'. Clark (1997:173) expresses related worries about their programme.

A more serious reason besides crowd control, is that current thinking in cognitive science is relevant to a variety of philosophical debates concerning mental causation and explanation but stands out off from such debates partly for the lack of much work linking the results of cognitive science with the metaphysical concerns of philosophers. One striking symptom of this is the continued reliance in most of the mental causation literature (for a recent example see Kim 1998) on a basically Humean conception of causation, when more and more of cognitive science moves towards a dynamical systems approach. Without some bridges being built between the two areas, parts of philosophy of mind are going to remain sadly out of touch. Rather than build a bridge from scratch we hope to approach our question largely by moving an available resource into position, that resource being Van Gelder's useful paper on 'The dynamical hypothesis in cognitive science' (1998). In that paper he distinguishes between a *nature hypothesis*, to the effect that cognitive agents actually *are* dynamical systems, and a *knowledge hypothesis*, which is the more cautious, or ontologically agnostic thesis that cognitive agents can be understood dynamically. His discussions will help us get to a position where we can make some headway with what was noted above as our central question, whether the structures identified by Kirsh and Maglio are real patterns in Dennett's sense.

Although Kirsh and Maglio's research is surveyed below, we should say now that we take it as uncontentious that they have pretty much

established a distributed version of Van Gelder's knowledge hypothesis with reference to human Tetris playing. By that means they have established that such players *can* be well understood as dynamical systems which include a range of components or interacting parts, only some of them neural. If there is a case that the nature hypothesis follows, then we have the result that human Tetris players actually *are* distributed systems, that the computational problems demanded by the game and solved by the players are not handled by the brain alone. Specifically, then, the subject of the paper is the *ontology* of the cognitive agents posited by accounting for the mind as dynamically realised by both body and world. The question whether the nature hypothesis follows is one which connects up directly with the debate concerning realist and instrumentalist readings of Dennett. We think that we can make a case for the conclusion that the dynamical structures identified by research like that of Kirsh and Maglio are indeed real patterns in Dennett's sense (as refined and debugged by Ross) or at least are candidates for being real patterns.

In what follows we proceed as follows: First we set out the outlines of Dennett's argument in 'Real Patterns' and explain a little of the context in which it occurs. Then we briefly survey the accumulating evidence for distributed cognition against the backdrop of more traditional expectations. With both of those topics set up we turn to a more detailed examination of Kirsh and Maglio's research on Tetris playing. This in turn allows us to begin dealing with our central question, which occupies the remainder of this paper, beginning with a discussion of Van Gelder on the dynamical hypothesis in cognitive science.

Dennett on Real Patterns

Dennett defends three different 'stances' as ways of looking at, explaining, and sometimes predicting the behaviour of part of the world, each of which serves particular purposes and has particular limitations. The three are the *physical*, the *design* and the *intentional* stances. Being a type of physicalist (although he generally uses the label 'materialist') Dennett thinks that the physical world is causally closed, which means that any physical state of affairs can 'in principle' be causally accounted for by reference only to physical states and processes. When we take the *physical* stance to

something we are concerned with the physical processes which are involved in it, and assuming that we had accurate enough measurements, physical understanding, and computing resources (the proverbial large enough piece of paper and sharp enough pencil) we could, again 'in principle', follow every detail of what happened and even make the best possible, although not necessarily deterministic[9], predictions of what would happen next.

In all except very rare cases we are not in a position to take a physical stance to any system in any detail. Furthermore, and Dennett is especially alive to this next point, what matters to us often doesn't depend on specific facts at the physical level. I just don't care what kind of polymer my drinking straw is made of, as long as it has two, and only two, holes and doesn't make my drink taste odd. This makes sense because often we're concerned with *functional* aspects of the large-scale activity of objects or systems. This is where the other two stances come in.

When we take the *design* stance towards some system we 'assume' (not necessarily as a deliberate decision) that it serves some function, and ask what that function might be, or whether it is working properly, and so forth, but largely ignoring what would be paramount from the physical stance (see Dennett 1998:229f). Design stance descriptions are *simpler* than physical ones, which is to say that they are often simple enough for us to make and understand them at all, which is almost never the case with the physical stance, but this tractability is bought at the expense of leaving out a lot of detail. This makes design stance descriptions and explanations somewhat vulnerable, but they can be very effective – with almost no physical stance information at all, I can predict pretty confidently that a drinking straw with, say, seven holes of the same size won't be much use for drinking.

Finally we can, in some cases, take the *intentional* stance, which means treating a system as having goals and some measure of rationality.

[9] Papineau (1993:16) puts it as follows: 'I take it that physics, unlike the other special sciences, is *complete*, in the sense that all physical events are determined, or have their chances determined, by prior *physical* events according to *physical* laws. In other words we need never look beyond the realm of the physical in order to identify a set of antecedents which fixes the chances of every physical occurrence'.

When we do this we might speak in terms of 'trying' to do *x*, or planning to do *y* or 'wanting' to achieve *z*, which is to say that we ascribe beliefs and desires, or something like them, the apparatus of 'folk psychology', to the system (see, *inter alia*, Dennett 1987, 1998). For Dennett intentional stance descriptions are always interpretations, and, furthermore, are always underdetermined. Nonetheless he thinks that they are the best available strategy for dealing with some types of system in the world, most notably one another. (Just try to describe anything that you or someone *else* is *doing* without referring to intentions, desires, beliefs and so forth.)

It is probably not surprising, given that Dennett thinks there is a physical stance to be taken, *even if* only 'in principle' that there has been some heated debate over the status of what is said from the other two stances, especially the intentional. After all, most of the contemporary mental causation debates concerns problems (real or not) which arise because thinking that physics *is* causally closed seems to make any non-physical causal claim, especially those referring to beliefs and desires, seem threatened with redundancy. Dennett sometimes makes Quinean noises here, suggesting that intentional descriptions (referring to 'propositional attitudes') are a kind of *façon de parler* or 'dramatic idiom' (e.g. Dennett 1987:110), which doesn't refer in the same way that fundamental physical descriptions do, but pay their way by being *useful*. But he also goes to some lengths to make clear that there are *true* explanations which are possible *only* from the intentional or design stances. So is he a realist or an instrumentalist?

He has said that he aims to place his view 'firmly on the knife-edge between the intolerable extremes of simple realism and simple relativism' but admits that 'this has not been recognized as a stable and attractive option by many others in the field' and that his 'critics have persistently attempted to show that my position tumbles into one abyss or the other' (Dennett 1987: 97). He also insists that his view should be seen as 'a *sort* of realism, since I maintain that that patterns [visible from the intentional stance] are really, objectively there to be noticed or overlooked' (Dennett 1987:97).

Part of the problem for Dennett here is that the adoption of the intentional stance seems to involve postulating entities and processes which are both abstract and observer dependent, which encourages an instrumentalist reading. On the other hand the facts about the success or

failure of the adoption of the stance, are to Dennett perfectly objective, so that beliefs and desires look likely in some sense to be real. In *The Intentional Stance* Dennett acknowledges the problem, and reiterates his choice of the knife-edge position:

> My thesis will be that while belief is a perfectly objective phenomenon (that apparently makes me a realist), it can be discerned only from the point of view of one who adopts a certain predictive strategy, and its existence can be confirmed only by an assessment of the success of that strategy (that apparently makes me an instrumentalist) (Dennett 1987:15).

There are two main components to Dennett's ongoing development of his position here: one is the defence of the *usefulness* of the intentional stance, and the other concerns the status of the 'patterns' referred to in the passages quoted above. It is the second component which is of interest to us, and we take it as more than established that the intentional stance enables the making of powerful predictions with greater efficiency than any other available method.

In pursuit of both projects Dennett (1987: 25-28) offers a thought experiment which for our purposes we will simplify and shorten. Suppose some Martians with superior intelligence were to land on earth. Suppose, further, they did not need the intentional stance or even the design stance to predict our behaviour in all its detail, but that instead they can comprehend the behaviour of people at the microphysical level in order to predict behaviour down to, for example the next bit of you, the reader's, eyelid. Essentially, we are to imagine, they can predict the individual behaviours of all the bodies they observe without ever treating any of them as intentional systems. (We are not to suppose that they do this with themselves or each other.)

From the Martian point of view, then, we really are not believers at all. If so, it might seem, our status as believers is not objective, but is instead an anthropomorphic notion, of utility given to our limited ability to track the world at the physical level. Dennett suggests, however, that even though these Martians may be able to predict the future of the human race at a purely physical stance level, if they did not also see us as intentional

systems, they would be missing something *perfectly objective*. Namely the *patterns in human behaviour* that are describable from the intentional stance, and *only from that stance*. The patterns that is, that support *our own* generalisations and predictions.

In this respect suppose that a Martian engages in a prediction contest with an earthling. The Martian, doing all his microphysical calculations, predicts the behaviour of an individual. Given that the earthling would be equally able to predict the behaviour of that individual, say a batsman after being conspicuously clean bowled, without access to the physical stance, the Martian would be left in amazement at how this was done. And it *is* amazing to think that an intentional stance prediction could sometimes, indeed often, do just as well as one generated from the physical stance, even assuming hugely unrealistic resources for pursuing the physical stance.

We know that we can make generally effective predictions of the behaviour of intentional systems without paying attention to their microphysical constituents. The question, though, is *why* we are able to do this, or what we are latching onto when we do. Dennett's answer is that there are high level *patterns* in the activity of intentional systems, and that we can and do track these patterns.

In fact, Dennett is committed to their being a range of 'higher level' patterns, not only ones associated with intentional systems. He illustrates the key idea by means of a discussion of Conway's *Life*, the cellular automata designed to test some of von Neumann's ideas concerning replication[6]. The universe of *Life* is a two dimensional grid, the cells of which can be either occupied or unoccupied. Time in *Life* moves discretely from generation to generation, and the state of the next generation is determined by the following rules, applied to each cell in the grid: If the number of occupied neighbours of a cell is *two*, that cell stays in its current state into the next generation; if the number of occupied neighbour cells is *three* the cell will be occupied in the next generation, and in all other cases the cell will be unoccupied. That's all simple enough, what is interesting is what can come out of such a system of rules.

In an important paper on general features of cellular automata Wolfram classifies cellular automata into types based on their dynamical

[6] Dennett (1995) also discusses *Life* in related ways.

properties. He divides them into four types (Wolfram 1984:5), those, class I, which converge quickly onto and remain in a homogeneous state. Those, class II, which result in separated simple configurations of occupied cells and/or converge on a periodic pattern of states, those, class III, which result in some or other chaotic pattern, and finally those of class IV, which produce 'complex localized structures, sometimes long-lived.' These options correspond roughly to standard dynamical possibilities, in that the first class amounts to a system with a point attractor, the second to a periodic attractor or limit cycle, and the last two to 'strange attractors'[1], or at least to configurations and sets of rules which take a very large number of generations to converge on some more stable state of class I or II. Classes III and IV are of most interest, since they manifest the most interesting structures and also have the property that prediction of what structures will result from the application of the rules to a given starting situation is not typically determinable by any means except direct simulation of the rules and configuration (Wolfram 1984:31; see also Wolfram 1983).

The structures possible in Life are rich and varied, including 'blinkers', 'gliders', 'spaceships', 'puffer trains', 'glider eaters' and so forth. A glider is most easily described as a shape which moves through Life-space, contorting through a fixed sequence of transitional states like a little digital caterpillar. Watching Life run on a computer screen and seeing a glider it is all but impossible to think of the glider as a moving thing, but there's the rub: from the bottom level of description there are 'only' cells which are either on or off, and no motion takes place. Are gliders, then, and their motion, or any of the other furniture and behaviour of the Life world, real?

Dennett points out that it has been established in principle that a Universal Turing Machine could be constructed on the Life plane, and invites us to imagine one which also happened to be running a chess playing program. The required array would be huge, requiring about $10^{10^{10}}$ pixels according to some estimates (Dennett 1998:109). A range of levels of description of such a Life array are possible—from one concerning single

[1] Lorenz (1963) is the classic reference here, although he himself does not use the term 'strange attractor' in this paper, referring to the time development of his artificial weather as 'quasi-periodic' instead.

cells, through ones concerning gliders and other objects, right up to ones interpreting macro states of the array as chess moves. The 'higher' the level of interpretation, the less work is needed to make predictions, including good ones such as 'the next move should be an attempt to save the Queen, probably QxP'. Indeed such predictions can be routine and relatively easy, whereas applying the transition rules of Life to trillions of individual cells is prohibitively demanding.

The 'objects' to which we refer when making these higher level descriptions, interpretations and predictions are abstract objects, according to Dennett, what Reichenbach called *abstracta*, the same type of objects as centres of gravity or 'Dennett's lost sock centre, the point defined as the center of the smallest sphere that can be inscribed around all the socks [Dennett has ever] lost in [his] life' (Dennett 1998:97). Reichenbach's *abstracta* were to be contrasted with *illata*—the more fundamental, level 'physical' properties of something, which in the Life case would be the states of individual cells. Dennett, arguing that metaphysically speaking all abstracta are equal, says we should concentrate on which ones pay their way by giving us explanatory, descriptive or predictive leverage, which he thinks intentional stance descriptions do exceptionally well. The 'good' abstract objects, he says, are those which 'deserve to be taken seriously, learned about, used' (Dennett 1998:97).

While partly dodging the demand for an account of reality which would answer the question whether he is realist or instrumentalist about even 'good' abstracta, Dennett also proposes that a kind of realism about patterns is plausible, actually 'demonstrably right' (Dennett 1998:97), drawing on the information theoretic notion of compressibility; patterns, unlike randomness, are compressible. Patterns, in this sense, 'are there to be described—whether or not we care to see them' (Dennett 1987:28). And, we might add 'whether or not we care to get metaphysical about them.' Before we can start asking our question about the status of the patterns discovered in Kirsh and Maglio's research, though, we need to say a little about distributed cognition.

Distributed Cognition

It can seem natural and innocuous to suppose that cognition, or thinking, is something done exclusively by the brain, or at least the central nervous

system. For some time now the dominant image of cognition has been that it is *some* kind of computation. While controversy and debate continue on exactly what kind of computation it might be, there is something like a consensus that some kind of information processing will be going on. Not only that, brains look like splendid candidates for organs which do the processing, especially in the light of the past few decades of work on neural networks and ever finer understanding of the functioning of the brain itself. Saying that brains are (*inter alia*) powerful information processing engines of some sort, though, is not the same as saying that all information processing goes on in brains. These two distinct claims can all too easily be elided, partly because it can seem intrinsically implausible that something which doesn't have a relatively conspicuous information processing architecture, which a huge collection of connected neurones clearly *does*, could be doing any computing at all. With Cowley and Spurrett (forthcoming) we call the view that all cognition does take place in the brain 'cognitive internalism'.

There is mounting evidence, though, that cognitive internalism is misguided, and that human and animal cognition, not to mention the most effective and flexible artificial cognition, exploits non-neural resources, i.e. bodily and/or environmental ones, in a wide range of ways. An excellent survey of research of this type is Clark (1997), although see also Thelen and Smith (1994), Brooks (1991a). Opposed to cognitive internalism, then, is the view that cognition not only can be, but typically is distributed. The notion of solving problems by leaning on the environment is, of course, not new. Vygotsky's (1968) 'zone of proximal development' concerns what the individual may achieve with the aid of an external guide. But recent work on distributed cognition tends to find that distributed cognition is not a developmental phase, but an ongoing feature of effective cognition. This work also concerns itself with developing finely structured models of the distributed processes which it identifies and studies[2]. For our purposes, as noted above, we will focus on a particular example, which is Kirsh and Maglio's research on Tetris-playing behaviour.

[2] For a brief survey of the case for distributed cognition see Spurrett (this volume).

Kirsh and Maglio on Tetris Playing

Tetris is an excellent arena for testing hypotheses about cognition. It is a game played at a computer, in which players attempt to position various falling shapes (which Kirsh and Maglio call 'zoids') so that they completely fill horizontal rows across the bottom of the field of play, and result in the player being awarded a number of points. Completed rows disappear and allow whatever is above them to move that number of rows downwards as a result. Incomplete rows become buried under poorly placed zoids, which reduces the amount of space, and hence time, available for placing new pieces. The game ends when the field of play becomes so clogged with incomplete rows that no new zoids can descend from the top. Besides the way in which misplaced zoids increase difficulty, the game automatically increases the rate with which zoids fall as a function of accumulated points. The player's task of matching zoid shapes into the contour formed by zoids already in place, is effected by four possible actions that the player can perform with single keystrokes. They are, (1) rotate a zoid 90 degrees counterclockwise, (2) translate it one step to the left, (3) translate to the right, and (4) drop. The latter involves moving the zoid instantly from its current position to the position it would eventually come to rest if no more keys where pressed. Only one zoid is ever in motion at a time, and a new one randomly selected from the available types appears the moment the preceding one comes to rest, or is dropped (Kirsh & Maglio 1994).

The game, then, has definite objectives, and unambiguous criteria of success and failure. It makes increasing demands which are highly time dependant, and it is possible to gather a range of temporally precise data about player behaviour some of which can be easily interpreted and evaluated given that fact the aims of the game are specified in advance[3]. As noted at Spurrett (this volume), Kirsh and Maglio were led by their investigations to propose a category of 'epistemic' actions, to be contrasted

[3] Similar advantages follow from Hutchins' (1995) decision to study co-operative cognition in teams by focussing on navigation, which yields a similarly quantitative backdrop helping interpret and evaluate the activities of navigation teams.

with 'pragmatic' actions[14]. A pragmatic action is one undertaken to bring an agent closer to some physical goal, whereas an epistemic action (also a physical action) is performed in order to modify the computational state of the system performing the action. A more detailed account of epistemic actions states that an epistemic action is one which improves cognitive performance by:

> 1. reducing the memory involved in mental computation, i.e., space complexity;
> 2. reducing the number of steps involved in mental computation, i.e., time complexity;
> 3. reducing the probability of error in mental computation, i.e. unreliability (Kirsh & Maglio 1994).

Kirsh and Maglio's empirical case for the occurrence of epistemic actions with respect to Tetris playing is founded on a contrast with what a classical cognitive internalist approach to the game would be. They develop a model of how such an approach would deal with the computational problems posed by the game, which breaks processing down into the following four stages:

> 1. Create an early, bitmap of representation of selected features of the current situation;
> 2. Encode the bitmap representation in a more compact, chunked, symbolic representation;
> 3. Compute the best place to put the zoid;
> 4. Compute the trajectory of moves to achieve the goal placement (Kirsh & Maglio 1994).

Using this model Kirsh and Maglio predicted the number of zoid rotations which would be optimal, and then observed the actual behaviour of human players. The model predicts that rotations should only take place *after* the optimal zoid placement has been established, which is not what happens at all: rotations and translations are far more abundant than the traditional

[14] Kirsh and Maglio (1992) describe an additional category of 'perceptive' actions.

view predicts, and they start earlier than it can account for. In fact, they often start *before* a new zoid is completely visible, which in terms of the classical model would imply that they were entirely unmotivated.

Of course Kirsh and Maglio can't, and don't, just say that these actions have to be 'epistemic' actions. Rather, by means of a series of subsidiary experiments, they show how the efficiency of human computation is improved by means of them, in a range of ways. These include demonstrations that rotated zoids can be more quickly identified; that engaging in physical rotations simplifies the problem of detecting 'fit' between a new zoid and the contour below; and that performing translations which 'bounce' zoids off the walls of the playing area reduce the likelihood of vertical alignment errors. Finding the rationale, or the pay off, for engaging in these actions, though, means abandoning not only the details but the foundations of the traditional view Kirsh and Maglio used to generate predictions and empirical contrasts. Computation, or Tetris cognition, isn't all 'in the head' but rather draws on a range of resources extending beyond the brain, which are closely coupled together to achieve optimal overall performance.

For our purposes it is important to note that one result of Kirsh and Maglio's work is a kind of specification of the high level, or functional, properties of a distributed cognitive architecture. The cognitive system they discovered and partly analysed is at some remove from a low level neural and muscular description, and vastly removed from a bottom level 'physical stance' description in terms of fundamental particles. Rather their analysis highlights, partly by means of a kind of 'reverse engineering'[15] what it is that brain, body and external systems (especially screen and keyboard) contribute to dealing with the computational demands of Tetris, and why these resources do a better job in just *those* combinations rather than others. Their work is not concerned with *how* exactly fingers, brains, eyes, screens and so forth do what they do, and so components which were functionally equivalent but structurally different would in principle be interchangeable would in a sense be indistinguishable to their enquiry. What they find and describe is at the level of the design stance, given that an overall intentional

[15] Dennett (1995:212-219) discusses reverse engineering and 'artefact hermeneutics' *inter alia*.

stance reading comes for 'free' with the rules and objectives peculiar to the game of Tetris.

Let's call for architecture that they discover a 'structure' to leave open whether it is a 'pattern' or not. As a structure, it is a specification of a distributed, or coupled cognitive architecture. Now we ask whether it is the type of thing a Dennettian should want to be a realist about. As we said above, we use Van Gelder on dynamical systems as something of a bridge between empirical and metaphysical questions, and having asked about the reality of Kirsh and Maglio's structure, we leave that question hanging for a while to take a look at Van Gelder's views. Although Van Gelder refers to a range of examples of empirical cognition research, Kirsh and Maglio's work is not on his list of examples.

Van Gelder on Dynamical Systems

In his 1998 publication, Van Gelder sets out to clarify and defend what he calls the 'dynamical hypothesis' in cognitive science, which is the hypothesis that cognitive agents are dynamical systems. He argues that this view could replace[6] what he takes to be the still-dominant computational hypothesis, holding that cognitive agents are digital computers. Van Gelder's contention is that digital computers and dynamical systems are importantly different types of system, and the ways in which he makes and defends this claim are important for our own argument here.

For his purposes Van Gelder takes a system to be a 'set of interdependent variables' where the state of the system is 'simply the state or value of all its variables at a time' and the behaviour of the system 'consists of transitions between states' presumably depending on the precise nature of the interdependence between the variables (Van Gelder 1998:616). On this wide understanding it is clear that both digital computers and dynamical systems are going to be examples of systems. Before turning to the problem of distinguishing the two, which he clearly most of his thesis is to make

[6] Van Gelder (1998:527) takes pains to indicate that whether the dynamical hypothesis will replace the computational hypothesis would depend on the outcome of 'sustained empirical investigation'.

sense, Van Gelder offers one further distinction, which is that between 'concrete' and 'abstract' systems.

A 'concrete' system is, for Van Gelder, a system such as the solar system in which the variables in the system (he is presumably referring to masses, positions, momenta and so forth) are actual features of the real world changing in real time in accordance with natural laws' (Van Gelder 1998:617). Although he makes no mention of Dennett (who does not appear at all in his list of references), note the similarity between a concrete system in this sense and the type of consideration which would be relevant when taking the physical stance. An 'abstract' system, on the other hand, is a set of abstract variables governed by mathematical rules' (Van Gelder 1998:617). Sometimes we can build or find a concrete system which 'realises' an abstract system. Van Gelder's example being a physical calculator which realises a given 'abstract computational system'. In addition to this, Van Gelder sets up what he calls a relationship of 'implementation' in which a low level system implements a higher level system only when the variables of the higher level system are 'somehow constructed' out of variables of the lower level ones. The implementation relationship is weaker than full identity, but does allow some degree of identification at the level of behaviour.

Van Gelder considers the most significant features of digital computers to be that variables and states of such systems are themselves digital, that time is discrete, state changes or computation depend on an algorithm, and that the states and behaviours of the system admit of systematic interpretation (consider the calculator example above). The level at which we are typically interested in a digital computer is that governed by the interpretation, rather than with details of the concrete system itself, on which the object of our interest is implemented.

With respect to dynamical systems on the other hand, Van Gelder emphasises the importance of the fact that variables or states in such systems are typically quantitative rather than discrete, that time does not typically operate discretely and further that quantitative states can depend on time, and that rates of change can play a significant role in fixing future states. Van Gelder proposes to take a system as dynamical 'to the extent that it is quantitative in one of the above senses' (Van Gelder 1998:619). An obvious objection at this point would note that on the basis of the characterisations

just set out digital computers simply *are* dynamical systems, which looks *prima facie* awkward for Van Gelder's main project. This is, of course entirely true—digital computers are at basic concrete systems in Van Gelder's sense, and they undergo changes which means that *some* dynamics must be true of them.[7] Van Gelder is entirely aware of this, and also of the fact that many of the dynamical systems he has in mind are studied by means of software which is run on the very digital computers which he thinks are a bad model of cognition. Before looking at how Van Gelder deals with the objection it is worth pausing to note a similarity with the case of Dennett. Recall the complaint that Dennett's account of intentional stance descriptions is supposedly doomed to collapse into instrumentalism just because the alternative physical stance description seems to be holding the causal trumps. Dennett's position is supposed to show how it can be OK to be a kind of realist about intentional descriptions without weakening his commitment to the physical stance. One can readily imagine a committed 'digitalist' saying that Van Gelder might well be correct to say that interesting and important things can be said about dynamical systems, but that at bottom we know about them because of things we have done with digital computers, and hence that digital approaches are still the ultimate causal and explanatory winners[8].

As noted above Van Gelder distinguishes between a 'nature' and a 'knowledge' hypothesis, each of which he takes to be a component of the dynamical hypothesis. The nature hypothesis states that cognitive agents *are* dynamical systems, while the knowledge hypothesis states that cognitive science should 'take dynamical form', i.e. that cognitive agents are best studied as dynamical systems. The objection presently on the table effectively grants the knowledge hypothesis but refuses to allow an inference from it to the nature hypothesis. Van Gelder attempts to cope with this difficulty by means of a notion of *simulation* where he grants that a digital computer can simulate a dynamical system, by means of software which in

[7] For a similar point see Spurrett (1999:2ch1) discounting Colliers' (1998) distinction between 'complex' systems and those which are 'merely complicated'.

[8] This objection has been made to Van Gelder in the responses to his paper. See, e.g., Quartz (1998) and Beer (1998).

turn is implemented on the low level dynamical system which the computer instantiates (Van Gelder 1998:620). This long route from instantiation to implementation to simulation stands in stark contrast to the case of digital computers running abstract cognitive models which are themselves digital, where Van Gelder holds that the relationship between the computer and the model is one of *realisation*, given the stronger structural similarities between the model and the basic properties of the computer. Whether this answers the objection, and to what extent Van Gelder's proliferation of types of system-system relation (identity, instantiation, implementation, realisation, simulation) bears up under rigorous scrutiny are crucial questions, which we postpone until later in this paper. For now we need to note that Kirsh and Maglio's work definitely counts as a kind of exercise in what Van Gelder sees as the proper method of cognitive science, an instance of the 'knowledge hypothesis' at work and delivering results.

Kirsh and Maglio, though, do not deliver (for reasons noted above) a detailed working dynamical description of human Tetris playing. They describe a system which is both distributed and dynamical, to be sure, but it is what Van Gelder calls a description of 'how agents are causally organized at the highest level relevant to explanation of cognitive performances' (1998: 619). One might say that as the cognitive internalist model fell in the face of the facts, Kirsh and Maglio's ongoing research revealed the broad outlines of a distributed system with distributed boundaries and a particular division of labour across itself. This understanding permitted significant testable predictions.

Consider, for example, the action of rotating a zoid *before* it has completely emerged from the top of the screen, which Kirsh and Maglio observed to be common in human players. Recall the classical planning model of Tetris playing, and how on its terms early rotation, i.e. rotation before the computational problem has been 'solved' can have no possible value. Kirsh and Maglio suggest that these actions make sense however, if we consider the possibility of an epistemic function, namely that the rotations are enacted to unearth information early in the game. Given the advantage of early identification, and the fact that zoids only partially appearing from the top of the screen are on most occasions ambiguous, Kirsh and Maglio (1994) predicted that 'if a strategy exists for disambiguating shapes early then good players would strike on it'. Rotating a zoid early to

expose initially hidden parts is such a strategy. Sometimes, however, rotations are not necessary to identify an emerging zoid, since if a player has expert knowledge of the game they will realise that certain shapes emerge only from certain columns and thus early rotation would reveal no new information. Interestingly, and in line with the thesis that early zoid rotation is epistemic, zoids that are completely ambiguous due to their initial partial images being identical in both shape and position, are rotated more than those which are not. Such a trend in the data is consistent with the thesis that the point of early rotation is to uncover new information immediately, since this bias in rotation rules out a purely pragmatic purpose. The action thus serves not to position the zoid towards the final pragmatic placement but rather to aid the computational end of revealing new information. Action and cognition are coupled.

The cognitive processing model which Kirsh and Maglio (1994) take their research to discover suggests that 'individual functional units inside the agent [are] in closed loop interaction with the outside world'. The result is a tight coupling between internal and external processes, which allows for 'offloading structure to the world, or for arranging things so that the world pre-empts the need for certain representations, or pre empts the need for making certain inferences' (Kirsh & Maglio 1994). Action, on their view, is part of the *way in which cognition proceeds*.

It would be possible to explain in further detail the ways in which Kirsh and Maglio's work fits with the finer points of Van Gelder's account of dynamical systems, in particular his emphases on dynamicists' interest in changes rather than states, parallel rather than serial processes, ongoing processes rather than ones which divide neatly into input/output stages, caution or even hostility to the standard view of the role of internal representations in cognition, and the relative importance of 'coupling' compared to state-setting. Such a discussion would be partly tangential to our central purpose, though, and possibly of future independent interest. The key point is that Kirsh and Maglio's work fits very well with Van Gelder's account of the knowledge hypothesis in cognitive science, and that it establishes a very plausible model of Tetris playing in which cognition is highly distributed, and where there are a variety of couplings between the agent as conceived from a cognitive internalist perspective and the body and environment of that agent. (As Van Gelder 1998:619 notes, the dynamical

hypothesis is open to the possibility that a properly understood cognitive system 'might include variables not literally contained within the agent itself'.) This higher level dynamical understanding makes sense of the observational data, and is further supported by the range of subsidiary experiments which were part of Kirsh and Maglio's project.

There doesn't seem to be any problem with saying that the higher level dynamical functions and structures identified by Kirsh and Maglio are patterns in Dennett's sense. They are not observed from the physical stance, but relate to functions and regularities observable at the level of the cognitive system as a whole, and given both some input from the intentional and design stances. Are they *real* patterns, though?

Nature and Knowledge

Let us grant for the purposes of argument that Kirsh and Maglio satisfy the requirements of Van Gelder's knowledge hypothesis. The question, then, is whether there is a defensible inference in this case to the truth of the nature hypothesis – that Tetris players cannot just be understood as distributed dynamical systems, but that they *are* such systems. The couplings they describe between mind and world, brain and game could be seen as useful constructs in conceptualizing and accounting for the actions of Tetris players.

Van Gelder (1998:615) would presumably say that the answer here is 'yes'. He goes so far as to suggest that the dynamical hypothesis is a strong candidate for being the 'law of qualitative structure' for cognition, to replace Newell and Simon's proposal that cognition should be understood in terms of physical symbol systems. With reference to his differentiation of the dynamical hypothesis into the nature and knowledge hypotheses, Van Gelder claims that 'the best evidence for the former would be the truth of the latter' (Van Gelder 1998:619).

Just how good, though, is the best evidence here? Van Gelder is careful to note that even the *nature* hypothesis is 'concerned in the first instance not with low-level systems but with how agents are causally organised at the highest level relevant to an explanation of cognitive performance, whatever that may be' (1998: 619). So an understanding of a putative dynamical system in terms of the knowledge hypothesis would be a

high-level dynamical model which, if correct, so that the nature hypothesis would be true in this case, would be *instantiated* by some real physical system. As noted above Van Gelder distinguishes instantiation from simulation, partly in order to make sense of how digital computers, whose qualitative structure is of the sort he wishes to say should not be taken as paradigmatic of cognition, can nonetheless be used to explore the dynamical systems he thinks are paradigmatic. Recall that the claim he made was that a digital computer can instantiate a simulation tool, which in turn can simulate a dynamical system of possible interest (Van Gelder 1998:620). Only the simulated system is considered to be a model of the target system on his view, rather than in what he calls ordinary computational modelling where *both* the concrete computational system and the abstract (higher level) digital system are considered models of the target system. The target system is in turn supposed be an instantiation of a dynamical system made out of concrete objects. (For good measure the target system in turn, for Van Gelder, realises an abstract dynamical system, to which it is hoped the system simulated on the concrete computer system will correspond. Happily this last layer is not crucial for our purposes.)

What should be manifest, if by no means obvious, is that the relationship between the target system and the concrete objects out of which it is composed is importantly similar to the relationship between an intentional system and the physical level of description on Dennett's account. The dynamical properties at the higher level of description are not of the same type as those at the basic level, they permit some variation between high and low level, which is to say that they can be multiply realised. This means that if we grant, for the purposes of argument, that the abstract dynamical system produced by some piece of cognitive science really does model a target system, even to the extent of being isomorphic with it, a question remains: Should we be realist or instrumentalist about the target system? If anything is to be considered an *abstracta*, it is, after all, the target system, a higher level dynamical system not identical with, and not reducible to, some set of concrete bottom level dynamics at the level of *illata*. So we're left on Dennett's knife-edge.

Ross (2000:147) quotes Dennett explaining why he is alarmed at philosophers who take 'Real Patterns' to be an exercise in metaphysics or ontology:

> I wouldn't want to rest my ontology and then find I have to spend the rest of my life defending or revising it, instead of getting on with what to me are the genuinely puzzling issues—like the nature of consciousness, or selves, or free will (Dennett 1993:212).

For his part Ross thinks Dennett *does* need a metaphysics, and that 'Real Patterns' goes most of the way to delivering it, but ultimately fails to do so because of the way Dennett buys into the distinction between illata and abstracta. He notes that the distinction can seem like a way of doing justice to the 'primacy of physics', which is clearly important for Dennett's materialism (Ross 2000:152).

Dennett is also, as we saw above, a deeply committed antireductionist. This cuts off one way of underwriting the illata/abstracta distinction, since an ontological reductionist who is also convinced about the primacy of physics can view any non fundamental kind as a posited kind (an abstracta) made up out of fundamental entities (the illata). Apparent compromise positions, represented by a wide range of supervenience theories, founder on the fact that they are typically wedded to untenable assumptions about the nature of physics. As Ross (2000:150) notes, supervenience theorists generally hold that science produces nomic generalisations, and that supervenient types can be multiply realised. Their antireductionism is based on the claim that the disjunctive set of physical tokens which is the base for a supervenient type will not be able to feature in the right kind of laws. But there is no reason to accept this view of how physics proceeds, and hence the required contrasts collapse, so that the supervenience theorist's 'defense of the reality of supervenient types threatens to collapse into instrumentalism: Our special sciences generalize over supervenient types *only* because we lack the epistemic resources necessary for identifying their disjunctive supervenience bases' (Ross 2000:156). But only that, on this view the abstracta, or supervenient types, end up being anthropocentric artefacts *of our* epistemic limitations.

Ross's argument (with which we are being very brief here) is that the way to proceed is to abandon the distinction between illata and abstracta, and the implication of degrees of reality, or of first and second class ontological citizens, entirely. His proposed account of reality is as follows:

> To be is to be a real pattern, and a pattern is real if
> (i) it is projectable under at least one physically possible perspective, and
> (ii) it encodes information about at least one structure of events or entities S where that encoding is more efficient, in information theoretic terms, than the bit-map encoding of S, and where for at least one of the physically possible perspectives under which the pattern is projectable, there exists an aspect of S that cannot be tracked unless the encoding is recovered from the perspective in question (Ross 2000:161).

The chief strengths of this proposal are that it does justice to the primacy of physics, is not reductionist and also does not permit an infinite ontology. Crucially, it is not anthropocentric, and hence could not possibly be instrumentalist. We also note that it is partly an extension and refinement of the information theoretic notion of patterns hinted at in Dennett's own talk of compressibility in 'Real Patterns'. We don't aim to defend Ross's proposal here, at least not on its own metaphysical terms, but we do hope to show how by helping Dennett it helps Van Gelder and answers our question about the reality of the patterns, or 'structures', discovered by Kirsh and Maglio.

Van Gelder, like Dennett, needed a justification for the move from the applicability of some approach (dynamical systems, the intention stance) to realism about the type of structure to which that approach seemed to be committed. In both cases instrumentalism presented itself as either a genuine option, or an unacceptable alternative to be guarded against. Note that Ross's account gives neither Dennett nor Van Gelder the conclusion that the things which matter to them, intentional systems or cognitive-dynamical ones, actually *are* real. What it does do, though, is provide a way of saying that they *could* be, and what it would mean if they were. What exactly is or is not real is, quite properly, on Ross's view an empirical question, and one which there is no guarantee we would be able to answer. The answer to our question, then, is that what Kirsh and Maglio discover could indeed be a real pattern in a Dennettian sense. That is to say that research on distributed cognition could provide not just a powerful way (one among many) of

describing cognition, but that it could bring us closer to a better understanding of the types of things there really are in the world.

Department of Philosophy
University of Natal

References

Bechtel, W & G Graham (eds) 1998. *A Companion to Cognitive Science*. Oxford: Blackwell.

Bechtel, W, A Abrahamsen & G Graham 1998. The Life of Cognitive Science. In W Bechtel, & G Graham (eds): *A Companion to Cognitive Science*. Oxford: Blackwell.

Beer, R 1995. Framing the Debate between Computational and Dynamical Approaches to Cognition. *Behavioral and Brain Sciences* 21:630.

Brooks, RA 1991a. Intelligence without Reason. *MIT AI Lab Memo* 1293, April 1991.

Brooks, RA 1991b. Intelligence without Representation. *Artificial Intelligence* 47:141-160.

Brooks, RA 1997. From Earwigs to Humans. *Robotics and Autonomous Systems* 20,2-4,291--304.

Cilliers, P 1998. *Complexity and Postmodernism: Understanding Complex Systems*. London: Routledge.

Clark, A 1989. *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*. Cambridge, Mass.: MIT Press.

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

Dennett D 1978. *Brainstorms*. Cambridge, Mass.: MIT Press.

Dennett, D 1987. *The Intentional Stance*. Cambridge, Mass.: MIT Press.

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

Dennett, D 1993. Back from the Drawing Board. In Dahlbom, B (ed), *Dennett and His Critics*. Cambridge, Mass.: Blackwell.

Dennett, D 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Touchstone.

Dennett, D 1998. *Brainchildren: Essays on Designing Minds, 1984-1996*. Cambridge, Mass.: Bradford Books.

Dupré, J 1993. *The Disorder of Things*. Cambridge, Mass.: Harvard University Press.
Feyerabend, PK 1978. *Science in a free Society*. London: Verso
Feyerabend, PK 1987. *Farewell to Reason*. London: Verso.
Feyerabend, PK 1993. *Against Method (third edition)*. London: Verso.
Feyerabend, PK 1995. *Killing Time: The Autobiography of Paul Feyerabend*. Chicago: Chicago University Press.
Hutchins, E 1995. *Cognition in the Wild*. Cambridge, Mass.: MIT Press.
Huxley, A 1954. *The Doors of Perception*. London: Chatto & Windus.
Kim, J 1998. *Mind in a Physical World*. Cambridge Mass.: MIT Press.
Kirsh, D & P Maglio 1992. Reaction and Reflection in Tetris. *Research Report D-015*. Cognitive Science Department. University of California. San Diego
Kirsh, D & P Maglio 1992. Perceptual actions in Tetris. AAAI Spring Symposium on Control of Selective Perception, Stanford University.
Kirsh, D & P Maglio 1994. On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science* 18:513-549. Note pagination corresponds to the version at: http://cogsci.ucsd.edu/~kirsh/
Kirsh, D & P Maglio 1997. Some Epistemic Benefits of Action: Tetris, a Case Study. Proceedings of the fourteenth annual conference of the Cognitive Science Society. Hillsdale, New Jersey.
Koyré, A 1957. *From the Closed World to the Infinite Universe*. Baltimore: Johns Hopkins University Press.
Lorenz, EN 1963. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences* 20:130-141.
Maglio, PP, T Matlock, D Raphaely, B Chernicky, & D Kirsh 1999. Interactive Skill in Scrabble. In Proceedings of Twenty-first Annual Conference of the Cognitive Science Society. Mahwah, NJ: Lawrence Erlbaum.
Papineau, D 1993. *Philosophical Naturalism*. Oxford: Blackwell.
Quartz, S 1998. Distinguishing between the Computational and Dynamical Hypotheses: What Difference Makes no Difference? *Behavioral and Brain Sciences* 21:649-650.
Ross, D 2000. Rainforest Realism. In D Ross, A Brook, & D Thompson (eds). *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge, Mass.: MIT Press.

Ross, D, A Brook, and D Thompson (eds) 2000. *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge, Mass.: MIT Press.
Rumelhart, DE & JL McClelland (eds) 1986. *Parallel Distributed Processing: Volume I & II*. Cambridge, Mass.: MIT Press.
Spurrett, D 1999. Review Article of Gellner, E,1998. Complexity and Postmodernism. London. Routledge. *South African Journal of Philosophy* 18(2):268-274
Spurrett, D (guest editor) The Case for Distributed Cognition: A Survey. *Automatica* 19,1.
Spurrett, D & Cowley, S (forthcoming). Out on Apes (Body and Language) Together again. To appear in *Language Sciences*
Thelen, E. & L. Smith 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, Mass.: MIT Press.
Varela, F, E Thompson and E Rosch 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, Mass.: MIT Press.
Van Gelder, T 1998. The Dynamical Hypothesis in Cognitive Science. *Behavioral and Brain Sciences* 21:615-665.
Vygotsky, LS 1968. Thought and Language. (Translation of 1962 edition). Cambridge, Mass.: MIT Press.
Wolfram, S 1983. Statistical Mechanics of Cellular Automata. *Rev Mod Phys.* 55:601-644.
Wolfram, S 1984. Universality and Complexity in Cellular Automata. *Physica D.* 10:1-35.

Language in Iterating Activity: Microcognition Re-membered

Grant Blair and
Stephen Cowley

1.0 Introduction

Radical as it's consequences may be, *pensa fecit* turns is not any implausible view of language. Linguistic theory notwithstanding, even casual observation suggests that during practical human activities, what we ordinarily call 'language' spreads across brains, bodies and world. Seeking ways of understanding it's 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) using word-based units or 'linguaforms'. Rather than stress internal processes, we highlight that much of the *work* of utterance activity arises at and across the boundaries of skin and skull. Much of what we mean and do is, initially, real time responding to prosodic and visible expression. Clarifying, we use Kirsh and Maglio's (1992; 1994) work on how successful (or failure) in computer games explores real-time activity. Using parallels between a player's microcognitive moves and events of utterance activity, we both endorse and criticize aspects of Clark's (e.g. 1997; 1998) extended 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. Thus thesis, we claim, throws conceptual light on Cowley's (1994; 1997; 1998; 2001b) empirical work on how prosody functions. Finally, we claim that

microcognitive acts enable word-based labels to promote remembering. What human bodies and voices do *together* affects how we conceptualise our enculturated 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 examine language scientifically, their models often present linguistic *abstracta*—static entities—as units susceptible to internal processing. 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 'grammar makes linguistic theory compatible with input-output views of what we do with words. Generally, then, what Clark&Wilkes-Gibbs (1995) first called 'linguaforms' 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 enculturated 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 more-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'.[3] On this view, what needs to be clarified, then, is how iterated social activity allows us to pick up on those 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, enculturated 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 enculturee 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' rests a triad 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 be ong neither to a subject nor the world

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

out, 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 wherever 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 skills, 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, 1996) 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 abel 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 learning '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 using a parallel view, Augustine is at pains to see off an alternative objection. Thus, he seeks to clarify how intentions are identified.

> Their intention was shown by their bodily movements, as it were the natural language of all peoples: the expression on 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, rejecting or rejecting something (PI 1).

If Wittgenstein's demolition of the Augustinian picture is conceptual, we want 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 irrefutable, 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.[8]

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 than prioritise linguaform, we ask how, on a given occasion, a goal-driven agent can set 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 on 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 material practices that this labelling game 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.[9] As Sparrow and Cowley (2002) 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

[8] Halliday and Matthiessen (1999) take just such a view as the cognitive grounding for systemic-functional theory. From our perspective this is no improvement on the state 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.

[9] 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 abstracts. In response, Cowley (2002) argues that this bottleneck may not be 'in the brain'.

leads to insomnia and a loss of memory. However, to cure loss of memory, the townspeople begin to spend night hours labelling things. So they put a big sign at the entrance to the village saying Macondo and, right next to it, being Catalan, a sign 'God Exists'. This all goes well. Entire households are overrun by little slips of paper, until the people begin to forget what each item is used for. After a time, this leads to, for now 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, that, 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 nodes with, above all, persons and external 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 on 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 enculturated 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 predictable; 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, indexes to a baby's *Umwelt*. Where it uses discrete entities ('mama', 'mummy'), it exploits an ability to hear abstracts that help shape social life. If, as we see it, we abandon a view of human rationality, then, we must examine human mentality with respect to its dynamics. To reach beyond skin and skull, our cognitive ancestors 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 words 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, neural 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 different 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 mid-level intelligence. Following Brooks (1991a; 1991b) and others, we accept that, in designing robots, it is dangerous to begin with *a priori* forms of 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 arises—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 the grid 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 this 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 interactive 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 1994:51). More surprisingly, since Tetris rotations and translations occur when a zoid moves into 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 processed by their epistemic counterparts. Both, therefore, are intrinsically cognitive. Rather than deriving from motor plans, the actions themselves are used to uncover 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 reading 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' is to be exclusively in the head. We also exploit actions that – by traditional criteria – exist in the world.

And Is Language 'out there'

The words used in Tetris are not inside. As we saw, in using these shapes, cognitive processes exploit actions. This enables us to see parallels between how we act with Tetris words and how we deal with people. With these questions in mind, we now bring out limitations in Clark's (1997; 1998) view of language. At times, we think that any linguaform externalization also has 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 passive and distributed nature of brain processes. Rather than compare acting with words and acting with people, he adopts a linguaform view that effectively treats words like end-items. Failing to consider how talk loops in and out of bodies, language is reified as part of the external world. Anything that is an external prop, which ingloriously 'do[es] cognitive[ly] cognition', he simply overlooks the question of how we find our way to exploiting this way of using words and language.

Clark (1998:166) looks at language as a transformer. It is a cognitive resource that 'allows patient 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 exploring 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 as (pitch) 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 she learns to put on a tie is—not saying something known but directing complex action. This illustrates activity is used to remember experience (provided by a parent or teacher) and, in so doing, to direct 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.* 1986:460).

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 contextualising is required. Success is more likely if the child uses a rhyme to suggest how actions are sequenced ('one, two, under, through') than if she uses linguaform to remember explicit instructions. While capturing the 'super-communicational' role of language and the importance of external facts, a focus on private speech overplays the world-like aspect of language. Thus, the bias tends to drive a wedge between discursive and other cognitive. If language loops previous experience with current goals, as we all function – 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 rooted *a priori*, talk, thinking and understanding must arise as sub-personal mechanisms direct, among other things, perception and action. In extending talk, then, we ask how events resemble a Tetris player's moves with zoids, or how acting interactively, 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 probed at the wrong level of description, we also stress that language stand-ins are unlike linguaforms. 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 encapsulation 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 turn, 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 first-person and vice versa). In particular, we ask how language loops across the skull. Since inner-a 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, Dennett-an 'real patterns');
- In utterance-activity, as in Tetris, we exhibit dynamical recall 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 sequentially. 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 passing 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 mutually shapes 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, during babbling, children develop 'gut nuts' towards their first birthday and 'combinators' 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, as 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 Uexküll 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 (I-U) and, of course, how it meshes with subsequent (U-I) 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 void 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 claim, contributes to why talk is episodic. Such a structure enables talking persons to engage in expansive 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 bluff), 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 context 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 micro-timing 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

of 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 orders, city centres, restaurants. Our claim is that these kinds of sign endow 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 or fortyfive utterance and thus invites the person to perceptual constancy (Clark 1998: 20). Of course, to learn, say, 'chastity' or 'ethnicity' we need recourse to other concepts. However, while Linguaform items as 'map' '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 with and linguaform 'charity' and 'charitable' activities.

In short, alongside stable and decontextualised linguaform properties, we must expect to find P-C and C-I microtemporal phenomena. Alongside its modular aspect, language knotties 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 perceptual action boundaries simultaneously in two directions. From the P-C side, we seek something analogous to gestures and movements as words are morphed away from their central senses and used to pick up information from other persons. On the C-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 improper pick-up information. Close analysis of utterance-activity not only reveals just this kind of organisation, but strikingly, shows its real-time dynamics to operate on 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 ordinary talk, we examine the former in detail. In succeeding, we look at events in a time domain where utterance activity crosses the boundaries of skill and skill. To avoid any process view, this 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 temporarily embedded context (see Keenan 1976). In other words, we examine conversation as coordinating (see Harris 1997; Crowley 2007b). In so doing, we stress this talk recruits to two or more brains (interactants P and I) activity that crosses human body boundaries.

I-I. Consider events from an Italian family conversation (for further description see Crowley 1995, 1996). While driving a rental car, Rosa tells to why a husband, Aldo (A), failed to cut his wife, Rosa (R), the pea-pods that she had requested. In raising this, Rosa is heard as complaining in her husband's hearing – to their daughter Monica (M). Hoping for assistance 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 as, paraphrasing, that 'a certain person is too lazy to cut pea-pods'. As it turns out, this is a 'false' re-membering of the events. In fact, as Aldo says (and she accepts), it did cut 15 min pods. This characteristic example of collective remembering clearly serves both social and self-personal ends. Not surprisingly, it comes nowhere near satisfying Rosa. Quite the contrary. Speaking verbally she indirectly acknowledges her mistake and says:

English version	Italian original
(10) R: You're right. They were you should have seen the pods or they're longer than this room if not longer.	R: Affatti se vedessi le baccheree ar, son piu' lunghe di questa camera (tsk) Non piu'

No brain that relied on complete pieces of linguaform could process the utterance-activity that crosses from C 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-C boundary crossing by causal chains of information that realise grammatical rules or functional goals. Far from being discretely planned or goal directed, the utterance is a rapid

move in a family game. Like a Tetris relation, 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:

(7) A: Can't do it	A: I Don't
(8) M: Course it is	M: I Don't
(9) A: (if) not longer	A: Non pin'
(10) M: No	M: Ya

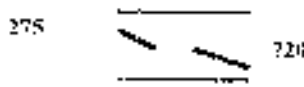
The mainline understanding effectively dismisses Rosa's talk as absurd. However, instead of being explicitly articulated, no use is made of logic or grammar. Instead, the contrastive stress on the face of it, as in claiming that the new toilet was bigger than the 4 metre room) evokes what Goffman (1978) calls a "response cry". What is transcribed as 'non' (and translated as 'come to it') 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 *dislocate Rosa, good sense ably and in harmony*. To grasp this, we need appeal neither to internal processes nor to manipulations of linguaforms. Rather, we argue that they exploit dynamic patterns that spread between the individuals concerned. Just as in Tetris, they rely on expert real-time 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 can't do more pea-pods is designed neither as part of a story nor as a propositional claim. Rather, it is action designed to probe attitudes that, as Mary responds, is like tapping a card. Of course, in making this IU move Rosa's speaking broaches her husband's IU boundary. Instead of hearing what she wants, he responds with what is transcribed as 'non'. Although saying nothing, saying 'non' alters each person's state of mind for he hastens to call for epistemic. Apart from anything else, as shown below, it prods the daughter into action response. Before examining in cognition in real time, we stress that talk can be jazzed able to successive speech acts. Often, what happens is simultaneously pragmatic, epistemic and perceptual action that primes and provokes further goings on.

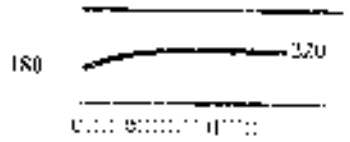
The talk also exemplifies a social payoff that arises in managing family members. On this occasion, Rosa gets what husband and daughter regard as not

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, as Clark's (1997:170) terms cited above, in "[t]ags the presence of underlying structure inviting perceptual commonality". What we add is that this enforces subtle properties onto a heterogeneous class of things, so that saying 'non' is a practice which, among other things, enables one to label something (or someone) as absurd. Even if no such word appears in conversation, the activity is constituted by, at times, saying 'non' while looking and gesturing in specific (Tetris) ways. Cognitive action spreads across the environment, even if, necessary, it seeps into individual brains.

Turning to how these abstract descriptions play out in moment-by-moment 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 pin' is represented iconically and measures are given for its acoustic peak and minimum fundamental frequency (the 'p' has no pitch).[9]

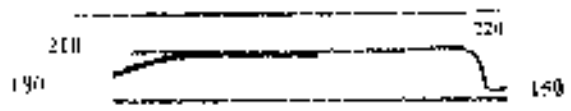


The message and the dynamic features of her, speaking prod Aido into his response cry. However, instead of using his usual (male) pitch range, he raises his voice into a female domain and says (duration is about 760 msec):



[9] All measures were made on a Kay sonagraph. For details, see Crowley (1998).

Given that Alde starts speaking during the 'pi' of 'pick', the meshing cannot be planned. Rather, the overlap shows that some kind of perception-action mechanism allows Jim to orient to the pitch of Rosa's voice. In producing 'beer' 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 entrapragmatic 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 'beer' but, as the acoustic record shows, orients to the same target. Both voices are shown below (Alde with ——; Monica with - - -)



time (sec) ... 1.0 1.5 2.0 2.5

From one perspective, this is Alde and Monica 'ridiculing' Rosa good naturedly and in harmony. Saying beer as described in these circumstances, is not rude. They ridicule Rosa insofar as how they are to speak, yield their voices meaningfully at hers. The 'good nature' is physically-based but clearly 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 Crowley 1990, 1994; 1997; 1998; Cooper-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 words, it relies on real-time responding. Third, while the example is useful because it occurs on 'beer', similar efforts 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 Tonis, 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 inexplicable in word-based patterns.

6.0 Microcognition, Talents 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 Tonis player's perceptual and corporeal nexus. 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 Donald (1969) implies, we think it likely that there is a special relationship between the human voice and the meaning 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 Marques and Wittgenstein, labels float on a sea of fleeting 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 at 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 largely unnoticed practices. Thus it has come to pass that activities such as making coffee connect – even if we know nothing about it – with sociohistorically based practices like talking cows and taking 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 as waves 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 Mazmalo. 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 generated – or by definition – out 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 human activities. Since they lend themselves to decontextualisation (we use a label for many purposes), they encourage cognitive reflexivity and sustain the myth that language is identical to cognition. However, as argued, the surface attention of labels is guaranteed only by literacy that extends bodily capacities and social practices. While labels indeed name objects, these are used for purposes that are inseparable from the patterns of life and 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 enculturated 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 prosody and visible activity, these must be basic to human sense making. It is this unlike 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 our past experiences. With Christiansen (in preparation), language *seems* to be vocal or written. 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. Affects nurtured in poetry, just as a child uses a rhyme in knotting its tie ('over, two, under through'), the businessman uses patterns to re-mind us and invite (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 crisscrossings of normal activity. While folk psychology would tell us that memories are attached explicitly to words, we see this as a cognitivist interpretation. 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 imputation 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, micro-temporal vocal dynamics, rather than words, are central to social goals. However, in both cases, C-U 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, biases are used and abused.

Practices exploiting labels both rely on sub-personal processes and the patterns of public behaviour. Much understanding, we think, depends rather on logic nor inferences but the co-ordination of behaviour. While Marques 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 (sensorimotor 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 rule governing early LFs. By de-emphasising written signs and monological ways of talking, we find reason to doubt that linguaform allows the "total installation of a new kind of computational device" (Dennett 1991: 210). While form-based language may lead, metaphorically, to mental 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 rapid recovery 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 capacities". Importantly, these include a capacity to respond quasi-reflexively to what we know, in part, in the body-as-Wittgenstein's-child. We have learnt this to how brain/body systems negotiate interpersonal life. Acting together, brains, bodies and the environment constitute enculturated 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', 'clarity', or even 'mouse'. 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. 1984. *The Contextualization of Language*. Amsterdam: John Benjamins.

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

Brooks, R.A. 1991b. Intelligence without Reason. A. Memo No. 1293, MIT AI Lab. (In: Proceedings of the 12th *International Joint Conference on Artificial Intelligence*. Morgan Kaufmann.). 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.ling.ohio-state.edu/~morten/cv/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.

Crowley, SJ 1991. The Place of Prosody in Italian Conversations. Unpublished PhD dissertation, University Of Cambridge.

Crowley SJ 1994. Conversational Functions of Rhythmical Patterning: A Behavioral Perspective. *Language and Communication* 14.4: 353- 376.

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

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

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

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

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

Crowley SJ & D Spurrett 2003. Purring Apes (Body and Language) Together Again. *Language Sciences* 25:289-316.

Crowley, SJ 2003. Distributed Cognition at Three Months: Mother-Infant Dyads in KwaZulu Natal. *Alternation* 10,2: 229 - 257.

Crowley, SJ, S Moodley & A Fiori-Crowley (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.

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

Dennett, 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 Matthiessen 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. (eds): *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol 2: Psychological and Biological Models*. Cambridge MA: MIT Press.

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

Sparrell, 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 Braten, 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 Uexkull, J 1934. A Stroll through the Worlds of 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.

A Communication Model for Industrial Theatre as a Negotiated Dramaturgy

Gary Mersham and
Gavin Baker

1. Introduction

This article proposes a model that effectively represents the communication process of the negotiated dramaturgy within the context of theatre performance. It begins by defining the negotiated dramaturgy within the context of industrial theatre before investigating several models of 'theatre as communication'. It analyses these models in terms of their applicability to the process of the negotiated dramaturgy. Finally, it proposes a model tailored for the negotiated dramaturgy developed by the authors.

2. Industrial Theatre—A Definition

Industrial Theatre can take on many different forms and have different objectives within an organisation. People generally use the term 'Industrial Theatre' with very little consensus regarding its meaning or function. It is used, in general, to define any theatre that occurs in an industrial setting. This broad categorisation ranges from a play informing people about HIV/AIDS, to actors role playing situations as part of staff training, to a piece of theatre that informs the aims and facilitates more effective functioning of that particular organisation. Practitioners use it to achieve different objectives related to, for example, issue management, internal communications, promotion, advertising and awareness campaigns. The circumstances that it addresses are industry and issue specific. For example, Industrial Theatre may be defined as the use of drama and drama techniques to create learning and change in business environments. This is achieved by creating synergy between the disciplines of drama, organisational development and psychology in the design and

practice. It ranges from Learning Theatre Organisation 2000). Industrial Theatre has also been referred as the use of drama and actors within a participative setting to sell a product or service, to put across a concept, to raise people's awareness and to enhance training (Actors Mean Business (2000) http://www.actorsmeanbusiness.com/industrial.htm).

The tradition of using song, dance, mime and theatre as a form of conflict resolution is as old as the history of civilisation. In Africa this tradition is deeply rooted in the culture of its people. Mbigi argues that 'In Africa you cannot introduce change with a memo you have to get the people emotionally involved in the process' (Khumalo *et al.* 2001:307). He calls this using the Ujamaa platform, a process of problem solving used in African cultures which calls for more group and process oriented than traditional Western practices. Communication theory tells us that to successfully communicate we must use the signs, symbols and cultural codes which are relevant to participants in the communication process.

Therefore, the model of Industrial Theatre set out here—that of a negotiated dramaturgy—is based on the two-way interactive nature of communication and the role of the theatre practitioner as a mediator and facilitator in the process of developing the negotiated dramaturgy within a specific set of cultural and social contexts.

3. The Development of a Negotiated Dramaturgy

The process of Industrial Theatre as a negotiated dramaturgy has five distinct stages. The first is the initial meeting which defines the mandate and the second the work of the small group forums in the creative-dramatic frame where one identifies the issues and where individual stakeholders negotiate the objectives and outcomes. The third stage is the rehearsal process where the practitioners create and structure the dramaturgy. The fourth stage is the dramaturgy (i.e., the interactive presentation within the dramatic frame). The stakeholders settle some of the issues in the forums while the practitioners deal with these and other issues in the dramaturgy. Within the dramaturgy stage one refers to 'in role forums' where the attributes of a forum are present (i.e., this case opportunities for free and open communication) but within the dramatic frame. The fifth and final stage is the 'feedback' stage where the practitioners report on the process to all stakeholders, not only to the organisation (usually management). This includes the plans of action, solutions and suggestions. We now look more closely at each of these stages.

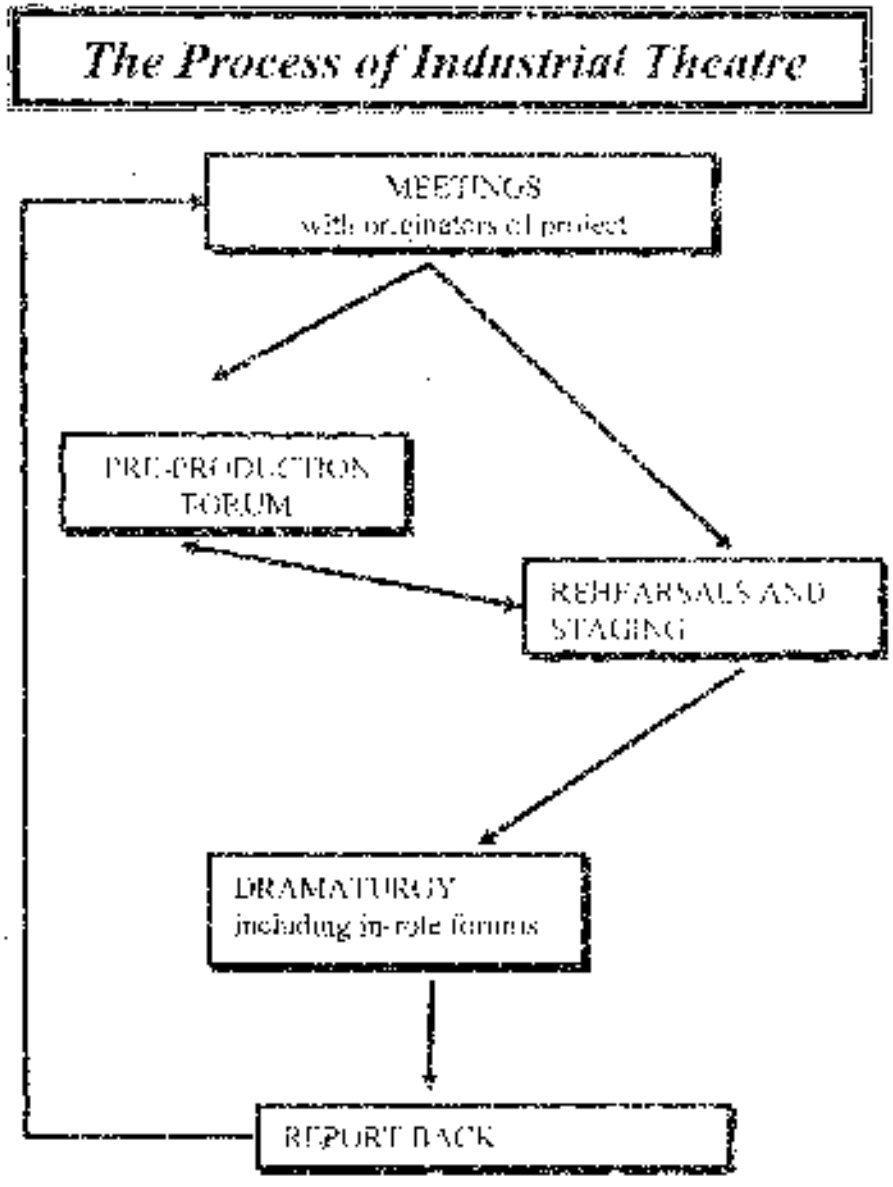


Figure 1: The Process of a Negotiated Dramaturgy

3.1 The Initial Meeting

The initial stage is the original meeting in which a particular organisation gives a mandate to a particular company to initiate an Industrial Theatre Programme. They present the topic and issues on which the practitioners will design their approach. At this meeting, the initiators provide the practitioners with information regarding the timing, issues, possible approaches, media and the sector of stakeholders involved in the process.

3.2 The Pre-production Forums

This stage involves the broadening of the stakeholder base to include all concerned sectors of the organisation. The stakeholders divide into small groups in an attempt to create synergies (Narayan *et al.* 1996:4). The practitioners keep the groups small to encourage the formation of interpersonal relationships among members. The practitioners facilitate discussion allowing for diversity of opinions and ideas while allowing all the stakeholders to present their views on the issues.

The forums attempt to resolve the issues within their small groups while gathering material for the dramaturgy. For example, when discussing the possible ways a supervisor character should act in the play, they are in fact debating the overall operational procedure of supervisors. The practitioners should ensure that while the participants discuss the negative or incorrect operational procedure, they establish the correct procedure as well. Often the different sectors of stakeholders might disagree or have a different approach. The forum can encourage the practitioners to facilitate a negotiated settlement.

Often a misunderstanding or lack of knowledge of the situation is responsible for the objective of the campaign. Practical experience has found that contextualising the play and, in particular, the creation of the characters, helps establish better understanding of and empathy for each other. For example, management might be unaware of the domestic situation of many of the workers. The situation directly or indirectly affects the productivity of the worker. This knowledge might create the desired empathy and understanding that the stakeholders may use to help resolve an issue or improve conditions. This exercise encourages the participant to become aware of the 'sociocultural and autobiographical circumstances' discussed later in this article.

3.3 The Creative Process

This process has its roots in the collective creative processes of production

rather than the more 'traditional' approach where a director interprets a playwright's text and creates a production through a particular production process.

The subject to creativity responsible for the creation of the negotiated dramaturgy addresses the 'tangential elements' that according to Hauptfleisch (1997:109) influences the communicative potential of the 'total performance event'. The stakeholders' realisation that they are part of the creative process enables them to identify with the 'image of the author' (themselves). The practitioner's performance within this creative process (his or her demonstrations of characterisation and use of artefacts in the pre-production forums) allows the stakeholders to identify with the 'image of the performer'. The site-specific venue directly relating to the objectives of the campaign and to the working environment of the stakeholders addresses the 'image of the place of performance' (the canteen or workshop or factory floor, for example). The stakeholders' knowledge of and participation in the creative process of the campaign as well as the intended perception that the campaign is for the benefit of the entire organisation, addresses the 'image of the audience'.

The practitioners use the principles of creative drama as the catalyst to create the dramaturgy. Creative drama is usually associated with children and the notion of 'play'. However, its principles are applicable to the creative process. Creative drama is a dramatic learning activity, guided by a leader that allows participants to imagine, enact, and reflect upon real or imagined human experiences. Pinciotti (1993) and Woodson (1999) argue that it nurtures both individual and group skills, enhancing the participants' ability to communicate their ideas, images and feelings with others through dramatic action. The stakeholders share ideas as their imaginations define the story, the setting and the characters. This process is highly theatrical as practitioners encourage the participants to demonstrate (act out) their ideas rather than just talking about them. Personal experience shows that participants find it easier to express themselves in this way. For example, when a stakeholder struggles to share an idea or wants to reinforce its potential he or she often gets up and acts out the idea, sometimes alone, sometimes using others in the group. In many instances, this spurs the creativity of the others, thereby guarding the process. Undoubtedly this varies from group to group but generally speaking this also develops high levels of cohesion (Moore 1997:84-93). It is essential for the practitioner to create this cohesion. From the outset, the various practitioners might see themselves as opposing each other (for example, management and trade unions). Johnston (1998:8-10) argues that practitioners can improv

one scene. These scenes become individual units within the overall structure of the dramaturgy. The scenes operate as if the narrative running through the entire dramaturgy was independent of time. This structure is likened to the episodic nature of Brecht's theatre (Brecht 1964:179; Mitter 1992:64).

Brecht's Epic Theatre makes use of closed 'playable' plays structured episodically and focuses on a moral dilemma while allowing for vital questions to be unconditionally aired with a view to their resolution (Brecht 1964:76; Counsell 1996:82; Styan 1981:140). This approach allows the audience the opportunity to make its own judgement (Brecht 1964:71). Brecht's theatre presents a structure that allows practitioners to use theatre for public discussion (Brecht 1964:130f; Styan 1981:126f).

The Negotiated Dramaturgy shares the intentions of Brecht's epic theatre. It is therefore inspired by its structure. Each scene may be a play in itself and not necessarily be one scene in an overall play, or the scenes may refer to each other in order to establish cohesion. The practitioners style each scene as a 'well made play' using the Aristotelian concept of a beginning, middle and an end. These scenes have a horizontal organisation structure (Hauptfleisch 1978:39), meaning that they are linear in their progression, ie, the commencement of the dramatic events; the development of the dramatic events; and the conclusion of the dramatic events.

Morgan and Saxton (1987:5-7) operating from an educational drama context list four stages in the linear progression, namely: 'exposition', 'rising action/complication'; 'climax/crisis' and 'denouement'. The negotiated dramaturgy approach favours this context as it clarifies the process and enables the participant to become more actively involved. Further clarification is gained by understanding that the 'rising action/complication' is achieved through dramatic tension. Listing the four stages mentioned above are: task; relationship; surprise; and mystery (O'Toole & Haseman 1986:19-39) creates a foundation from which to create the dramaturgy.

These self-contained episodic scenes form part of a larger structure that is styled on Boal's Forum Theatre. Using Boal's (1992:180) 'rules of the game' to formulate the structure of the dramaturgy it should have the following characteristics.

The dramaturgy must clearly represent the nature of each character, identifying them precisely and accurately, so that the audience can easily recognise the ideas and beliefs of each character.

The solutions proposed by the practitioner must contain at the very least one unconsciously devised workplace error which will be analysed during

action by allowing the stakeholders to become aware of their similarities and of what they have in common (the general good of the organisation and what each person gains from it) and the need to focus on a common goal (the creation of the story and characters).

The pre-production forums establish the needs, formulate the objectives, formulate the message and establish the time frames. The participants create the plot and the characters to deal with the needs, objectives and message of the campaign, while also defining the dialogue. They then pass these on to the practitioners who use them in the rehearsal process when structuring and rehearsing the dramaturgy.

The practitioners rehearse the dramaturgy and conduct the pre-production forums simultaneously. This creates a 'feedback' channel between the stakeholders and the practitioners that allows for constant interaction on the story and characters as well on the technical correctness of the action and terminology.

3.1 Rehearsals and Staging

The practitioners base the story line and characters on the objectives established in the pre-production forums.

This dramaturgy is designed to give substance and form to the message (Rhodesh 1986:7), while evoking 'feedback' on the negotiation issues. It functions as a three-dimensional demonstration incorporating speech, movement and sound, concerning the positive and negative attributes of the negotiation issues (De Marinis 1982:137, Williams 1996:131f). The dramaturgy allows the stakeholders to become the communicators with the same status and credibility as the 'original' communicator(s) or initiators of the project (usually management). Management becomes one of several stakeholders.

3.2 The Structure

The structure of the dramaturgy is vital for effectively communicating the intention of the message. Hauptfleisch (1978:37) states that the playwright manipulates the 'vocabulary' of a play, particularly the scenes and episodes within the total plot, for a variety of purposes. In Industrial Theatre, the aim of the dramaturgy is to share an intended message with the stakeholders. The message (decided upon by the stakeholders) is broken down into objectives, which form the basis of each scene. Similar objectives are grouped together in

the in-role forum. The play must present a mistake or a failure, inducing the stakeholders to find problem solutions and inventing new ways of confronting the issue. The presenters must pose good questions, allowing the stakeholders to supply good answers. This action must open the channels of communication, creating a free, enabling environment within which all parties can participate.

The negotiated dramaturgy can be of any genre. Surrealism or the irrational should, however, be avoided as the participants find these styles difficult to understand, thereby limiting their chances of becoming actively involved in the process. This tends to alienate the audience, causing them to struggle to understand the objectives rather than being able to concentrate on the issues that the objectives present. Therefore, the style does not matter, as long as the objective is to theatre, through the medium of theatre, concrete solutions.

Practical experience shows that limiting each scene to two or three characters is the ideal. It keeps the scenes simplistic while enabling sufficient character development to deal with the issues and objectives.

3.6 The Characters

The characters are dramatic creations of typical people found within the organisation. Often they are stereotypes and caricatures designed to appeal to as many of the stakeholders as possible. Experience shows that these characters should be blatant in their actions and their intentions. The characterisation must be clear to the diverse audience and represent an employee of the organisation. The concept of Brecht's 'Alienation Effect' assists the practitioners in this style of characterisation. Today scholars prefer

[1] For example, the researchers conducted a campaign at Richards Bay Coal Terminal (RBCT) dealing with 'Value in Diversity'. In this campaign, the stakeholders collectively created characters belonging to management and unions. The characters were created from their own experiences, drawn from their backgrounds and facing the issues that they had to deal with as individuals and as members of the organisation. This specific detail was used to create a 'stereotypical' character that would be found at RBCT. The characters were identifiable to all the stakeholders. This also assisted the stakeholders with understanding each other and went a long way to establishing empathy for each other. This technique was instrumental in the campaign successfully achieving its objective.

the term 'distancing effect', as many believe that 'alienation' is an inaccurate translation of the German Verfremdungseffekt.

Brecht's 'distancing effect' draws the audience's attention to a particular point or object (Brecht 1964:143). Quite by eliminating the 'magic' of the theatre. Often theatre is associated with the creation of a magic reality in which the audience sits back and enjoys the spectacle. The magic of the theatre absorbs the audience and therefore they are not required to think about what they are watching. The 'distancing effect' encourages the audience to adopt an attitude of enquiry and criticism in their approach to the story of the play (Brecht 1964:136; Counsell 1996:102; Mitter 1992:44).

3.7 The Dialogue

The choice of dialogue is very important for the success of the campaign as a whole. This choice is more involved than just choosing a language, for example English and/or isiZulu. The participants need to choose a medium that guarantees the highest level of communicative success. Hauptfleisch (1997:89) outlines three premises that shape dialogue in South African theatre. These are applicable to Industrial Theatre and need to be considered by the practitioners when devising the dramaturgy.

The first premise concerns the general attributes of dramatic dialogue. His third premise is that:

> • Dialogue on stage is artificial, it is a distillation of and selection from everyday language for the purposes of communicating a specific message under particular circumstances.
> • Dialogue in performance is an integral part of the single communicative transaction.
> • The playwright is not the sole creator of his dialogue form – it is also determined by a number of external, social and cultural factors.

The second premise states that the nature of dramatic dialogue is shaped in part by three demands made by the dramatic form itself:

> • a play has to communicate its message directly, by aural and visual means, to an audience at the very first exposure;
> • Dialogue needs to be 'performable', i.e. it is to be a spoken language.

> • dialogue in a performance must be understandable to an audience made up of a wide spectrum of the general public, and having a variety of backgrounds

His third principle is that:

> Normal, everyday language as spoken by the average man in the street, has a very limited range of expression. The aim of any artist – particularly a verbal artist – is to transcend the limitations of 'normal' human communication, to somehow say more than words can. Hence the enormous weight given to the non-verbal elements of performance, and hence too, to the basic absurdity of an enterprise which aims at being so much more than a mere mirror to be held up to 'nature' (Hauptfleisch 1997:89).

Hauptfleisch (1997:93-95) notes two types of dialogue found in South African theatre. He refers to 'tsotsitaal' as the dynamic ever evolving language spoken on the streets of South Africa, and 'theatretalk' as the language used in theatres in South Africa. His argument is that theatre is 'an artificial representation of life, not life itself' and that a 'play is a defined, purposely structured world created for a specific communicative purpose' (Hauptfleisch 1978:80, 1997:93-95). Playwrights manipulate language usage to show diversity within one language code, for example using Standard South African English and Black South African English to differentiate between the race or socio-cultural differences of the characters. The context of the negotiated dramaturgy is different to the context in which Hauptfleisch (1997:93-95) bases his argument. Industrial Theatre is intended for a multicultural, multilingual (and often a socio-economically diverse) audience. For this reason, it makes use of elaborated codes to facilitate sharing the message. Practitioners are encouraged to use code switching within the dramaturgy to reach this diverse audience. This puts greater emphasis on the non-verbal codes used in the dramaturgy thereby improving the chances of the message being understood. Added to this is the need for terminology specific to the organisation, essential for contextualising the dramaturgy and for achieving the objectives of the campaign. For the purposes of this study, Hauptfleisch's (1997:94) 'theatretalk' can evolve to become 'industrial theatretalk'.

3.8 Performance Space

Generally speaking, the performance space is a site-specific venue within the organisation, for example the canteen, training centre or shop floor. A physical theatre (building) is not essential in Industrial Theatre, (nor is any theatre, for that matter). As Black Mkhwane of Sinc Clarence productions puts it "it doesn't have to happen on a stage with lights. It can happen under a tree" (cited in Bell & Seery 2000:16). Hauptfleisch (1978:126,182) argues that "drama takes place wherever an actor and an audience meet".

One of Peter Brook's (1972:11) most famous statements was that theatre could take place in any "empty space". Within this mindset, the negotiated dramaturgy makes use of Boal's (1995:16) concept of the "aesthetic space". Boal (1995:181) states that:

> All that is required is that, within the bounds of a certain space, spectators and actors designate a more restrictive space as 'stage' or 'aesthetic space'. This space may be a corner or centre of a room, or floor level, or it may be a raised space like the back of a truck. The space may be stationary or it may be mobile. For example, the researchers presented a campaign from the top of a flat-bedded truck moving around the site from one section to another.

Hauptfleisch (1978:183) and Heilbo *et al.* (1987:9) argue that the relationship of the audience to the actors and the physical environment can be a significant factor in the overall meaning of the play. This is particularly relevant to the Negotiated Dramaturgy, as the stakeholders are actively involved in it. They should have the sense that the 'aesthetic space' belongs to them and that they are part of that space. The stakeholders' participation in the in-role forums (whether from the stands or in the space provided for the practitioners) automatically include them in the 'aesthetic space'. Within the creative drama approach at the pre-production forums the 'aesthetic space' is constantly changing too. As the participants act out their ideas, their immediate space becomes the 'aesthetic space'.

3.9 The 'Report back' Stage

The final stage of the negotiated dramaturgy is the 'report back' stage where the practitioners report on the successes and or failures of the process to the originators. They comment on the issues raised during the pre-production forums and in the dramaturgy. This includes the plans of action, solutions and suggestions. The type of 'report back' varies according to the 'brief' given to the practitioners. The 'Report back' may be an informal discussion where the

relevant points are raised, a detailed written report that discusses the experiences and observations of the practitioners (This can include video and 35mm or digital still photography). It can also include conducting post-performance research to determine the effectiveness of the campaign, using questionnaires and compiling a research document that addresses the objectives of the campaign in detail.

4. Performance Studies

The investigation of theatre as communication has conventionally fallen under the ambit of performance studies. This area of study investigates the meanings that theatre creates. It encourages the study of meaning from both the senders' and the recipients' point of view. Helbo *et al* (1987:5-7) divide the investigation up into the following categories:

Text: This is either the written script that precedes the performance or the spoken form in relation to other performance codes. The text is repeatable and enduring, which allows it to be transformed by the performance into voice, which is an ephemeral phenomenon. This transformation justifies the distinction between dramatic (written) text designed to be read, production text (stage direction and dramaturgy), and theatre text (the ensemble designed to be performed).

Speech: Speech poses the problematical question of the power of language. It therefore has to be approached through its interweaving with other performance codes such as gesture, facial expressions and props.

Stage Design: Included in stage design is lighting and sound effects. Lighting can be used to focus the audience attention on a certain part of the stage while drawing their attention away from another. The ongoing technological advancement of stage lighting has considerably increased the impact that lighting has on theatrical expression. Sound effects create a mood and impose a rhythm. It can also be used to structure space and punctuate a performance. The set encompasses the overall scenography of the theatrical event. Sets can be mimetic or symbolic. They can have a dynamic function or static function, or can be used to demonstrate the plasticity of the human body or it can be dispensed with entirely (cf. Braun 1982 & Grotowski 1968).

Stage/Audience Relationship: The interrelation between the performance space and the audience raised questions concerning theatre aesthetics and actor training. The performance may or may not demand an active response from the audience. The emphasis in actor training varies as a consequence of the performance functions desired: explanation through theatre of the mechanics of everyday life, translation of universal myths, arousal of emotions, exploration of the self, improvisation, acceptance or refusal of chance occurrences to mention a few.

The preceding categories are based in an empirical listing of detail concerned with the material substance of the performance rather than the object of knowledge. This approach largely ignores exploration of the signifying relationship that takes place as the spectator constructs meaning by making connections across the spatio-temporal axis of the performance and develops structures of congruity. This approach further suggests that the theatrical sign is constructed and defined exclusively through the prior existence of the performance tradition. The first three categories are primarily concerned with the directorial vision, and do not establish how the performance object is constructed and at what level it exists. The latter investigation is at most philosophical and aesthetic, yet it has attracted many theatrical scholars. One such group of scholars was collectively known as the Prague Circle.

4.1 The Prague Circle

Initial research into the systematic theorisation of performance was conducted by members of the Prague Circle. These scholars focused on the semiotisation inherent in the theatrical phenomenon. Honzl (qtd:118-127) rejected approaches restricted to the material reality of the stage. He argued that,

> ... total art can be seen to negate theatrical expression; the latter is ultimately no more than the sum, the juxtaposition, the "co-ordinated presentation" of a number of material forms: music, text, actor, décor, props, lighting. The principle of total art however, involves recognition that the impact of theatrical expression, in other words the strength of the impression received by the spectator is a direct function of the number of percentages flowing simultaneously to the mind and senses of the spectator.

Veltrusky (1976) argued that the process whereby all stage signs are rendered artificial is the basis for the transformation of all phenomena marked by theatrical convention into intentional signs. He argued that all events in the theatre are necessarily 'resemanticised' by the spectator. For example, the spectator perceives an unintentional sign such as a stutter or a scratch of the head as meaningful.

Bogatyrev (1971:517-530) reinforced the concept of stage semiotics through his notion of the excess of meaning inherent in theatrical signs. He argues that this is what distinguishes theatrical signs from signs of everyday life. Mukařovsky (1934/3:10; 1978) also studied the theatrical sign and believed that the performance signifier or 'text' was associated with a signified established by the collective mind of the audience.

The Prague Circle also considered the system of stage meaning and they claimed that the denotative/connotative network was activated dialectically by the actor (Helbo *et al.* 1987:8). The overdetermination of the stage signifier led to the study of theatrical codes. Honzl (nd:118-127) observed the interchangeability of signifiers and the lack of limitations on the class of signifiers to which they may refer. The Prague Circle thus introduced the distinction between static and dynamic codes. These scholars were also interested in the hierarchy of codes in particular the way that meaning is generated and the interplay between verbal and non-verbal codes during the performance. These studies led them to create the concept of layering of codes (Helbo *et al.* 1987:8).

4.1.1 Variations in the Prague School

Following in the linguistic tradition, Georges Mounin (1969) used communication as the frame of reference for his analysis of the theatrical phenomenon. He used communication in the traditional linguistic sense (i.e. the intentional transmission of a message from the sender to recipient, perceived as such and entailing a response through the same channel). Mounin (1970:95) argued that authentic communication does not take place in theatre because communication can only take place if the receiver can respond to the sender through the same channel, in the same code or in a code that can fully translate to the codes of the original message. Mounin cited in De Marinis (1993:139) and Elam (1980:33) argues that for this reason theatre is stimulation and not communication.

This analysis was only applicable to the fictional world or stage. The stage/auditorium relationship seen in this perspective excludes any response

from the spectators other than merely applauding, hissing or booing. This point of view has since been largely abandoned by those scholars who wish to study theatre as a sub-section of the field of communication (Helbo *et al.* 1987:9). Ruffini (1974a:401) argues that in order for communication to take place in theatre, the communicator and recipient must know each other's codes. These codes do not necessarily need to translate, coincide or occur along the same channel (Elam 1980:33). De Marinis (1993:144) argues that '*all* performances intend to communicate *in the same way and to the same degree*, or that *everything* in a performance is meant to communicate' (p. 144). He argues that the audience responds through linguistic, paralinguistic and kinesic signs. The audience is aware of the codes used by the communicators and this enables them to comprehend and interpret the shared message. This act of knowing the codes while not necessarily knowing how to use them (Ruffini 1974a:401) is what categorises theatre as communication (De Marinis 1993:140-143).

The idealistic notion of the gap between the pre-production and the production has been replaced by a materialistic approach *to production* to the performance event. Scholars are investigating the recognition of inversion, aberrant decoding (see Eco 1977; 1978), and the delegation of pleasure (see Helbo 1975, 1979; 1983; 1985). Scholars now emphasise the reciprocal functions of the actor and audience in the theatrical event. They have established that the stage/auditorium relationship is socially situated, meaning that it is linked to a particular audience and its socio-cultural context. Studies have also moved to focus on the language of theatre perceived in its production or reception functions within the context of a shared social experience. This focus created the use of the terms 'performance codes' which are conventions specifically applicable to the performance genre and historical period. 'General codes' are the linguistic, ideological/cultural and perceptual codes used in the investigation, while 'mixed codes' are the general codes that function in a specific performance context. An adequate description would be to describe communication in theatre as an enunciating collective that consists of two elements. The first is a discourse or combination of two enunciative acts. Theatrical discourse establishes a specific genre in that it displays its own rules of operation. It makes these explicitly readable in their own context while disassociating them from the everyday experience. The second is a situation of enunciation that conjures a dynamic set of relationships and contracts (either pre-existing or constructed by the performance) determined by the popular ideology (Helbo *et al.* 1978:9).

Performance studies constitutes the detailed analysis of the systems of production and reception. In the performance studies context, production is concerned with the work of the actor, speech, the relationship between fiction and the physical performance, the use of space and the construction of the performance text. In this context, reception is concerned with the visual composition and juxtaposition, the relationship between the recorded and the visceral emotions, the role of the audience member, communication affby the spectator and the verbalisation of the spectator (Helbo *et al* 1978:13f).

As this article is concerned with the negotiated dramaturgy as corporate communication, it is necessary to investigate the various models that represent the theatrical process and to propose a model that best represents the negotiated dramaturgy.

4.2 Models of Theatre as Communication: The Pfister Models

Pfister (1988:3) graphically represents a unidirectional mode of theatrical communication as experienced by a reader of a dramatic text (figure 2). Pfister (1988:2) distinguishes between a narrative and a dramatic text. He argues that the difference is the speech situation (cf. Hempfer 1973:160-164) as the communicative relationship between author and receiver. He acknowledges Plato's *Republic* as the source of this distinction. Plato first pointed out this difference in Book III (1945:94f) where he discusses the difference between "narration" (reporting) and "imitation" (representation). These actions depend on whether the poet (playwright) himself/herself is speaking or whether it is the character speaking. Pfister (1988:3) states that dramatic texts may be distinguished from narrative texts in that the former are consistently restricted to the representational mode. In this mode, the playwright never allows him or herself to speak directly.

Figure 2: The Pfister Communicative Model for Narrative and Dramatic Texts

In this model, S4 is the actual author in his or her extra-literary role as producer of the work (in example as Henrik Ibsen is the author of the play *Hedda Gabler*. S1 stands for the ideal author, the character that the text implies is the subject of the entire work, which is the actor. S2 is the fictional narrator whose role in the work formulated by the narrative medium. In our example the character Hedda Gabler. S/R1 stands for the fictional characters communicating with each other through dialogue, e.g. Hedda, Tesman and Judge Brack. R2 represents the fictional addressee at S2, while R3 is the implied or ideal recipient of the work, the audience. R4 stands for the recipient who reads the play at that time or all those that read the play at a later stage. The darker area represents the internal communication that occurs during a performance. The lighter shaded area represents the mediated communication system. In dramatic texts the positions of S2 and R2 are left vacant, as there are no actors or audiences involved in this process.

This model does not consider the other factors that influence the communicative process. Pfister (1988:3) has a modification to the above model which considers the other communicative factors (figure 3), i.e. the model is a representation of the external communication of dramatic texts (Pfister 1988:27).

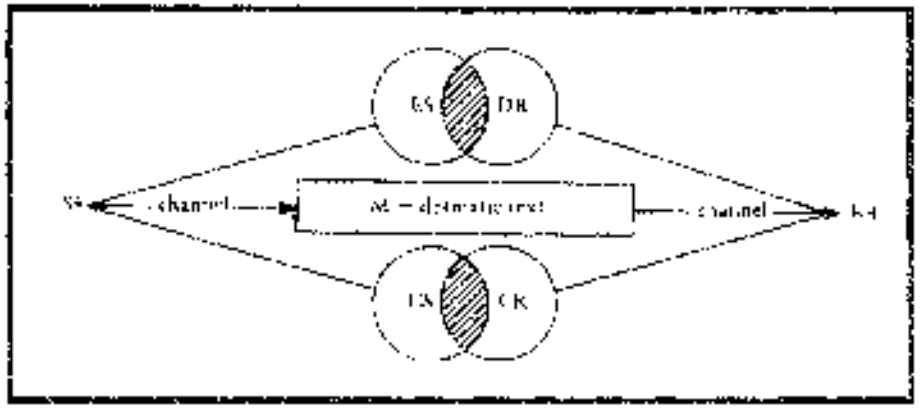


Figure 3: The Pfister Model for External Communication System of Dramatic Texts

In this model, DS stands for the encoding code of the sender and DR for the decoding code of the recipient. CS stands for the content as encoded by the sender and CR for the content decoded by the recipient (Pfister 1988:27).

Pfister (1988:27) argues that the communicative process needs a channel, a message, code and a content as well as a sender and recipient. This channel forms a physical and psychological link between the sender and the recipient. The message is transmitted as a complex of signs along a channel. The codes enable the sender and recipient to encode or decode the message respectively, thus revealing the content of the message. Pfister (1988:27) states that the senders' and recipients' codes are only identical in an ideal context. They overlap in the 'real world' to a greater or lesser degree. Therefore, the content decoded by the recipient is not identical to that encoded by the sender. The message itself is distorted by the channel's own noise.

4.2.1 Critique of the Pfister Models

The Pfister (1988:27) Communicative Model for Narrative and Dramatic Texts is essentially a reader's model. It therefore cannot effectively represent a dynamic process such as the Negotiated Dramaturgy. As this model is the basis for Pfister's Model for the External Communication System of Dramatic Texts (1988:27) it is necessary to analyse it in terms of the Negotiated Dramaturgy.

The dramaturgy in the Negotiated Dramaturgy is a performance. Although initially there is a 'dramatic text' and a 'production text', the dramaturgy is a 'theatre text'. Using Pfister's (1988) terminology, the practitioners are the ideal authors (S1) while the stakeholders are the ideal recipients (R1). The Negotiated Dramaturgy emphasises the use of creative drama, role play and improvisation (see chapter 2) in the creation of the text during the pre-production forums and in the rehearsal stage. The dramaturgy is interactive and dynamic. This makes both the stakeholders and the practitioners the actual authors (S4) of the text. Similarly, both practitioners and stakeholders are the recipients (R4) of the text. This is applicable to both the Pfister (1988) models. The Negotiated Dramaturgy encourages direct communication between the ideal recipients (S3) and the characters during the in-role forums which are part of the dramaturgy stage of the process. The process should be bi-directional, demonstrating a direct link between the recipients (stakeholders) and the authors (practitioners). This uni-directional representation does not reflect the nature of the Negotiated Dramaturgy.

Pfister's (1988) Model for the External Communication System of Dramatic Texts incorporates encoding and decoding essential in any communication. However, the uni-directional nature of this model only allows for the actual authors (S4) to be the encoders and the recipients (R4) the

decoders. The Negotiated Dramaturgy encourages interaction between the practitioners and the stakeholders thereby allowing for a closer 'fit' between content and meaning.

The Negotiated Dramaturgy is a process consisting of five stages described above. The models described above do not effectively represent the process of Negotiated Dramaturgy, even though they have merit when examined as representations of the communicative process of 'traditional' theatre (the context for which they were intended). They The study will now focus on two further models in an attempt to find one that represents the process.

4.2 Models of Theatre as Communication: The Hauptfleisch Models

The contemporary South African theatre theorist Temple Hauptfleisch (1997:3) sees theatre as a system and his diagram, Theatre as a System of Process (figure 3), demonstrates this. Hauptfleisch intended this as a model that 'delimited the domain of "theatre research"' (Hauptfleisch 1997:6). He argues that the model demonstrates how theatre operates as a complex and dynamic structure of inter-linked processes, to generate a particular theatre event within the wider systematic context of a specific community or society' (Hauptfleisch 1997:3). Theatre is an open system, which is constantly changing as it interacts with the larger systems of society. The model consists of numerous elements, which are also processes. These processes are heuristic in nature and encourage individual research. As this model was intended to demonstrate the overall domain of theatrical research, the entire model is too complex for the purposes of this study. It is therefore necessary to delimit the process and to investigate the elements that make up the overall process.

One of these elements is addressed in his earlier work where Hauptfleisch (1978) investigates the play as a method of communication. Hauptfleisch (1978:26) is concerned with the meaning shared in the theatrical encounter. He explains that:

> the *meaning* of a play is the product of the total effect achieved by a control ed transaction which involves all the cues (verbal, non-verbal, situational) provided by the author through the 'play in performance'. The play-in-performance thus becomes a communicative act within a specific temporal and spatial setting, between communicators with

...specific individual and group characteristics, specific role-relations, and certain conventional and shared assumptions (e.t.c.).

Hauptfleisch (1978) visually illustrates his reasoning as follows in figure 4:

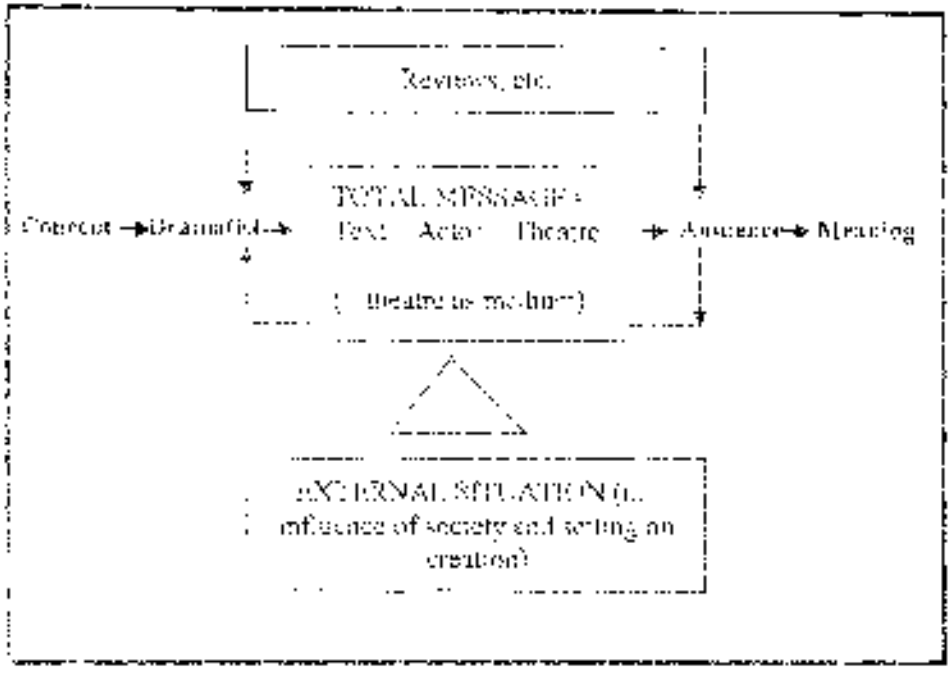


Figure 4 The Hauptfleisch Model of Theatre

In this model, Hauptfleisch (1978:266) suggests that communication between the dramatist and the audience takes place by means of the 'total message' consisting of the textual and theatrical elements. The play in production is central to the analysis for Hauptfleisch. Feedback takes place in two ways. It occurs directly between audience and dramatist by means of reviews, box office returns and book sales. Indirectly, it occurs through the audience response in the theatre itself. The concept includes the story, incidents or myth being communicated as well as the text, both verbal (dialogue, narrative, songs) and non-verbal (theme, plot structure, characters, rhythm, symbols, choreography, subtext). The cultural situation is the physical theatre building

and the stage. It includes the technical qualities and facilities, the atmosphere, reputation, the stage-audience relationships, traditions and comforts etc. The actor represents certain verbal elements (such as improvised dialogue) and non-verbal elements (such as timing, voice, gesture, movement, make-up, wardrobe, lighting, set, props, music, dance spectacle, style). The external situation consists of elements including society, country, age, theatrical tradition and environmental circumstances.

Hauptfleisch (1978:24) believes that theatre is transactional. He argues that the 'total message' is shaped by the audience's influence as they act as a 'homogeneous group of interacting individuals'. He furthermore states that this influence is not only exerted on the communicator (the performer) but also on the audience's own receptivity to the message. Hauptfleisch (1978:241) argues that the awareness of the audience is an essential part of theatrical work. Knowledge of its role is essential to the understanding of its effect in the communicative process.

Hauptfleisch (1978:88) illustrates the transaction between the performer and the audience as follows in figure 5:

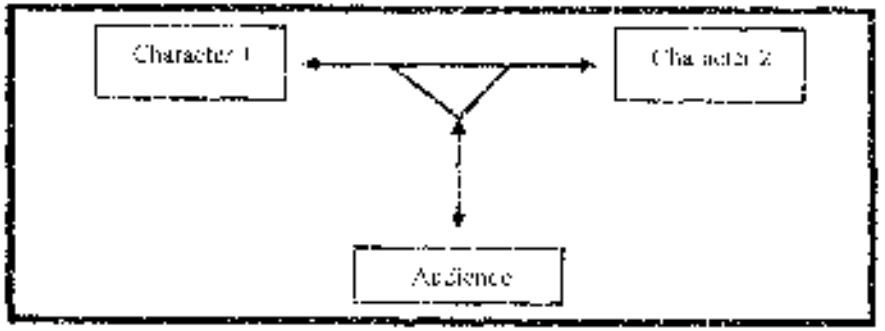


Figure 5: Hauptfleisch's Model of the Transaction between Performer and Audience

This model represents a direct bi-directional communication channel between the characters in the theatrical event. The audience interacts with the channel and not with the characters directly. Hauptfleisch cites Kennedy (1983:11) who proposes this form of transaction between audience and characters. Hauptfleisch (1978:88-89,180) debates whether practitioners are able to circumvent the artificiality of the theatrical event. He believes that if the

characters are able to interact with the as he/she sincerely feels

> the world on the stage is a self sufficient and closed community, with its own conventions for interaction, and this includes linguistic conventions. The relationship between that world and the one inside and outside the auditorium is thus circumstantial rather than direct and/or inevitable.

4.3.1 Critique of the Hauptfleisch Models

Hauptfleisch (1978) created his Model of Theatre by 'translating the general model of human communication' (Hauptfleisch 1978:25). This model is therefore unidirectional, moving from dramatist to audience. The researchers do not dispute Hauptfleisch's emphasis on the importance of the actor[1] in the real-time. One accepts that the actor is an 'iconic sign *par excellence*, a real human being, who has become a sign for a human being' (Elam in Hauptfleisch 1997:56). However, in the Negotiated Dramaturgy the practitioners'[2] role in the Negotiated Dramaturgy is larger than that proposed by Hauptfleisch (1978). He or she is not only a performer but also functions as a facilitator during the pre-production forums. The practitioner is the initiator of the message. He or she interprets, shapes, selects, edits, emphasises and de-emphasises information that constructs the overall message. This includes what type of character he or she will play and what that character will say and do. It is functioning removes the practitioner from the 'total message'.

The practitioner relies on the stakeholders to contribute to the concept and to the creation of the dramaturgy. Hauptfleisch (1978) has acknowledged a link between the audience and the dramatist. The 'internal feedback' reflects the audience's response to the theatre and does not reflect their participation in the creation of the concept and dramaturgy. Hauptfleisch's (1978) 'external feedback' has little effect on the Negotiated Dramaturgy as box office returns and critics' reviews do not have the same impact as they would in 'traditional' theatre. The audience member has been established at the beginning of the campaign and it has no direct bearing on the overall financial profit of the Negotiated Dramaturgy. Similarly, the production is not piloted on good reviews in the media to attract people to come and see the production. Word

[1] This is the 'traditional' concept of the actor as used by Hauptfleisch.

[2] Practitioner in this refers to professional people who work with the Negotiated Dramaturgy.

of mouth is relevant as it may assist with encouraging stakeholders to attend the pre-production forums and encourage stakeholders to attend performances particularly if there is more than one performance of the dramaturgy. In particular, the Negotiated Dramaturgy has a 'feedback' stage where the overall campaign is evaluated. The criteria for this evaluation include aesthetic and functional effectiveness.

Hauptfleisch (1978) recognises the importance of the *external situation* on the 'total message'. He acknowledges the role that society, environment, theatrical tradition and demographics play in the perception of the message. These elements are also important and have bearing on the Negotiated Dramaturgy. Where we also need to differ from the model is that the 'stage' of the negotiated dramaturgy is neither a self-sufficient nor a closed community. The relationship between world inside and the world outside is not circumstantial nor rather direct and inevitable. This emphasises the fact that Hauptfleisch's (1997) model is not a accurate representation of the Negotiated Dramaturgy. This model cannot represent the communication that takes place in the pre-production forums.

As with Pfister's (1988) models the Hauptfleisch (1978; 1997) models cannot depict the dynamics of the Negotiated Dramaturgy. While they may in part represent sections of the process they fall short in representing the process in its entirety. One can argue that the Negotiated Dramaturgy is a communicative process that uses theatre as well as other media to communicate. It is therefore inappropriate to use theatrical models to represent the entire process. A communication model needs to be developed to represent the process and components of the negotiated dramaturgy.

In most industrial theatre there is a developmental imperative present. As Mowlana (1987:5) observes,

> the meaning and philosophy of development is inherent in the value system of any community and nation in which a variety of economic, political, social and cultural activities are under examination.

The communication model representing the Negotiated Dramaturgy should draw attention to economic, political, social and cultural aspects that form part of the developmental imperatives. Similarly, it must reflect ongoing, dynamic two-way communication where all players are involved in the process on equal footing in terms of the creation, expression and interpretation of messages. If the participants are to communicate openly and freely, they need

to consider such elements as unique individuals with their own biographical circumstances and as members of diverse cultural and social groups. Therefore, the model should reflect the civilisation history, society and culture of both the communicators and recipients and place these within a broader, societal and international context. As the interpretation of the message is important in the Associated Dramaturgy, the model should clearly represent the medium as the central part of the process. The model should reflect that the message can be housed in a variety of media made up of different codes.

5. A Graphic Communication Model for Development

In the following paragraphs we discuss a communication model (Figure 6) that can be used in the development and analysis of negotiated dramaturgy within the context of industrial theatre.

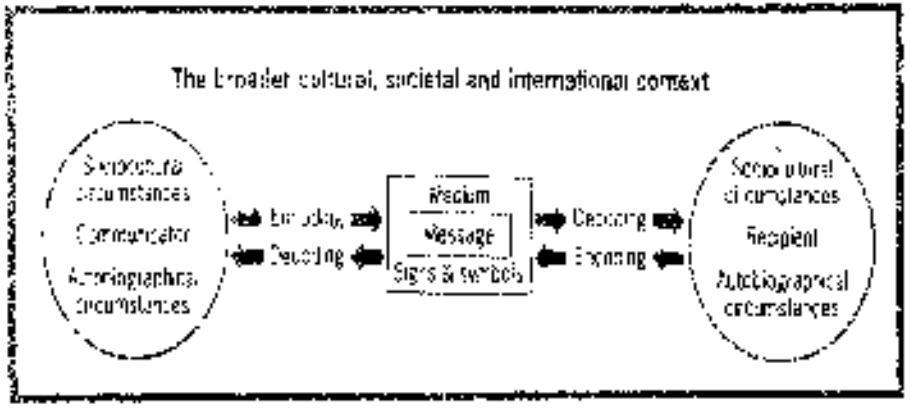


Figure 6: **The Graphic Communication Model for Development** (Mersham *et al.* 1995:55)

5.1 The Communicator

Typically, scholars refer to corporate communication practitioners as the source of the communication and the ones who are responsible for choosing the message (Heinrich 1986:6; Rensburg 1996:60; Verwey 1996:67). Mersham *et al.* (1995) argue, as many others have done in the past (e.g. van Schoor 1979; 1986) that the communicator 'exchanges roles' transactionally with the recipient on an ongoing basis

Mersham *et al.* (1995:54) further argue that in corporate communication, the professional corporate communicator needs to consciously and explicitly become a recipient or a 'listening post' for messages originating from stakeholders such as employees. In this particular context, the meaning content and flow of the communication emanates from the stakeholders

5.2 The Recipient

The recipients are actively involved in the process of creating and sharing of a message. In this model, the context of the communicator and the recipient is identical. This article has already defined the communicator as everyone involved in the process. *Inter alia*, the recipient is also everyone involved in the process. The context of the recipient is then identical to that of the communicator.

5.3 The Message and the Medium

The medium is a message receptacle or that which provides the platform for the signs, symbols and codes of the message to be conveyed (Mersham *et al.* 1995:55). Each medium has its own set of encoding possibilities and structures. In this regard the model is influenced by McLuhan's (1964) observation that the 'medium is the message'. McLuhan (1964) argues that the message content and presentation is shaped and influenced by the medium through which it is expressed.

Mersham *et al.* (1995) argue that the communicator must use appropriate media, drawing attention to the oral nature of the majority of communication messages within an organisation. They also indicate that the inappropriateness of certain other media such as print and video in the development of corporate cultures in the South African corporate context, where the broader workforce is involved. They argue for skills in encoding messages in the mother tongue of the communication partners as well as skills in the technology and techniques of the medium in question. Similarly, the model implies that it must not be assumed that all partners share the requisite skills in encoding and decoding the 'mediated' message where more sophisticated technologies (e.g. corporate videos, CD Roms) and codings (e.g. technical jargon) are employed (Mersham *et al.* 1995:55).

This approach reminds us that signs and symbols are devoid of meaning in themselves. They only 'receive' meaning only in the sense that the

source a sign from a specific meaning and if the recipient attaches a specific meaning to them. The sourcing of a sign depends not only on personal interpretation but also on collective agreement that may change across culture, space and time (Mersham *et al.* 1995:55-56).

The recipient's active participation in the communication process should be encouraged and recognised. The interpretation process may be active enough to transform the message into the recipients 'own' message when it is re-expressed. Often too much emphasis is often placed on the so-called 'effect' of the message upon the recipient. Mersham *et al.* (1995) therefore argue that it is important for the communicator to take active steps in encouraging the recipient to manifest his or her interpretation through a medium with which all the participants are comfortable. This makes it possible for the understanding or meaning attached to the original meaning to be evaluated.

5.4 The Autobiographical and Sociocultural Circumstances

In the model, the communicator and recipient are encircled by their sociocultural and autobiographical circumstances. Recognition of the sociocultural circumstances of persons involved in the communication process are essential in a country such as South Africa which is characterised by its sociocultural diversity. Language, culture, race, living conditions, social status and identification with specific communities and ethnicities are obvious factors that impinge upon communication in South Africa.

At the same time, individual circumstances must be recognised. Regardless of commonalties that link people in social structure, no two lives are the same in terms of individual experience (Mersham *et al.* 1995:57).

The sociocultural and autobiographical circumstances influence the perception and credibility of all the communicators and their messages in the negotiated dramaturgy. The knowledge of these circumstances establishes a definite context for the communication process. Mersham *et al.* (1995) stress the importance of the way in which the communicator manifests his or her ideas. In order to transfer thoughts, information, feelings and attitudes, the communicator needs to manifest these in a form that is accessible to all involved in the process. The model also suggests consciously monitoring the interpretations that are attached to messages by communication partners preventing them from remaining inner thoughts that the communicator assumes have been shared.

5.5 The Broader Cultural, Societal and International Context

This model also points to the importance of considering the broader societal circumstances in which the communication interaction takes place. This is illustrated graphically by the spheres surrounding both the communicator and recipient and the box around the triptych of the communicator, message and the recipient to represent the broad cultural, societal and international context. Mersham and Skinner believe that the societal (national) context is an important aspect of communication. For example, the trends towards affirmative action, transformation, restructuring and privatisation have all impacted on corporate communication. (Mersham & Skinner 2001:91-120, 145-170).

At the international level, globalism and global competitiveness, new digital communication media technologies, the concept of the African Renaissance, and South Africa's leadership role in many initiatives of the developing countries to play a greater role in international affairs, have also impacted on corporate communication and cultures (Mersham 2000; 2001; Meyer 2000.)

The return of South Africa to the community of democratic nations in the post apartheid era has exposed its peoples to many more stimuli and factors that effect the way in which individual South Africans express and interpret messages. The model draws attention to the transactional, cross-cultural nature of communication, and also the need to take into account the broader national and international contexts that impact on the South African workplace.

5.6 Critique of the Model

The model contextualises the basic elements of communication: the communication triptych within the broader cultural, societal and international context. It is important for organisations to re-align their positions in the height of the new context in which they find themselves. Organisations are constantly attempting to improve their productivity so that they may be world players. Many Industrial Theatre campaigns are intended for this purpose. For example, the researchers conducted a campaign at Richards Bay Coal Terminal (one of the largest in the world) where the stakeholders were made aware of their role in the international business community.

The Negotiated Dramaturgy encourages the stakeholders to become part of the process at every stage. They are responsible for establishing the

objectives and proposing ways in which these objectives can be realised. The stakeholders are also responsible for creating characters and a storyline. In the dramaturgy they are encouraged to respond to the scenarios presented to them. In this direct communication with the practitioners they suggest possible solutions to the problems dramatised. They also take part in 'feedback' conducted at the end of the campaign. In order for this to take place effectively, they must be empowered as communicators sharing ideas, suggestions and comments with the other participants. The model represents this, assigning communicator and the recipient equal status within the process. It also shows that those roles are easily reversed showing the communicator and the recipient both encode and decode the signs and symbols that construct the message.

By acknowledging the individual's sociocultural and autobiographical circumstances the participants are able to overcome potential barriers to communication. Experience shows that such communication (considering the sociocultural and autobiographical circumstances of both the communicator and recipient) increases as the process proceeds. It becomes one of the 'rules' for practitioners in the facilitation process.

The original Merskam *et al.* model (1995) may be criticised because fails to represent 'noise' or factors that function as physical barriers to effective communication. The dramaturgy may be hampered by the nature of the physical environment in which it is performed. This can include bad acoustics, sightlines, seating, lighting and physical characteristics of the venue. For example, in one instance, Baker (2001) conducted a campaign where the dramaturgy was performed in a tent next to the din of a generator. In another example, the dramaturgy was presented outside, under trees, next to a noisy factory. The same researcher has also had experience of working during the pre-production forum in a venue situated in or next to a noisy canteen. Although the practitioners try to eliminate this type of 'noise' they are often assured by the organisation that certain venues are ideal, only to find out later that they are not.

6. Conclusion

Despite the above omission, the model is useful in a heuristic way to every stage of the Negotiated Dramaturgy, challenging and interrogating many of the assumptions made in textbook approaches to internal organisational communication which suppose the communication process as unproblematic

and straightforward. It draws attention to the culturing value of the ideal typical form of the communication model and challenges the assumptions of using 'standard' wave-stepping techniques employing print media, videos, mission statements in languages and cultural codes that may not be mutually comprehended by participants.

Department of Communication Science
University of Zululand

Industrial Theatre Consultant
United Kingdom

References

Actors Mean Business 2000. *Industrial Theatre* <u>www.actorsmeanbusiness.com</u>. 23 March.

Baker, G & M van der Merwe 1999 Community Empowerment - Creating the Tools from within: Creative Conflict Management in KwaDlangezwa, KwaZulu/Natal. Paper presented at the South African Political Studies Association's Biennial Congress on 'Political Studies at the Turn of the Millennium', 29 June - 2 July.

Baker, GR 2001. *Negotiated Dramaturgy: Industrial Theatre as Communication in the Organisation.* Unpublished Doctoral thesis submitted to the Faculty of Arts in fulfilment of the requirements for the degree of Doctor of Philosophy (D.Phil.) in the Department of Communication Science at the University of Zululand.

Bell, S & B Geary 2000. 'Revving the Engine of Industrial Theatre'. *The Independent on Saturday* 31 July: 10.

Boal, A 1979. *Theatre of the Oppressed.* London: Pluto.

Boal, A 1992. *Games for Actors and Non-actors.* London: Routledge.

Boal, A 1995. *The Rainbow of Desire.* London: Routledge.

Boeren, A 1992. Getting Involved: Communication for Participatory Development. In Boeren, A & K Epskamp (eds): The Empowerment of Culture: Development Communication and Popular Media. *CESO Paperback* 17:47-59.

Braun, E 1982. *The Director and the Stage: From Naturalism to Grotowski.* London: Methuen.

Brecht, B 1964. *Brecht on Theatre: The Development of an Aesthetic.* Willett, J (ed and trans). London: Methuen.

Brook, P 1972. *The Empty Space*. Harmondsworth: Penguin.
Bogatyrev, P 1971. Les signes du theatre. *Poetique* 8:517-530.
Counsell, C 1996. *Signs of Performance: An introduction to Twentieth-century Theatre*. London: Routledge.
De Marinis, M 1993. *The Semiotics of Performance*. O'Healey, Áine (trans). Bloomington: Indiana University Press.
De Saussure, F 1969. *Cours de linguistique générale*. Paris: Payot
Duran, P 1994. Towards a Better Understanding of Code-switching and Interlanguage in Bilingualism: Implication for Bilingual Instruction. *Journal of Educational Issues of Language for Minority Students* 14.Winter:69-88.
Eco, U 1977. The Semiotics of Performance. *The Drama Review* 21,1.
Eco, U 1978. Pour une reformulation signe iconique. *Communications* 29. Paris: Seuil.
Elam, K 1980. *The Semiotics of Theatre and Drama*. London: Methuen.
Esslin, M 1987. *The Field of Drama*. Methuen. London.
Fourie, C 1981. *Aspects of the Theory of Communication*. Pretoria: Academica.
Grotowski, J 1968. *Towards a Poor Theatre*. London: Methuen.
Hauptfleisch, T 1978. *The Play as Communication: A Study of the Language of Drama*. Pretoria: University of South Africa. Unpublished D.Litt. et Phil. Thesis.
Hauptfleisch, T 1997. *Theatre and Society in South Africa. Reflections in a Fractured Mirror*. Pretoria: van Schaik.
Helbo, A, Johansen, J Dines, P Pavis & A Uhlenfeld 1987. *Approaching Theatre*. Bloomington and Indianapolis: Indiana University Press
Hempfer, K 1976. *Gattungstheorie*. Munich: Spanen.
Honzl, J (ed). Dynamics in the Sign of the Theatre. *Modern Drama* 2,1:118-127.
Huebsch, JC 1986. *Communication 2000. 2nd edition*. Durban: Butterworths.
Johnston, C 1998. *House of Games: Making Theatre from Everyday Life*. London: Nick Hern Books.
Jones, GR 1995. *Organizational Theory: Texts and Cases*. Massachusetts: Addisen-Wesley.
Kelman, J 1996. *The Leader-Trainer-Facilitator in Small Group Communication: Theory and Practice*. Cathcart, Samovar & Henman (eds). Madison: Brown and Benchmark.
Kennedy, AK 1983. *Dramatic Dialogue: The Duologue of Personal Encounter*. London: Cambridge University Press.

Kennedy, J 1987. *The Theatre of Grotowski*. London: Methuen.
Learning Theatre Organisation (The) 2000. *Learning Theatre Organisation: Introduction and Overview*. http://www.learningtheatre.com/overview.htm. 23 March
McLuhan, M 1964. *Understanding Media: The Extensions of Man*. UK: Routledge.
McQuail, D 1987. *Mass Communication Theory: An Introduction*. London: Sage.
Mersham, GM 2000. Australia and South Africa: Neighbours or Strangers in the South? – Economic, Political and Cultural Comparisons. *Mass Media: International Journal of Plural Societies* 2:—April
Mersham, GM 2001. *The Impacts of Digital Technologies on Corporate Communication in Southern Africa*. Workshop presented to TSAMI (Tanzanian Southern African Management Institute). Durban. 31 October
Mersham, GM & JC Skinner 1999. *New Insights into Communication and Public Relations*. Johannesburg: Heinemann
Mersham, GM & JC Skinner 2001. *Organisational Communication*. Johannesburg: Pentagram.
Mersham, GM, RS Rensburg & JC Skinner 1995. *Public Relations, Development and Social Investment: A South African Perspective*. Pretoria: J.L. Van Schaik.
Meyer, MJ 2002. Globalisation: An Issue of Contestation and Struggle in South Africa. *Mass Media: International Journal of Plural Societies* 12 - April.
Miller, S 1993. *Systems of Relevance: Stanislavsky, Brecht, Grotowski and Brook*. London: Routledge.
Moore, CW 1986. *The Mediation Process: Practical Strategies for Resolving Conflict*. San Francisco: Jossey Bass.
Morgan, N & J Saxton 1987. *Teaching Drama*. Portsmouth: Heinemann
Mounin, G 1969. La communication théâtrale. *Introduction à la sémiologie*. Paris, Minuit.
Nowotna, H 1957. *Development: A Field in Search of Itself*. Budapest: International Association for Mass Communication Research
Mukařovský, J [1934] 1976. L'art comme fait sémiologique. Actes du huitième congrès de philosophie à Prague. In Matejka and Titunik.
O'Toole, J & B Haseman 1988. *Dramawise: An Introduction to GCSE Drama*. Oxford: Heinemann.

Plato 1955 *Republic*. Lindsay, AD (trans). London: JM Dent and Sons Limited.

Pfister, M 1988 *The Theory and Analysis of Drama*. Cambridge: CUP.

Prescott, P 1993. Creative Drama and Young Children: The Dramatic Learning Connection. *Arts Education Policy Review* July/August,94,6:24-29.

Rensburg, R 1994 Public Relations Communication Media. In Lubbe, BA & C Puth (eds): *Public Relations in South Africa*. Isando: Heinemann.

Ruffini, F 1974. Semantica del reale: riconoscimento degli stati. Bibliotheca teatrale, 9.

Ruffini, F 1979. Spettacolo comunicante/spettacolo significante: Una comunicazione spettacolare. Paper at conference on Text, Communication, and Performance, Cosenza, 14-16 September.

Samovar, LA, L Herman & S Ng 1990. Small Group Process. Cathcart, LA Samovar & L Henman (eds). *Small Group Communication: Theory and Practice*. Seventh edition Madison: Brown and Benchmark.

Samovar, LA & RE Porter 1995. *Communication between Cultures*. Belmont: Wadsworth.

Samovar, LA, RE Porter & JC Nern. 1981. *Understanding Intercultural Communication*. Belmont: Wadsworth.

Skinner, C, L van Essen & GM Mersham 2001. *The Handbook of Public Relations*. Johannesburg: Oxford University Press.

Steinberg, S 1996. Organisational Communication. In Rensberg, R (ed) *Introduction to Communication. Communication Planning and Management*. Cape Town: Juta.

Styan, JL 1981. *Modern Drama in Theory and Practice 3: Expressionism and Epic Theatre*. Cambridge: CUP.

Valtinoki, I 1975. *Poetics of Literature*. Assen: Royer de Ridder

Van Schoor, M 1979. *The Idea of Communication*. Bloemfontein: OU of Willies.

Van Schoor, M 1986. *What is Communication?* Pretoria: Van Schaik.

Verwey, S 1994. Communication Theory. In Lubbe, BA & C Puth (eds): *Public Relations in South Africa*. Isando: Heinemann.

Woodson, SE 1999. (Re)Conceiving "Creative Drama": An Exploration and Extension of American Metaphorical Paradigms. *Research in Drama Education* September,4,2:201-225.

Williams, R 1981. Argument, Text and Performance. In Hooley, M & S Wills (eds): *The Twentieth Century Performance Reader*. London: Routledge

Cognition, Persons, Identity

Simon Beck

SECTION 1: The Story so Far

During the last twenty or so years, our particular view as to what constitutes the identity of a person has come as close as anything is likely to come to being the received view amongst philosophers working on the topic. The trend has been to follow the lines of Derek Parfit's update of Locke's view that what makes someone the same person over time is a matter of facts concerning continuity. Overlapping, one is formed most importantly by apparent memories, continuing beliefs, desires, projects and so on (while not forgetting emotional attachments) are held to form your identity (Parfit 1984:205ff, 222). To be more accurate, this continuity is aimed to be what constitutes personal *survival*, while its holding in a 1:1 relationship would amount to personal *identity* (or *strict* survival).

While this view can be called the standard view now, it certainly was not always standard. Reaction to this cognitive view, or what I will call (following current usage) the 'psychological' view, led a number of philosophers in the 1960s to espouse versions of a *physical* criterion of identity, most notably Bernard Williams (1960; 1970), Sydney Shoemaker (1963) and Terence Penelhum (1970). They argued that at least some overtly physical or bodily continuity was necessary for personal identity. But these attacks no longer have any teeth. Shoemaker's case was based on a

[1] Parfit's (1984) view updates Locke's in both expanding the psychological links to include more than just memory and in substituting the less question-begging notion of *apparent* memory (which does not imply that you are the person who had the "remembered" experience) for the notion of *real* memory (which does). It also famously rejects the idea that identity is what matters

misguided semantics (Beck 1994) and he has since repented (Shoemaker 1970, 1999). And the attacks of Williams and Perelhuin affect only a cognitive account like that of Locke which reduces the identity of x and y to x's being able to (really) remember y's experiences—that is, an account much more naive than Parfit's.

But the nineties brought a new resurgence of non-cognitive views, based on new and intriguing arguments that are aimed at the sophisticated standard view. These views are united in classifying the interests of cognitive scientists as utterly distinct from the important field of metaphysics. In this paper I will look at three influential theories along these lines—those of Peter Unger, Peter van Inwagen and Eric Olson (although I will suggest that the latter two are at least the same). I hope to show that the orthodoxy survives all of these attacks. All of the attacks take some form of materialism for granted, and I am happy to accept that. One of the central contentions, as we first see, is that you cannot be a serious cognitive scientist and hold the standard view of personal identity. The standard psychological view, it is contended, is inconsistent with materialism and thus with serious cognitive science. I will argue that this is false.

The arguments I am concerned with present different challenges to the psychological view. The one is, however, in a sense, softer than the others, requiring less change of the standard psychological view. In the first part I treat this softer challenge and then proceed to engage and reject the harder ones.

SECTION 2: Unger's Soft Challenge

In his *Identity, Consciousness and Value*, Peter Unger sets out a criterion for personal identity that, he is at pains to point out, is at odds with the received view. While he accepts that cognitive factors are of some importance, Unger wishes to champion an incorruptible physical criterion. While I cannot hope to do anything like justice to the full complexity of Unger's case for this criterion here,[2] I nevertheless wish to question what on this alternative can hope to oust the orthodoxy.

[2] I present a detailed argument against many more considerations he raises in Beck (2001).

Unger's view is not a crude physical criterion counting just all there is to personal identity is the existence of the same human body. Nor does he even require the continuous existence of the brain. He accepts that to count as the same person you need to retain your 'core psychology'—which is made up (roughly) of the capacities for reasoning, self consciousness and for forming simple intentions—the capacities which all normal humans share (Unger 1990:68). He denies that 'distinctive psychology'—the continuity of beliefs, desires, projects, etc.—on which the received psychological criterion places all the emphasis, is crucial at all. For Unger it is the case of person x (at time t) and y (at some earlier time t') if and only if there is sufficiently continuous (and unique) realization of x's core psychology between the physical realizer of x's core psychology at t and the physical realizer of y's core psychology (Unger 1990:69).

Now, those who are persuaded that our survival is ultimately a matter of psychological connectedness are likely to complain that, even if we accept in this criterion to ensure survival. After all, a continuity of the capacities for reasoning, self consciousness and for forming simple intentions—as found in all normal human beings—does not sound much like *me*. And it hardly seems enough to ground the claims of responsibility, commitment, and so on, which make the personal identity debate so important. They may also complain that this is no more than an overly watered-down version of their own view of identity, or not appropriately described as a physical criterion at all. On the second count, however, Unger has not simply replaced continuity of distinctive psychology with continuity of core psychology—the physical aspect is indeed fundamental in his view. For, according to Unger, it is precisely the continuity of the physical realizer that makes those core mental capacities *your* core mental capacities, as opposed to *other*, exactly similar, capacities (Unger 1990: 113-117).

Even so, the continuity of these core capacities is necessary for survival on Unger's account, and we can still ask whether this is true. Surely who you think you are and whose experiences you seem to remember are more important to who you are than whose basic capacities you enjoy? Sydney Shoemaker (Shoemaker ?) has also asked just this, and has argued that 'individual' psychology is more crucial to our surviving than core psychology (Shoemaker 1992:14) 149). On the face of things this seems

plausible enough – and Shoemaker moves to support this intuition with his description of Brainland:

> Brainland is covered with brains, each linked to a life-support system and to the rest of the network of brains. All of the brains have the capacity for active thought but only some are active at any time. Should an active thinker decide (for example) to move in a specific direction, that causes that direction would then be activated while the original part becomes dormant. Its thoughts moving across to the activated brain. In this way, discussing psychology will move around Brainland, being supported by different brain psychologies (following Unger's account of Gert) as they go. When it comes to discussing the persons in Brainland, Shoemaker claims that we have a clear example of continuity of physical psychology coupled with psychological shifts as the brains in Brainland stay put, the persons move around' (Shoemaker 1992:161).

In the face of this challenge Unger retorts from 'Forget Brainland', he says,

> here's the painful case. A machine records the exact nature of, and the relative arrangement of, all your atoms. The information is then sent to a computer device that, from now matter, makes a qualitatively identical person. Now, instead of destroying anything much – the very process by which the information is recorded also *momentarily* does this: it affects your *brain* in such a way as to make it the brain of an amnesiac moron. As is very clear, you survive, just becoming such a moron; it's *someone else* who now has the (sort of) distinctive psychology that, before, only you had (Unger 1992:161-162).

Unger feels that this thought-experiment removes misleading descriptions, and thereby permits a better description of what goes on in Brainland: persons may just 'move' as they are, becoming amnesiac morons while the adjacent brain acquires a distinctive psychology exactly like theirs used to be (Unger 1992:162).

SECTION 3: Unger's Challenge Considered

At first glance, Unger does indeed take the intuitive high ground here: my

first response to his thought experiment is that I would be the moron. But this first response may be too superficial – it's not like my first response that the rabbit has actually disappeared from the magician's box – and it does merit some investigation.

If the thought-experiment under discussion only involved a machine that turns me (or to put things in neutral terms) what used to be me, into an amnesiac moron, then my response is clear: that I would be the moron. To confirm that this fits with my deep views on 'ordinary survival', we can apply Unger's crucial test of whether I would accept a full measure of pain now so as to avoid that moron's undergoing great pain in the future. The result is that I respond unequivocally that I would take the pain now. My response is also clear, although more complicated, when the great future pain test is applied to Parfit's 'Branch-line' thought experiment. In that case (Parfit 1984:267), a machine like Unger's one creates a copy of the person from new matter. But instead of turning the original person into a moron it leaves him as he was psychologically – only damaged in such a way that he will die in a few days. In this case, I would take a fair amount of pain beforehand to reprieve the fatally damaged person from torture. I would also take pain for the reprieve of the surviving copy – but not as much pain. To borrow some terms from an earlier work of Unger's (Unger 1982:1) my *dominant* response to the branch-line case is that I stay in my old body, but I nevertheless also feel a *diminished* response that I am the surviving recipient.

This brings us back to Unger's case which combines the amnesiac moron scenario with Parfit's branch-line. Having thought about the branch-line case, I am much less eager in my response to Unger's story. The response which he suggests as evinced by his story is really only forthcoming from the simpler version in which the only candidate for identity is the amnesiac moron. When his story is considered as a development of the branch-line case it is no longer at all clear that I should accept much pain to reprieve the moron, or that I should take less pain for me than I should for the sake of the psychologically similar replica. Changing the branch-line case in this way has an important effect on my intuitive response in that thought experiment, namely of reversing the order of dominance. At the very least it renders my response that I survive as the moron much less dominant, to such an extent that it cannot be said to provide anything like independent confirmation for Unger's criterion. And significantly, the factor

that has been responsible for the change in response is the removal of *continuing distinctive psychology* – the very factor Unger wishes to sideline through his thought-experiment.

In the light of these considerations, Unger's alternative to Shoemaker's thought-experiment cannot be seen as the helpful case that it was advertised to be. His claim that it is 'an example ... which does less than Shoemaker's to generate contexts where misleading tendencies have much force' (Unger 1995:167) is just not true.

Rather than Unger's version, I contend that *this* is the helpful case as far as the contest between core and distinctive psychology is concerned: Brown and Robinson (yet one more time) have their brains removed from their bodies, using a brain state transfer device along the lines envisaged by Bernard Williams (Williams 1970) – which can extract information stored in one brain, clear it of information from another and replace it with the scanned information – distinctive psychologies are swapped between the removed brains. Now Robinson's brain is placed into Brown's body and vice-versa. In Robinson's body is someone who thinks they are Robinson in Brown's body is someone who thinks they are Brown. The question as to who they really are seems 'fair to receive a straightforward answer. The crucial point to notice, though, is that Brown-body has *Robinson's core psychology*, given the continuity of its realizer – Robinson's brain. The same goes, *mutatis mutandis*, for Robinson-body. Is there any temptation to say that a body swap has occurred in fact, with what Unger's criterion decrees? The great future men test certainly provides no support for such a description. The only diagnosis for someone who insists that Brown is now in Robinson's body seems to be that they have 'fallen into the trap that Unger (1990:13) warns us against and allowed too great an "attachment to a certain approach" to warp their intuitions.

SECTION 4: The Hard Challenge I—Van Inwagen

Unger's case, even if it fails, only asked that we bring physical factors into the picture as a way of marking the distinctness of a particular core psychology. The challenges to the psychological view now to be considered ask much more than this. They demand that we ignore psychological factors altogether in characterizing what constitutes a person's identity over time.

Both arguments that I will examine have this consequence, but I will examine that presented by van Inwagen first, since it is the more direct of the two. It also forms an interesting contrast with the case presented by Unger insofar as views about method are concerned.

Van Inwagen professes to be deeply opposed to thought experiments of the sort used by Unger and Shoemaker (and in my response to Unger). He points out that the most influential arguments offered for a psychological view of identity depend on the intuition that people can 'swap bodies' which is drawn out by thought experiments involving brain state transfer devices and the like. He is unimpressed by this method and the bet it's it evokes, no matter how widely they are shared. He is, he says, 'an avowed enemy of the philosophical employment of fantastic thought experiments' an employment that is, 'I believe, the result of the widespread adherence of philosophers to the nonsensical idea of "logical possibility"' (van Inwagen 1997:308). Despite this aversion, the argument he presents here does not turn on a rejection of the method of thought experiment. Rather, he allows brain state transfers for the sake of argument, but contends that even if your psychology were fed in some way into another body, that would not mean your identity would go with it. As he puts it, he sets out to show that informational transfer cannot turn one material thing into another.

Here is the core of van Inwagen's case:

> You believe that human persons really exist. And you are a materialist. So what do you think these human persons that you think really exist are? Well, material things, obviously. Perhaps whole human organisms, perhaps parts of human organisms ... but some sort of material thing. For the sake of having a concrete example, I'll suppose that you think, as I do, that human persons are whole human organisms .... So you are a certain living organism, *x*. And if you hold views like Shoemaker's on the possibility of bodily transfer, you believe that there could be another, numerically distinct living organism *y* such that, if the right sort of information flowed from *x* to *y*, you would become *y* .... But when the matter is put this way, it is evident that your belief is simply impossible, a violation of the very well established modal principle that a thing and another thing cannot become a thing and itself (Van Inwagen 1997:310).

The general premise in this argument – the presentation of materialism as the view that all real things are material things, and that this 'are' is to be understood as representing strict identity. Thus if you are *x* in this sense and you become *y* (after your psychology is transferred to *y*) while *x* still exists, then you are both *x* and *y*. And since, by assumption, *x* and *y* are distinct individuals, this is in conflict with the principle of the necessity of identity. In the light of this *reductio*, van Inwagen contends, we cannot hold both materialism and a psychological view of personal identity. He is clear as to the way forward: we must accord psychological factors any of the weight they are traditionally allowed, and must realize that we are organisms (or some such material thing) and no more than that – our identity is a matter of organism-identity.

The *reductio* premise outlined above is a contentious one, however. I do not mean to drag the issue of relative identity back into the debate here. Van Inwagen sees that doing so would be one possible response, and I am happy to accept his contention that this would not lead to an acceptable solution[3]. The view that the relation between persons and their bodies, organisms, or whatever, is strict identity is contentious because it has been widely denied, even by those of a strong materialist bent. Van Inwagen (1997:311) cites Shoemaker as expressing just this denial:

> Shoemaker recognizes that, although his brain-donor and brain-recipient are (on his account) the same *person*, they are not the same animal or the same human being. His solution to this difficulty is to say that ... persons are not *strictly* animals or human beings .... Rather, persons ('human' persons, anyway) 'share their matter with' and 'occupy the same space as' and have 'the same non-historical properties' as human beings. And, of course, he holds that it is in principle possible for one and the same person to bear these relations to different human beings at different times.

Van Inwagen (1997:312) is not about to accept this response. It just misses the point as far as he is concerned.

[3] I am also happy to accept his contention that a four-dimensionalist approach won't offer an acceptable solution.

> Let us grant, for the sake of argument, that ... persons ... are not strictly identical with any human being or any animal. Nevertheless, if one is a materialist and if one believes that persons really exist, then one must concede that every person is strictly identical with *some* material thing. Someone who holds views like Shoemaker's is therefore committed to the proposition that there could be two simultaneously existing material things such that one of them could become strictly identical with the other simply in virtue of a flow of information between them. ... The argument does not essentially depend on the assumption ... that the two objects in the case considered are 'human organisms'. Substitute in the argument any material object category you like for 'human organisms' and the point of the argument is unaffected.

SECTION 5: Van Inwagen's Challenge Considered

I want to argue that van Inwagen is wrong to insist that a materialist cannot follow something like the route that he accuses Shoemaker for attempting – trying to defend a psychological view. Consider familiar materialist claims about the relation between psychological states like beliefs and material things. Most materialists nowadays accept some sort of functionalist account of the psychological: psychological states are defined by the functional roles they play (how they are typically caused by external stimuli, how they interact with other internal states and what behaviour they typically cause). These defining roles, it turns out, are played by material states[4]. In the light of this general picture, the claim is often made that beliefs are material states. At first glance, this analogy seems to be going van Inwagen's way. For if my belief that p is a brain state, then we have the same sort of relation that he is holding us to between persons and material things, the one which led us into the *reductio*. But to stop there is to ignore the crucial place of *realization* in the story.

The functionalist account is that brain states *realize* the roles that define psychological states. Psychological states like beliefs are higher-order

[4] I will stick with van Inwagen's term 'material' rather than the usual term 'physical', simply for stylistic reasons.

states defined (at least) in terms of causal relations) over a domain of first-order states, and these first-order states turn out to be brain states. In this way there is no threat to the principles of materialism, yet no requirement that beliefs be strictly identical with brain states—it is on precisely that point that functionalists differ from sixties-style mind-brain identity theorists. Van Inwagen's case against the compatibility of the psychological view of personal identity with materialism rests on the demand that something has no place in the materialist ontology unless it is identical with something material, yet this is patently not the case with the items closest to the materialist's heart. Should materialists—most of whom are functionalists—indeed be convinced by van Inwagen's principle they would have to be eliminativists, since their denial of the identity of beliefs with material states would mean that there are no beliefs. In other words, the materialism that van Inwagen sees as incompatible with the psychological view is not materialism as it is usually understood at all.

This reply has said nothing about what the relation is between persons and human beings nor even argued that the relation between beliefs and brain states is suitably analogous to this relation. I think showing the central premise in this argument to be faulty is enough to block the argument, but I concede that it may nevertheless not be a fully satisfying response. I will return in Section 8 to a more detailed account of the relation between persons and humans, but in the meantime something can be said to strengthen my case.

Although a person is a different sort of thing from a belief (persons *have* beliefs, but not vice-versa) there are some important analogies between the concepts. The sort of materialist envisaged above sees beliefs as states that are defined by their functions—by what they *do* rather than what they *are* (in the sense of what they are made of or how they are internally structured, say). Something very similar seems to be true of persons. Persons are defined by the distinctive role of moral agent and moral object. Their standard description is given in functional terms. According to Locke's (1694: 59) classic account, a person is something that can think and 'can consider itself as itself, the same thinking thing, in different times and places'. Something is a person as long as it can do the things a person does—in this very clear sense one of the features of the concept *person* is that it is a functional concept. What it is that performs these functions is not specified.

As a result, just as in the case of beliefs and brain states, there is nothing in the concept itself that demands identity between a person and what it is that functions in this way.

Now we can return to the issue of Head. *Person* as a functional concept is a higher-order concept, and we can see persons as being related to (material) things of an order below, even if that relation is not precisely the relation of realization. Persons are (perhaps *constituted by* human organisms[7] or whatever turns out to play the relevant role) in a way similar to which beliefs are realized by brain states. This would mean that persons are not strictly identical with material things, but that does not make them *non-material* things. Organism *y* may thus cease to exist when you when information is transferred from organism *x* to *y*, yet there is no conflict with the necessity of identity since this is the reason (as there was in van Inwagen's version of things) to say that you are still alive as well as being *y*. That means there is a way of reconciling material constancy with the psychological view of identity, despite van Inwagen's insistence to the contrary. And importantly, the analogy that serves to show the way comes from precisely the source that van Inwagen cites as his main premise—namely, materialism. Have we given up the other central premise in van Inwagen's argument—taking this route—that is, are we denying the reality of persons? Persons may not feature in the list of basic things in our ontology, but that makes them no less real than any psychological thing.

SECTION 6: The Hard Challenge 2—Olson

Olson's challenge to the psychological view turns on issues similar to those raised by van Inwagen, although it is a more sophisticated case and places its emphases differently. Like van Inwagen, or Head, Olson requires us to ignore cognitive factors where our identity is concerned, and to see it rather as a matter of the identity of organisms. The case he presents against the psychological view centres on two consequences of such a view which he claims are strikingly counter-intuitive.

[7] Note that I am not relying on Lynn Rudder Baker's definition of constitution here. As will emerge in Section 8, I do not think that persons are ontologically on a par with humans as Rudder Baker does.

Olson argues that if the psychological view is taken seriously, then it follows that you and I were never foetuses (Olson 1997a), and that you and I are not (even contingently) animals (Olson 1995). He contends that since it is obvious both that we are animals and that we were foetuses, the psychological account of identity is to be rejected out of hand and replaced by an account which does not have these absurd consequences—such as his 'biological view' that the conditions for personal identity are precisely those for the identity of a human animal.

Olson explains why we cannot have been foetuses following the psychological view as follows (Olson 1997a: 95-97). What makes me the same person I was, according to the standard psychological view is the existence of psychological connections between that person then and myself now or, if no such direct connections exist any longer, of a chain made up of overlapping links formed by such connections. These connections will usually be those of memory, continuing beliefs, and so on. According to best evidence available, a five month old foetus has no psychological experiences which could be remembered, and no mental contents which could count as a belief. Its cerebrum is not in a state to support anything which could even be called a capacity for conscious awareness. There is thus no possibility of any psychological relation, direct or indirect, between myself now and a five month old foetus. This would mean that I was not that foetus.

> So if the Standard View is right, nothing could be a fetus—or at any rate a fetus that cannot yet think—at one time and a person later on. No person ever was a fetus, and no fetus ever becomes a person (Van Inwagen 1997a: 96)

The matter does not just rest there either, Olson insists. He wishes to show that the 'foetus problem' is a major embarrassment to the psychological view. To bring this about, he points to what this view can say about what becomes of a foetus. There are two options, he says, since the foetus does not become a person: either the foetus dies when the person comes into existence, or it survives and comes to share its matter with a person numerically different from it (Van Inwagen 1997a: 101). But neither of these options is acceptable. The foetuses of all other animals survive and become adults. It seems absurd to say that humans are different here. So that leaves

us with the option that I (this person) share my matter with another thing—a human animal—which is a distinct being from me altogether since it was the foetus I was not. This has the very strange consequence that although I am made of living organic tissue[3], just like other animals—in fact, of exactly the same tissue as one animal—I am not an animal at all. This is what Olson calls the 'animal problem'.

The animal problem does not end there. This is because the animal whose matter you share and from which you are indistinguishable, though distinct, is also indistinguishable from *the person* you are. I see you, it thinks it is a person, but it is not. And how then do we know that we are not making the same mistake as it does? You are, after all, no more certain that you are a person than it is. The problem is also more general than I have made out so far. It is not just a problem concerning one option of how to explain the fate of foetuses: it is a problem any psychological account must face anyway. Because if the identity conditions of a person differ from the identity conditions of a living human, then there are two distinct beings. And given a psychological account of personal identity, this difference is inevitable—the human vegetable left by removal of your cerebrum is still a living human, but it would be something numerically distinct from you. And that seems to mean that the psychological view cannot avoid the embarrassing consequences just outlined.

Olson's response to these problems is to adopt the radically different theory mentioned above—the 'biological view' that I am essentially a living organism. It is this essence that sets my identity conditions, rather than being a person. That means that I remain the same as long as this biological life continues. I was this a foetus and I am an animal: the problems facing the psychological view just do not arise here. Psychological matters are, on this view, simply irrelevant to personal identity.

SECTION 7: Olson's Challenge Considered

Two things should be noted about Olson's case and the view he takes it to support. One is its reliance on the principles of David Wiggins's account of

[3] Olson assumes agreement that we are material objects—pointing out that almost all who accept the psychological view accept this as well.

substances, and the other is the way it flies in the face of the thought-experiments that have generally been taken to offer support to a psychological account of identity. It is these two points that will shape my response.

According to Wiggins, the identity (or persistence) conditions of a thing are governed by the type of thing it is essentially – as he puts it, by its 'substance concept'. Any thing may fall under a number of concepts, but most of those will only apply at certain stages of its career – as the concept *tadpole* applies during only a part of a frog's career. They are mere 'phase sortals', but what makes something one and the same thing throughout its career will be determined by its substance concept (Wiggins 1980:24). Olson's diagnosis as to why so many are misled into accepting the psychological view is that they take *person* to be our substance concept. Given that substance concept, my identity conditions will centre around psychological matters, with the highlighted consequences that I am not an animal and was not a foetus.

On the other hand, the biological view takes our substance concept to be *unified organism* or *human organism*. The ensuing identity conditions allow that I was a foetus and am an animal. The view asserts that *person* is just a phase sortal, that applies only to part of my career, the part when psychological relations are to be observed. After all, a human with its cerebrum removed can still be a living organism – and thus *me* – even though it has no psychology.

This brings me to the second point mentioned above. Olson's biological view is by no means as free of counterintuitive consequences as this passage might suggest, and it is time to highlight some of the costs of his view. His case turns on the costs the psychological view must pay for its truth, but his own view must itself pay the costs incurred in rejecting the case for the psychological one. As Olson acknowledges, the most powerful considerations militating acceptance of a psychological view has been the possibility of a body-swap. Ever since Locke started the debate, thought experiments outlining what seemed best described as a person swapping from one body into another have been taken as compelling evidence that our concept of ourselves allows that we can do this. The crucial test would be whether I survive the brain-state transfer device in another body or even (slightly less fantastically) the transplant of my cerebrum – that is, the part of

my brain supporting my individual consciousness – into another body. As we have seen with van Inwagen, not everyone will tolerate consequences in these cases, but once considered they have drawn the strong response from all (or almost all – barring Unger as discussed), 'Yes, I do'. Even Olson acknowledges that he responds intuitively in this way (Olson 1997b: 161). Nevertheless, given the biological view he proposes, he must say, 'No, I do not survive – I stay with the living but mindless body'. And I stay, not because this body has my capacity for consciousness, but merely because it is still the same living body. The other body will think it is me, love my family, have all my apparent memories, beliefs and desires, and all of these will be causally grounded in my own experiences – this all may seem crucial, but none of it is acknowledged as relevant to who I am. And consider what reaction the biological view would require to a case of looking in the mirror and seeing what appears to be someone else's face looking out. You must say, 'Oh look, I'm not who I thought I was!'

I do not wish to push this particular point too hard, since Olson (1997b:Chapter 7) does offer an argument to ameliorate his ignoring of the widely-held intuition that we go with our psychology, one which I do not have space to deal here[7]. There are nevertheless further counterintuitive consequences beside the ones just outlined. One is a direct analogy to his own argument against the psychological view. He claims that that view implies that I am not an animal: his view implies (although he never makes this clear) that I am not a person. On the biological view I am (strictly identical to) a human organism, and personhood is a phase I go through. But that is just it; I am a substance, and a person is not. There is not really any such thing as a person; like *adult*, *person* has no identity conditions (Olson 1997b:27), such things do not attach to phase-sortals. To say 'I am a person' must then be to speak metaphorically, since it cannot really be an identity claim in the way that 'I am an animal' is. And yet if we are faced with the choice of denying either 'I am an animal' (along with what Olson demands of the psychological view) or 'I am a person' (along with the biological view), it seems much easier to deny the latter. Even more importantly, this role the biological view casts for *person* means that the psychological view does not just provide the wrong identity conditions for persons, but that it is

[7] Although I am quite prepared to argue that his attempt fails.

needlessly to avoid even the question posed by body-swap thought-experiments. 'Is this (different human) the same person?' If the biological view is right, that would be like asking, 'Is this (different human) the same adult?' To say the very least, that is deeply counterintuitive.

As far as counterintuitive consequences go then, it is not at all clear that the biological view is any worse off than the psychological view. If anything, the psychological view appears to fit better with our conceptual scheme. But perhaps this is because we are not taking our metaphysics seriously enough. That brings us back to the first point I raised in this section, the grounding of the biological view in Wiggins's metaphysics. Even if the psychological view wins on the intuitive level, the battle at the deeper metaphysical level can still be decisive, and decisive the other way.

On this deeper level things do look bad for the psychological view. If we cling to the view that I am essentially a person, and that *person* is the concept which determines my identity conditions, then we are taking *person* to be a substance-concept. That may sound fine, and most proponents of a psychological view may be happy to accept this. But Olson contends that *person* is simply not fit for it is a substance concept – it is the wrong sort of concept to play that crucial role. A substance-concept in the Wiggins system (and this point is rooted in the history of discussion of substances going back to Aristotle) cannot be dependent on some more fundamental concept. *Person* seems to break this rule – for all the features which are definitive of persons (the having of intentional states of sophisticated kinds, and so on) are features we have in virtue of being humans. Wiggens outlines the task of a substance concept as being, not to answer questions like 'what does this thing do?' or 'how much does it weigh?', but rather ones like, *'what is it* that is doing such-and-such?' (Wiggins 1980:151). *Person* does not answer this sort of question, as Olson points out:

> To say that something is a person is to tell us what it can do, but not to say what it is. To say that something is a person is to say that it can think in a certain way ... but it doesn't tell us *what* it is that can think in that way (Olson 1997b:32)

With this case behind him, Olson is in a position to stand up to the problem posed by our intuitions about ourselves. The metaphysical case

wins, he argues, and we must accept that my concern for the individual with my distinctive psychology is a mere practical concern. It is only in this practical sense that I am the same person as the human who inherits my psychology in a cerebrum transplant: that individual is not numerically identical with me – he is not *really* me.

Before we accept defeat for the psychological view, however, we must be prepared to concede all that has been claimed about the deeper metaphysics here. Some of it I have already conceded. In Section 4 I responded to van Inwagen's challenge by claiming that *person* could be functionally defined, and that is precisely what Olson has contended is the case. All the same, I wish to argue that we can tell a story other than Olson's which fits better with the case made by our intuitions, and goes some way towards lessening the costs which Olson has claimed adherence to the psychological view incurs.

Section 8: Persons Located

Here is that story. I am essentially a person. *Person* is not a substance concept, though[8]. While the concept *person* is what fixes my identity conditions, a person is not a substance in Wiggins's terms. Olson is correct that *person* does not answer the questions that substance concepts must, and the distinctive properties of persons are properties they have in virtue of being something else, in a way that substances do not. But while *person* is not a substance concept in this sense it is also not a phase-sortal, the only other option that Wiggins offers us and which Olson takes it to be. The person in the body that inherits my cerebrum and psychology is the same person that was in my old body. Those two humans are distinct substances, and phase-sortals only apply in the career of one substance. We need to add a third sort of concept to the system – and *person* is one of those. Olson sees only substances as being genuine sortal concepts in the sense of having identity conditions attached to them, but simply because *person* is a functional concept (unlike *human* or *frog*) that does not mean it is not a sortal concept as well. What the psychological view offers us precisely an

[8] Nor does it fit all the requirements for being a 'primary kind' concept – Lynn Rudder Baker's (2000) substitute for substance concepts.

account of what it is that we and only persons have in common, and of what it means to be a person over time". *Person* is not a concept unique in being functional and sortal – numbers are also functional and sortal – although it is unusual in being a 'continuant' concept as well, when number concepts are not. But there is no reason, at least we certainly have not met one in this discussion, why it cannot be such a concept. persons are massively interesting things, and that is the role our conceptual scheme grants them. The 'I' that Olson and van Inwagen attempt to locate in the theory of human organisms (to use Frank Jackson's picture of metaphysics) is not the 'I' which attaches to our folk concept of a person or to that of ourselves.

What does all this mean? It means presumably that *person* is not a fundamental metaphysical concept (at least on the Wiggins system). It is a concept that applies only to certain human beings in virtue of the features they have – the psychological features referred to above. The problems raised by Olson can now be dealt with. 'I' is an honorary role only appropriately used by persons and its proper use follows the identity conditions that the concept *person* dictates (viz. psychological continuity). So while I *am* an animal, this is not in any essential sense but rather in the sense that I am constituted by an animal as discussed in Section 5 (and in the sense that I could come to be constituted by another animal or even a non-animal, should my psychology end up there). There is no problem of my sharing any matter with another thing – a 'human animal' as contended by Olson, because there are not two substances there. Persons are higher order things, and in that sense are not independent of the things that constitute them – just as beliefs are not independent of their first-order realising states (see Kim 1993:200), even though their identity conditions differ. When my cerebrum is removed and transplanted into another body I will go with it; the living animal left behind will not be a person at all, and it would be inappropriate to use 'I' to refer to it, even if the operation were to fail.

As far as the foetus problem goes, I was not a foetus. That is the role 'I' is not appropriate to: foetuses – it applies only to persons. But following my story that does not mean the foetus died when I came into

[7] Olson simply asserts that no identity conditions can be provided for persons, ignoring the strong claims of the psychological view to meet our commonsense view on this point.

existence, nor that the foetus continues to live alongside me, for 'I' does not name a substance (or at least not directly). Of course, I may still use 'I' to refer to the foetus from which my substantial animal developed, and everyone will know what I mean. But this is not to commit me to being essentially an animal – it is really just a metaphorical extension from the strict usage, just as in fiction I may refer to the dead body that remains after my death as 'me'; Olson would do that too, even though his theory also implies that the dead body is not me – it is in this same sense that Unger's moron is me[8].

Does it matter that we are not substances in the sense that Olson assumes we must be if we are real things? The debate since Aristotle has seen substances as the things of major importance, but is totally unclear that we lose anything by not falling into that category. We belong to a genuine sort, having identity conditions, and we are continuants – in those ways we have some of the most important features of substances. True, we exist in virtue of the existence of more basic things, but there is nothing terrible about that.

This account makes room for some matters of importance that are left out on Olson's account. For while *person* is no longer to be seen as a fundamental metaphysical concept, it remains a fundamental moral concept. Olson does gesture at our 'practical considerations' which mirror a psychological view of personal identity, but his account of real identity has to play these down. He seems prepared to bracken this out, but there is an enormous amount at stake here. The whole impetus of the debate around personal identity in which Olson locates his theory[9] is provided by the moral consequences and significance of personal identity. When Locke calls personal identity a 'forensic concept' this is part of what he means – it is of

[8] A sense that Unger (1990:13,fn14) explicitly acknowledges: 'And what of our everyday talk of people? Not only do we often say of a person recently dead that he no longer exists, but, in other moods, we say that he is in the funeral parlor. Whatever the linguistic analysis of these utterances, the former sort of discourse is, by far, the better expression of our more serious and our more accurate, thought about ourselves'.

[9] His book is subtitled 'Personal Identity without Psychology', and he calls his theory the biological view *of personal identity* (Olson 1997a: 166).

interest and importance because of its role in law and morality. Outside of that context, it is conceptually has no place. Olson's account attempts to remove it from that context, but that removes all the point from the debate. To lovers my account, we can keep that point, and face the consequences Olson ascribes without embarrassment.

Philosophy
School of Human and Social Studies
University of Natal, Pietermaritzburg

References

Beck, Simon 1997. Humans, Persons and Thought-experiments. *Journal for the Study of Religion* 6.

Beck, Simon 2001. Let's Exist Again (Like we did last Summer). *South African Journal of Philosophy*, 20.

Jackson, Frank 1998. *From metaphysics to ethics*. Oxford: Clarendon Press.

Kim, Jaegwon 1997. The Mind-body Problem: Taking Stock after Forty Years. *Philosophical Perspectives* 11.

Olson, Eric 1994. Is Psychology Relevant to Personal Identity? *Australasian Journal of Philosophy* 72.

Olson, Eric 1995. Human People or Human Animals? *Philosophical Studies* 80.

Olson, Eric 1997a. Was I Ever a Fetus? *Philosophy and Phenomenological Research* 42.

Olson, Eric 1997b. *The Human Animal*. Oxford: Clarendon Press.

Parfit, Derek 1984. *Reasons and Persons*. Oxford: Clarendon Press.

Penelhum, Terence 1970. *Survival and Disembodied Existence*. London: Routledge.

Rudder Baker, Lynn 2000. *Persons and Bodies*. New York: Cambridge University Press.

Shoemaker, Sydney 1963. *Self-knowledge and Self Identity*. Ithaca: Cornell University Press.

Shoemaker, Sydney 1970. *Persons and their Pasts*. In Shoemaker, Sydney (ed): *Identity, Cause and Mind*. Cambridge: Cambridge University Press.

Shoemaker, Sydney 1984. *Identity, Cause and Mind*. Cambridge: Cambridge University Press.

Shoemaker, Sydney 1997. Unger's Psychological Continuity Theory. *Philosophy and Phenomenological Research* 52.

Unger, Peter 1982. Towards a Psychology of Common Sense. *American Philosophical Quarterly* 19.

Unger, Peter 1990. *Identity, Consciousness and Value*. Oxford: Oxford University Press.

Unger, Peter 1992. Reply to Reviewers. *Philosophy and Phenomenological Research* 52.

van Inwagen, Peter 1997. Materialism and Personal Identity. *Philosophical Perspectives* 11.

Wiggins, David 1980. *Sameness and Substance*. Oxford: Blackwell.

Williams, Bernard 1960. Bodily Continuity and Personal Identity. In Williams Bernard (ed): *Problems of the Self*. Cambridge: Cambridge University Press.

Williams, Bernard 1970. The Self and the Future. In Williams Bernard (ed): *Problems of the Self*. Cambridge: Cambridge University Press.

Williams, Bernard 1972. *Problems of the Self*. Cambridge: Cambridge University Press.

Anthropocentric Chauvinism

Jacek Brzozowski

Introduction

The question of colour has been a source of great anxiety for philosophers, as most philosophers are split between two seemingly incompatible positions, both having conceptual appeal, yet neither seeming to wholly satisfy our conceptual schema. The objectivist view appeals to our belief that colours are real properties of objects (either microphysical properties (Jackson 1998) or surface spectral reflectances (Gilbert 1987)) in the external world, while the subjectivist view appeals to our experience of colours (colour as redness, greenness, blueness, etc.), arguing that as these phenomenological aspects cannot be found in the perceiver-independent world, colour must be a product of our mind.

This has led to a third position in the debate. This position knows as the secondary quality view (or dispositionalism) which boasts success in finding a middle ground, saving both the externality and the phenomenology by arguing that colours are secondary qualities in that they consist in a power or disposition to produce a sensory experience in a perceiver, this power being grounded in primary qualities. This is a relational stance as the identification of colours is dependant on the experiences of the perceiver (or at least to the experiences of what is often referred to as the normal perceiver), and if there ceased to be perceivers so too would there cease to be colour.

Recent empirical evidence, from various comparative studies of visual systems across species, has given theorists[1] reason to argue that

[1] See Thompson (1992, 1995) who then argues further that this gives us reason to doubt that any objectivist position can be attained. See also Mather (1997) who proposes a neural realism.

dispositional accounts are motivated by a species-specific chauvinism, and further that this anti-claim can be extended to other objectivist positions. If this is right then it appears that we have good reason to reject the objectivist views as possible candidates for explaining the ontological status of colour.

In this paper I will argue that even though we should reject the identification of colour with dispositional properties, we need not reject all objectivist positions on these grounds.

I shall first, in Section 1, set up the general structure of the dispositionalist view showing its dependence on standard perceivers and standard circumstances. In Section 2, I will set out a brief account of how the visual system works. In Section 3 I will reveal the chauvinism that underlies the dispositional account. Finally, in Section 4, I shall propose that the primary quality view (the alternative objectivist position), by acknowledging the chauvinism in dispositionalism and limiting this only to pragmatic use and by identifying colours rather as disjunctive microphysical properties that surpass our experience of them, avoids similar criticisms of chauvinism within their proposed ontology.

Section 1. *Dispositionalism*

Johnston (1992) argues that both the subjectivist and objectivist views 'are each in their own way perfectly true' (Johnston 1992:222), depending on how inclusively' one talks about colour. His proposed methodology is that we draw out our beliefs about colour, separating out our 'core' concepts (those which we use to define the subject which we are investigating) from the more 'peripheral' beliefs (those that we hold towards the defined subject, yet can be changed without changing the subject). From this we are able to 'rectify' which concepts of colour are central to our colour-beliefs and -practices, thereby identifying those that we are able to give up in pursuit of a coherent theory without changing what we are talking about when we speak of colour.

Following this methodology, a number of philosophers have rejected what they see as the two extremes of objectivism and subjectivism -arguing

[2] *'Even so inclusively speaking* the external world is not coloured. *More or less inclusively speaking* the external world is coloured' (Johnston 1992:222).

that by accepting either one, one must give up too many of our core beliefs – and embraced a (in some way or another) modified Lockean dispositionalism that they argue allows one to salvage both externality as well as phenomenology.

Lockean dispositionalism stems from Locke's distinction (this distinction was earlier made by Boyle, however Locke gave it the mature form from which dispositional theories were created) of primary and secondary qualities, whereby colours (much like tastes, sounds, odours, etc.) are taken as secondary qualities in that they consist in a power to produce ideas in a perceiver, this power being grounded in primary (perceiver independent, physical) qualities.

> Material things interact with one another in regular causal ways; hence we can say that each thing has various powers. To say that a certain thing has a certain power is just to say that it would affect or be affected by another thing of a certain sort in some specific manner. A power is not the cause of such and such an effect; rather to have the power is to be such as to cause the effect. The cause will be some set of properties ... of the thing that has the power. It will generally be, or at least include, some set of properties of the minute parts of that thing, of the collection of particles of which it is composed... That is, material things have powers to produce sensations and perceptions in us, and these powers, like any others, have grounds or bases in the intrinsic properties of things ... *Secondary qualities* ... – of which he gives examples of colours, sounds, tastes, etc. –, he does identify with powers: they are not true in the object themselves but powers to produce various sensations in us by their primary qualities (Mackie 1976:9-12).

Locke however acknowledged "the possibility that "by the different Structure of our Organs" different ideas of colour could be produced in different minds by the same primary attributes" (Ayers 1991:209), thereby if two perceivers, who's sensory organs were sufficiently different, looked at the same object, it is possible that one could have a yellow colour experience while the other could have a blue colour-experience. However recognising this problem, he was careful to make it clear that he did (very reasonably) believe that people's senses in fact function in similar ways to similar effect" (Ayers 1991:209), thereby bringing about a standardisation by dispositionalists of colour perceivers, the standard colour perceiver being any one of us who's perceptual system functions (more) like the majority of our perceptual systems function. Through reference to standard perceivers, dispositionalists then avoid criticisms from (imagined) possible cases of inverted colour perception as well as from the more common cases of colour-blind individuals (these being non-standard perceivers).

Secondary quality theories of colour therefore (roughly) hold that the colours of objects are the dispositions (Locke referred to these as 'powers') of these objects' surfaces (oddities/funkish dispositions aside[4]) to

[3] I believe we have good reason (though I shall not argue this point in this paper) to reject subjectivism in that our use of colour terms is not in reference to the natures of our colour experiences, but rather to features of objects that we take as causing these experiences (the feature which our colour-experiences represent), where "[t]he evidence for this is that we look at the objects, we do not introspect ... [and further] we take it for granted that it is a good idea to look carefully, and in good light, and that others may be in a better position to rule on the object's color than we are" (Jackson 2000: 153f). By denying this intuition, thereby accepting a subjective stance, one would have to claim that the world is invisible, or at least that we do not *see* the objects, as we *see* objects by seeing them as coloured (as having a colour-property that causes us to have a visual experience). As our pre-theoretical colour terms are based upon the assumption that the world is coloured, a denial of our objectivist intuition would therefore involve a radical change in our colour language.

[4] Lewis (1997) argues for a way to reform the conditional analysis of how things are disposed to respond to stimuli, in light of oddities/funkish dispositions (a disposition that when put to the test would vanish as the stimulus for it's manifestation may cause the disposition to go away and therefore prevent its manifestation). His proposal is that: "Something *x* is disposed at time *t* to give response *r* to stimulus *s* iff, for some intrinsic property *B* that *x* has at *t*, for some time t', *x* and *x*'s having of *B* would jointly be an (at least) complete cause of *x*'s giving response *r*" (Lewis 1997:157).

prod. be certain perceptor experiences (Locke referred to these as 'ideas') of colour in standard viewers under standard viewing conditions (standard conditions are those conditions under which objects are usually observed, thereby avoiding problems of colours viewed under darkness or through coloured glasses, and so on). Thereby dispositionalists attempt to incorporate both the causal properties of objects that underlie the disposition, as well as the 'ineliminable subjective element in the analysis of colour' (Thompson 1995:31). Modifications of this account therefore tend to vary mainly in where they choose to place emphasis: either on the physical surfaces (thereby more in the objectivist camp; for example Johnston 1992[5]), or on the perceived colours (thereby more in the subjectivist camp; for example Harvey 2000). However with this general structure underlying most dispositional theories, dispositionalist hold that they are provided with 'grounds for speaking of objects' being "coloured", meaning that they *produce some percept of colour or other in viewers*' (Harvey 2000:138).

Dispositionalism is therefore a relational theory, in that it relies on physical properties of objects, the (standard) circumstances in which perception is taking place, the physical perception mechanisms of (standard) perceivers, as well as the phenomenological (subjective) aspects of the colours perceived. The disposition is then used to unify these separate aspects (to a lesser or greater degree): the physical properties of objects that are coloured play a causal role (reflecting different wavelengths of light) in such and such circumstances, stimulating the physical perception mechanisms (the photosensitive pigments found in the cones), which then have a causal effect on the post-receptoral channels of such and such perceivers, causing these perceivers to have such and such experiences (through brain stimulation from the perceptual mechanisms) of colour (red-ness, green-ness, purple-ness, etc.).

[5] He proposes that colour be taken as a constituted response-disposition, which he defines 'as a higher-order property of having some intrinsic properties which, idealties aside, would cause the manifestation of the disposition in the circumstances of manifestation' (Johnston 1992:234). Therefore an object has intrinsic properties, which are the constituting basis of its disposition to produce a mental response in such and such a subject under such and such circumstances.

I have left the dispositionalist account above rather general, thereby allowing it to accommodate most dispositional theories. Most alterations to this general structure are responses to various objections, and depend on the dispositionalist's metaphysical leanings, though I will ultimately argue that it is this general structure that is inadequate in offering us the ontological status of colour.

Section 2. *The Visual System*

Before we turn to what has been argued should be seen as an unfounded bias, which motivates the desire for a dispositionalist account (I call this 'anthropocentric chauvinism'), we need to first review (briefly and somewhat abstractly) how the visual system functions.

Light of various wavelengths is reflected from the surface of objects as well as their surrounds. This light enters the eye and stimulates the photoreceptors/cones (S, M, and L wave sensitive pigments (having ~445nm, 535nm, and 570nm peak sensitivities respectively)). These pigments are however not individually responsible for our seeing spectral stimuli as blue, green, red, or yellow. Collectively these pigments respond throughout the spectrum, but the visual system responds only to the differences in their relative levels of activity (Thompson 1995:33f), whereby the pigments give off response-signals (S, M, and L receptor signals, known as cones) that are compared (roughly through addition and subtraction) to determine their relative strength, and then recoded in the three post-receptoral channels: An achromatic channel (L+M; as there is no difference measured, it is not spectrally opponent, therefore does not signal difference in wavelength), which determines the luminance of the colour perceived, and two spectrally opponent chromatic channels, a red-green channel (L−M), and a yellow-blue channel (S−(L+M)).

Each channel can be seen as a relative visual response curve. The achromatic response curve determines the whiteness (also known as lightness or brightness[6]) component, in that it specifies the amount of energy

[6] Though, as Thompson (1995:146n) points out, brightness refers to colour appearances (due to dazzling) in the aperture mode, while lightness refers to colour appearance (grey scale of black to white) in the surface mode.

necessary at each wavelength for a given observer to first detect the stimulus (Thompson 1995:62), this being to the exclusion of chromatic responses. The two chromatic response curves then represent the four basic hues (red, green, blue, yellow), with red and green as opposites on one curve, and blue and yellow as opposites on the other curve, each curve passing through neutral balance points where the chromatic response of that curve is zero, and therefore that channel is nulled. Colour experiences are then determined by the interactions of these channels. If, at a wavelength (around 475nm), the red-green curve crosses its neutral balance point, then the red-green channel is nulled, and the colour experience is determined by the value of the corresponding blue-yellow channel. If the value of the blue-yellow channel is negative, the chromatic response will be blue (as the blue is below the null point while yellow is above the null point, this is however merely convention, and one could indicate the blue as positive and the yellow as negative, as long as one adjusted the red-green curve accordingly). As, in this case, the red-green curve is nulled, we have a unique blue, a blue that has neither red nor green. Further there are points on the curves where the two curves cross each other, either in the positive (whereby red and yellow cross) or in the negative whereby blue and green cross). In such cases one experiences a colour that contains equal quantities of the two hues, this is known as a balanced binary colour (i.e., red and yellow, will result in a balanced orange). Finally, the saturation of the colour experienced at a given spectral location corresponds to the ratio of the responses of both chromatic channels at that location to the sum of the chromatic and achromatic responses at that location (Thompson 1995:63).

We can thus divide the visual system into three parts: the receptoral colour space, determined by the various possible cone triplets, the post-receptoral colour space, determined by the three visual response curves corresponding to the three channels, and a phenomenal colour space, determined by the phenomenal aspects of visual experience (i.e., the experience of colour as the combination of hue, saturation and brightness/lightness/whiteness)[4]. It is not altogether clear that one must

[4] The visual system is far more complicated than what I have given above, and there are other visual spaces (such as cortical colour space), though what I have given is sufficient for the purpose of the following discussion.

divide the receptoral from the post-receptoral colour spaces as these two can quite easily be worked as one single colour space, a process of "stimulation-response-opponent processing". Matthen states "Information about spectral distribution is extracted from the outputs of [the L, M, and S] cells by sampling the relative strength of a signal in selected wavebands. This process, called "opponent processing" (in effect) computes the following functions: (L+S)- M, (L+M)- S" (Matthen 1999:19)[5]. The first function gives us the red-green opponent pairs, where the second function gives us the blue-yellow opponent pairs). I have chosen to divide them merely for ease of explanation of the move from stimulation to phenomenology, though it is not important to this argument whether they are treated like this or not, as long as one takes note of some 'stimulation-response-opponent processing' process.

Section 3. *Chauvinism*

The case given above is what is often referred to as colour perception by a standard perceiver by dispositionalist accounts. If an object reflects light which stimulates the three cone-types, which in turn send a triplet response signal that is recoded in the three post-receptoral channels that determine the hue, lightness, and saturation that is experienced by the perceiver, as long as the perceiver has these colour spaces formulated (properly) in this way, and the experience in the phenomenal colour space corresponds, more or less, to that of the experiences of other perceivers looking at that object, then that perceiver can be said to be seeing the colour of that object, as colour is the disposition of an object to manifest such and such an experience in a standard/trichromatic perceiver (one having three cone types cross connected to three post-receptor channels) in standard conditions.

We can now turn to the case of colour-blind individuals. These are individuals who due to a failure in the functioning of their photoreceptors (lack of function or non-existence of one of the L, M, or S pigments in their cone) are not able to have certain colour experiences. They are known as dichromats, as they only have the function of two of their cone-types

[5] Thompson (1995:66) speaks of 'trivariance' to refer to the 'receptor-channel linkage'.

Colour-blindness comes in two main forms: 1) individuals not able to distinguish colours along the red-green phenomenological axis, these are either protanopes who cannot see red due to lack of L pigment, or deuteranopes who cannot see green due to lack of M pigment; 2) individuals not able to distinguish colour along the blue-yellow phenomenological axis, these are tritanopes. These are therefore individuals who are classified as non-standard perceivers, as their visual systems are somewhat lacking and therefore they are not able to have the colour experiences that trichromats have when faced with a certain object. If a deuteranope looks at a green object, the dispositionalist would say that the deuteranope does not see the green as the perceiver is non-standard, therefore does not see the disposition to manifest such and such an experience in a standard perceiver. However, the dispositionalist will insist that even though the deuteranope does not see the object as green, the object is nonetheless green, in that if the perceiver were standard then the object would be disposed to manifest a green experience in that perceiver.

This amounts to the claim that the trichromat, who has a better visual system (a visual system that has three functioning cone-types, and is therefore able to better discriminate between colours, offering a greater range of colours), is able to see colours that the dichromat is not able to see. But now let us imagine an individual who is an even better colour discriminator[9] than the trichromat, a tetrachromat (who has four visual pigments/cone types). One could imagine that this individual had an extra visual pigment that was sensitive to wavelengths that fell in the ultraviolet end of the spectrum. This individual would therefore have UV, S, M, and L wave sensitive cones. This individual is sensitive to the wavelengths that the trichromat is sensitive to, and in addition, is sensitive to an extra wavelength, the UV wavelength, that through the limitations of the trichromat's visual system, the trichromat is not sensitive to. Thereby it is argued that the tetrachromat phenomenological colour space, of which the trichromat's phenomenological colour space (consisting of unitary and binary hues) has an extra dimension (one that could have ternary colours) that the human phenomenological colour space thus lacks. One can therefore conclude

[8] Averill has a similar idea, referring to 'unusual human observers' (Averill 1985:390).

from knowledge of how the visual system works as a 'stimulus-response-opponent processing' process) that the tetrachromat is able to see colours—'UV colours'[10]—that the trichromat is not able to see. Imagining such a situation does not require a great stretch of one's imagination, as when one looks to nature one discovers that many species are tetrachromats (e.g. pigeons[11]), and therefore as we have evidence of animate beings that are dichromats, the possibility of human beings who are tetrachromats is not a great leap of faith[12].

Now let us assess the situation as hand. We have a dichromat (the deuteranope) who cannot experience green due to a lack of M pigments. We have a trichromat that can experience green, as the trichromat has S, M, and L pigments, though s/he cannot experience UV colours as s/he lacks UV pigments. Finally, we have a tetrachromat who can experience both green and UV colours, as s/he has UV, S, M, and L pigments.

The problem for the dispositionalist should now be clear. The dispositionalist defines colour as a disposition of an object to manifest such and such an experience in a standard perceiver. Therefore the dispositionalist would have to say of the situation I have described, that the trichromat sees green where the dichromat fails to see green, while the tetrachromat sees green but only as a UV illusion. It is unclear why the UV colours should be labelled as illusions. After all, the reason that the tetrachromat can see UV colours is not because their visual system is malfunctioning, but rather because it is a better (of greater sensitivity) discriminatory device, in that it can cover a larger part of the visual spectrum than the visual system of the trichromat. This is the same reason why the trichromat's visual system is taken as better than the dichromat's visual

[9] As it is not exactly clear what these would be like phenomenologically, when I refer to them as UV colours all I mean is that they involve light from the near-UV (shorter-wavelength) end of the visual spectrum (wavelengths that do not feature causally in the human visual system).

[10] See Thompson (1995:148-150), and Mather (1999:81). It was been suggested that pigeons may even be pentachromats, though many other avian species are tetrachromats (therefore could be substituted for pigeons above).

[11] See Thompson (1995:166-168) for a discussion on the real possibility of tetrachromacy in a portion of the human female population.

system. Saying that the tetrachromat experiences a UV illusion is as odd as claiming that the trichromat experiences a green illusion. The dispositionalist is forced into holding some mysterious double standard. I am arguing that maintaining this double standard is motivated by what I am calling anthropocentric chauvinism. Once again it is the bias of seeing humanity (or the major part of humanity-as-it-is-now) as the centre of the universe.

The double standard becomes even more mysterious when one looks at colour language. The dichromat is not able to know what the experience of green is like, yet is still able to speak about green things, such as the 'go' (green) traffic light, and actually believes that that light is green, even though she cannot experience it's green-ness. Why then is the trichromat not in the same position when speaking about UV colour? Surely commonsense dictates that as in the case of the deuteranope and green, the trichromat should say that certain things are a UV colour, though not being able to have a UV colour-experience. Just as the deuteranope would have to ask someone if a certain object was green, and then only after being told that the object is green (by someone with a better visual system, i.e., a trichromat or a tetrachromat) refer to that object as green, so too the trichromat would have to ask someone if a certain object is a UV colour (someone with a better visual system, i.e., a tetrachromat) and only then refer to that object as a UV colour. Through advances in science, one can no longer appeal to the argument that 'whereas that dichromat can, through language, appeal to the trichromat's experiences about the existence of extra colours, the trichromat has no one to appeal to' as through an understanding of the workings of the visual system, one can conclude that tetrachromats do see other colours, even though we are not able to conclude what the colour experiences caused by these colours would 'look like'. When conceptualising trichromats and tetrachromats, one would not have to conclude, as some dispositionalists would have us believe, 'that the best way to describe the two groups of viewers is to speak of their having two different colour-languages' (Harvey 2000:54), in fact one would be mistaken to do this. They *would* share a colour language - speaking about which objects are which colours (like the dichromat and the trichromat both speaking of the 'go' traffic light as green), though they would not share certain colour experiences. This however is merely a difference in the phenomenological colour space, which

is either more or less sensitive to colours depending on the receptoral and post-receptoral colour spaces of the particular organism.

If one looks at colour concepts in the real world (our trichromatic human being aside), then when told of how the visual system works, and then told that pigeons do have an extra UV-sensitive pigment, and therefore have an UV experience [3], one is *not* cast into some conceptual confusion. As Smart (1961:135) points out:

> Since Wittgenstein we surely know better than to think of a word's having meaning as consisting in its evoking a certain sort of mental image. Meanings are not mental images [rather] the meaning of a word consists in its *use*, not in its associated imagery.

One is therefore able to accept that the pigeon does see extra UV colours, even though we are not able to know what its colour experiences *is* like. This merely places us in a position similar to that of the deuteranope who is faced with the 'go' traffic light, which is a perfectly acceptable position, a position that in no way conflicts with (but rather, is a part of) our commonsense view.

Just as we once held that the sun revolves around the earth, now the dispositionalist wants us to hold that colours revolve around our (trichromatic) visual system. Both the former and the latter claims can be attributed to mere anthropocentric bias, an appeal to some kind of 'human-special-ness'. The dispositionalist has no justification as to why we (trichromatic) human beings should have some special colour-perceptibility. As we have seen from science, other species have visual systems that are more sensitive than our visual system. As we have seen from colour language, we *do* separate out colours from colour experience (as in the case of the dichromat). Therefore as we are able to speak of objects having colours other than those that we experience those objects as having, and as we do speak of visual systems that are more sensitive as giving us a greater ability to discriminate between colours than those less sensitive, we have no reason (as the dispositionalist would have us say) to conclude that tetrachro-

[3] As the tetrachromat probably has three post-receptoral chromatic channels this could be a situation where both the red-green and the blue-yellow channels are nullen, while the UV channel has a response value.

cats when responding to UV light, not having a colour illusion.

> A general conception of colour and colour experience should allow us to treat the human and the pigeon systems as instances of a general kind. Or else it should give us a principled reason for excluding the pigeon (Madden 1999: 51).

As no principled reason is given it seems that the dispositionalist is not just flat-footedly appealing directly to standard human observers in arguing for an ontology of colour. The dispositionalist needs to acknowledge this and provide, within existing metaphysics, some grounding for a theoretical distinction.

Section 4 *The Primary Quality View*

Objects are disposed to look coloured, this is more in that some objects do look such and such a colour to such and such a perceiver under such and such circumstances. The argument that I have presented has not left colours as dull. Dispositions are however relational and highly relative concepts. A colour-disposition is a relation between properties of an object and a certain kind of observer in certain circumstances. Therefore an object could be disposed to look red and such a colour to a cat and such a one over under such and such circumstances while disposed to look green and such a different colour to a human but to look other different colours to other perceivers in other circumstances. This is not to appeal to chameleon colours, but to emphasise the complexity of colour.

The primary quality view of colour, at first sight, appears to be at odds with the secondary quality view. Primary quality theories (following Jackson 1998) identify colours with certain physical properties (perhaps very complex disjunctive ones) that underlie the representations and the tasks described with representations[4] in phenomenologically robust colour vision (leaving aside, unfortunately, the difficult problem of how it is that we are to conceptualise certain

[4] Though the question of what it means to have such representations, and exactly how the primary qualities are represented is another problem altogether, and would require a separate paper to be adequately discussed.

properties that are represented in colour-experience by typically causing colour-experience. This separation may at first seem slightly counterintuitive, though I believe that even though colour and colour-experience are often conflated by us into a single concept, this is a mistake on our part made largely by the fact that these two concepts do not often have occasion in our day to day life to come apart. I shall however argue that when we analyse our concepts of colour we find that these two concepts do come apart, and evidence for this is given through language use, in that they often are taken apart in language when need arises in certain circumstances. It is precisely in circumstances where we speak of extra colours (above) that we can see our concepts coming apart, in that here we address colour as separate from colour-experience. It is this separation of our concepts that made these circumstances problematic for secondary quality theories.

Despite the problems that the secondary quality view has in deviant circumstances, it did have initial conceptual appeal, and therefore if one is to uphold our concepts of colour, one should maintain that dispositions are in some way relevant. Prior, Pargetter, and Jackson (1982) argue successfully that when it comes to causation of manifestation it is *not* the disposition that is doing the causing, as dispositions cannot (strictly) cause their manifestations, but rather it is the first order properties that constitute that disposition that cause the experience. Jackson and Pettit (1990) even argue further that even though the dispositions are not causally efficacious (in that it is the constituting bases that do the causal work) they may still be causally relevant, in that their 'realization programs for the realization of a lower order efficacious property' (Jackson & Pettit 1990:115). This works, in that by this higher-order property being realized it ensures that there is some lower-order causally efficacious property that is doing the work.

Dispositions thereby can be seen as causally relevant in that an object's disposition to manifest a colour-experience in a certain kind of perceiver under certain circumstances informs that perceiver that the object has a (causally efficacious!) colour property that is causing them to have such and such a colour experience. Colour-dispositions are therefore useful (pragmatic) in that they provide a group specific account of the perceiven colours of certain objects under certain circumstances. The problem however was that the dispositionalist theory is too reliant upon these perceived colours, the particular group of perceivers and the circumstances in which

they are perceived. The dispositionalist account was left trying to do too much in that it tried to incorporate our colour and colour-experience concepts under the single label colour from a single group perspective, and because of this it was not flexible enough to accommodate our extended (non experiential) colour concepts adequately, instead it imposed certain counterintuitive positions upon us when circumstances deviated from the norm.

It is for this reason that one should rather conceptualise colour dispositions merely as heuristics for identifying colour, as each colour disposition of each object would provide a certain group of perceivers (those being united by similarity in their visual system) with information about which colour they are observing, this information is in most cases somewhat crude as it is limited according to limitations of that groups visual system. Dispositions thus are causally relevant in informing a group of the colours *for that group*, though this does not mean that that group has full information of the colours proper.

I propose that groups of colour perceivers should therefore be placed on a continuum, with dichromats at the one end (as this is the minimum visual state to experience colour) and *n*-chromats at the other, with trichromats, tetrachromats, and so on, in between. The positing of *n*-chromers does not however mean that this continuum is endless, with each organism with ever increasing number of kinds of photoreceptive pigments (and corresponding post-receptoral and cortical visual systems) placed further and further along the continuum (ad infinitum), as after a certain *n*-chromatic system has been reached, an *n+1* chromatic system would no longer offer the organism any visual advantage over the *n*-chromatic system, as the difference in sensitivities of the visual systems would be finer than the difference in colours, therefore both visual systems would be equally efficient in discriminating between colours by representing all colours in colour experience (where the continuum ends is however a question for science to solve). This is just part of our (extended) commonsense intuition, as shown in the case of extra colours, whereby the trichromat (standard human perceiver) discriminates more precisely and represents more accurately the colour properties of objects than a dichromat (a colour blind human perceiver), while a tetrachromat (i.e., a pigeon, or our imaginary tetrachromat human perceiver) would discriminate more precisely and

represent more accurately the colour properties of objects than a trichromat.

One can therefore see the way that colour properties of objects and the colour experiences come apart. The primary quality view therefore maintains that colours are primary qualities of objects that cause (in being causally efficacious) colour-experience. The question that one then needs to ask is whether the primary quality view can then account for our commonsense colour beliefs in standard cases as well. Johnston (1992:334-336) points out, through his example of Zinka the canary and a colour photograph of Zinka, that the primary quality account would have to maintain that as the physical property (P_1) which causes the canary yellow experience when looking at Zinka is very different from the physical property (P_2) which causes the canary yellow experience when looking at a photograph of Zinka, canary yellow must be a disjunctive property consisting of disjuncts P_1, P_2, as well as other disjuncts that cause canary yellow experiences. The problem for the primary quality view then, is that our belief commonsense belief is that yellow is the property that causes yellow-experience, however it is dubious to say that such a disjunctive property is the cause, as it seems to put forward to many different possible causal candidates. This is mostly, though not entirely, right. Jackson (1998) argues that disjunctive properties can be causes, using the example of the depth of a wound that is responsible for the death of a victim, arguing that it is not the precise depth that is important, but rather that the depth fall in the range of depths that would count as deep enough to be fatal, claiming further that 'it is arguable that most things we cite as causes are more or less disjunctive' (Jackson 1998:106), though adding that such disjunctive properties are causes as long as 'the disjunction is not excessively disjunctive' (Jackson 1998:108). When faced with disjuncts that are excessively disjunctive, one rather makes sense of these as separate causes, whereby the claim is about either one or the other disparate disjunct, where an excessively disjunctive claim such as 'either a 10cm knife wound by Frank or a 15cm bullet wound by Mark caused the death of the victim' should rather be understood as 'either a 10cm knife wound by Frank caused the death of the victim or a 15cm bullet wound by Mark caused the death of the victim'.

Returning to Johnston's example, the reason why I refrain from saying that his analysis of the primary view commitment is right, is that he

cites the property P_1 as 'very different' (Jackson 1998:235) from P_2.[5] If these properties were very different (though one need not say that they are), then one would have to conclude that the disjuncts are excessively disjunctive, and thereby have to rephrase the situation as done with the knife wound and the bullet wound above: either P_1 caused the canary yellow experience or P_2 caused the canary yellow experience. P_1 and P_2 should then be identified as two different colours as, through being excessively disjunctive, they are separate causes. However the two properties do cause a single colour experience in human beings, which is identified by us, phenomenologically, as a single canary yellow-experience. We therefore need to accept that even though they do cause a single colour-experience (canary yellow experience) in human beings, P_1 and P_2 should however remain separate colours (canary yellow$_1$ and canary yellow$_2$) as they have separate causal properties that cause the same effect in human beings *only because of the limitation of the human visual system*. These would however cause two different effects in beings that have visual systems that are finer discriminators (are more sensitive) of colour. Those beings would then have a separate colour-experience/representation of canary yellow$_1$ from the one our experience/representation of canary yellow$_2$[6] (thereby they would have the ability to separate out canary yellow$_1$ from canary yellow$_2$ through perceived/phenomenological differences between the representations caused by the two properties). We should therefore accept that two objects that we see as having the same visual colour could actually differ in colours. Further

[5] This would however be the view held by primary quality theorists such as Smart (1975), whereby Smart identifies colours as disjunctive physical properties underlying our (normal observer) colour experience under normal lighting conditions. Therefore Smart would have to accept that even if the microphysical properties causing Zinka's canary yellowness and the microphysical properties causing the yellowness of the photograph of Zinka were to me to be very different (excessively disjunctive), they would still be the same colour.

[6] By using 'represent' I do not here tie myself to the view that the representation is the object of perception (i.e., a sense datum view), but rather that the external object of perception is perceived by me by being represented in a certain way through the use of my visual system.

if this difference were pointed out, it could then be that in some cases we should refer to them as separate colours (which however cause the same experience in us). One could imagine that the male canary (which we can imagine has the ability to discriminate between canary yellow$_1$ and canary yellow$_2$) is attracted to female canaries because of their canary yellow$_1$-ness, while is simply unmoved (or even repelled) by canaries that have been painted with a canary yellow$_2$ paint (because of a dislike of the canary yellow$_2$ colour, and not because of it being paint etc.) In such cases, when speaking about the mating activities of canaries, an ornithologist would say that the colour that canaries are attracted to is canary yellow$_1$, while they are repelled by canary yellow$_2$. An example (given by Jackson 1998: 112) that lends plausibility to this explanation (as a way that we do actually react to similar cases) is that of the ornamental stone jade, whereby when it was discovered that there were two different forms of jade – nephrite and jadeite – this did not lead people to denounce the existence of jade, rather one says that there are two kinds of jade. The lay person, may refer to these as both being jade, and in a way s/he would be right, though a mineralogist would then be able to correct them and exclaim that there are in fact two different kinds of jade where the lay person saw only one. So too then could an ornithologist correct a painter who by using canary yellow$_2$ paint believed s/he was painting a wall the same colour as that of a canary. This however, does not mean that the painter need say that the room is not canary yellow, s/he needs merely accept that technically it is a different canary yellow to that of a canary.

One does not then however have to extend this division to every disjunct of the disjunctive property, in that one would not be compelled to claim that each strawberry red disjunct is then a different colour, i.e., strawberry red$_1$, strawberry red$_2$, strawberry red$_3$. The primary quality theorist is *only* compelled to say that the excessively disjunctive properties are separate colours, as argued above, though disjuncts that are not excessively disjunctive would be labelled as a single colour. I have equated colours with microphysical causal properties, therefore those properties that are not excessively disjunctive are properties that are unified under their causal role (though as noted above, this causal role is not a role identified by the effects on a human being (trichromat), but rather on the effects of a tetrachromat). What therefore prevents a colour from being an excessively

disjunctive property is that it is highly implausible that excessively different disjuncts would be able to play the causal role required for a disjunctive property to count as a single colour. It is therefore the role that restricts which causally efficacious microphysical properties fall into which disjunctive colour property. A point that needs to be noted is that even though I speak of microphysical properties that cause experience in the *x*-chromat this does not mean that there has to be such an *x*-chromat, or any perceiver at all for that matter. Colour properties can be causal even if they cannot have the colour-experience effect (i.e., they cause light to be reflected in such and such a way; one could here possibly appeal to something like Hilbert's (1987) Surface Spectral Reflectance[17] as the common cause that holds the disjuncts together). The use of the *x*-chromat is merely there to indicate that the physical colour properties are disjuncts that are grouped together by some commonality between them that surpass our experience of them and that can be adequately discriminated by a perceiver if such a perceiver has an appropriately sensitive visual system. Colour properties are however (external real) properties that are not relational (in the sense that they would exist in a world where there were no colour perceivers) even though they do (in a world where there are such and such colour-perceivers) play a relational role in causing colour-experience in perceivers.

Conclusion

> Perception of difference points strongly to the real existence of such differences. Failure to perceive differences points much less strongly to the absence of differences (Armstrong 1968:286).

From comparative studies we have found that species with visual systems that are more sensitive than the human visual system are able to make finer discriminations between colours and are therefore, in some cases, able to see two separate colours where we see only one. This revealed that

[17] Surface Spectral Reflectance is the percentage of light that a surface reflects at each wavelength across the entire visual spectrum (from 300nm-700nm)

dispositionalist theories of colour are motivated by an unfounded anthropocentric chauvinism that is unable to account for this empirical evidence, as well as unable to deal with our extended colour concepts.

However, by acknowledging both that dispositions are species-specific causally relevant properties that program for the underlying causally efficacious microphysical properties, and that our visual systems are limited discriminatory systems, the primary quality account is able to save objectivism. The primary quality account holds that colour properties are disjunctive (microphysical) properties (the disjuncts are grouped together as a disjunctive property by their common, perceiver independent, causal role) that typically cause colour-experience. The experience is however limited by how fine a discriminatory visual system is representing the causal properties. As the human trichromatic visual system is limited, we shall often represent two different disjunctive properties as a single colour-experience. However to claim that this means that these are metaphysically a single colour would be indulging in anthropocentric chauvinism.

Through limitations in the human visual system, the primary quality theorist must accept that their colour-perception is corrigible (this is however no different from perception in general) and thereby must accept that when in day to day activities she uses colour terms these are (for pragmatic reasons) based in an anthropocentric chauvinism (as colour words were made by human beings and used to identify colours through anthropocentric colour experiences). However as our colour concepts *do* notion our colour-experience, the primary quality theorist is able to maintain that colours, metaphysically, surpass our experience of them, and that if we were finer discriminators of colour we would then be able to identify a greater array of colours in the world[18].

School of Human and Social Studies
University of Natal, Pietermaritzburg
Research School of Social Sciences
Australian National University

[18] The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the National Research Foundation.

References

Armstrong, DM 1968. *A Materialist Theory of the Mind. Revised Edition.* London: Routledge.

Averill, EW 1985. Color and the Anthropocentric Problem. *The Journal Of Philosophy* LXXXII, 6 (June):281-30.

Ayers, M 1991. *Locke Volume I: Epistemology.* London: Routledge.

Harvey, J 2000. Color-Dispositionalism and Its Recent Critics. *Philosophy and Phenomenological Research* LXI, 1:137-156.

Hilbert, DR 1987. *Color and Color Perception: A study in Anthropocentric Realism.* Stanford: Center for the Study of Language and Information.

Jackson, F 1998. *From Metaphysics to Ethics: A Defence of Conceptual Analysis.* Oxford: Clarendon Press.

Jackson, F 2000. Philosophizing about Color. In Davies, S (ed): *Color Perception: Philosophical, Psychological, Artistic, and Computational Perspectives.* New York: Oxford University Press.

Jackson, F & P Pettit 1990. Program Explanation: A General Perspective. *Analysis* 50:107-117.

Johnston, M 1992. How to Speak of the Colors. *Philosophical Studies* 68,3 (December):221-263.

Lewis, D 1997. Finkish Dispositions. *The Philosophical Quarterly* 47,187 (April):143-158.

Mackie, JL 1976. *Problems from Locke.* Oxford: Clarendon Press.

Matthen, M 1999. The Disunity of Color. *The Philosophical Review,* 108:4 (Oct.):47-84.

Prior, EW, R Pargetter & F Jackson 1987. Three Theses about Dispositions. *American Philosophical Quarterly* 19,3 (July):251-257.

Stuart, JJC 1961. Colours. *Philosophy* 36:128-142.

Smart, JJC 1975. On Some Criticisms of a Physicalist Theory of Colors. In Chung-ying Cheng (ed): *Philosophical Aspects of the Mind Body Problem.* Honolulu: University of Hawaii Press.

Thompson, S 1992. Novel Colours. *Philosophical Studies* 68. (December): 321-349.

Thompson, E. 1995. *Colour Vision: A Study in Cognitive Science and the Philosophy of Perception.* London: Routledge.

No Time Like the Present: A Cognitive Approach to Time Differentiation in Discourse

Marie Spruyt

Introduction

The term 'cognitive' is used in linguistics to refer to the perception that aspects of human experience and cognition are reflected in the structure and functioning of language. Langacker (1987a; 1987b; 1988) establishes a connection between research done since Fillmore (1968), Rosch (1973; 1975; 1978), and others like Johnson (1987). Lakoff (1987), Fauconnier (1985; 1985) and Givon (1979; 1982), by defining one of the main principles of cognitive grammar: meaning must be reduced to conceptualisation, which implies, inter alia, that the semantic structure of a language reflects our conceptualisation of a physical, social and language-related world. A semantic structure is therefore a conceptual structure that takes cognizance of the fact that language functions in a people directed world, so that one can expect to find signs thereof in the language system that one studies.

Image Schemas

The theory of image schemas within cognitive linguistics were formulated for the first time by Johnson (1987). Edelman (1999:109-148), a renowned neurologist, refers to Johnson's work as an explanation for the way in which we construe mental images, by the use of image schemas that are related to bodily attitudes in the organization of thought and language which are reflected as metaphors in language. To prove that image schemas are important in language use, Edelman refers to Johnson's (1987) study of propositional structures which can only be realised as a complex web of non-

propositional schemas structures based on human experience. We will therefore recognize objects, events and stories because we think in images schemas which are based on sensory and motor experience (Lakoff 1987; Johnson 1987; Talmy 1988; Turner 1991; 1996). The container image schema, for instance, allows us to recognize several objects as containers – not only bags, packets, cups, bottles, a valley, a drawer, rooms or houses, but also our bodies and our heads.

An event can be construed by using several combinations of image schemas as conceptual basis. Movement, for instance, can be construed in terms of a source-path-goal schema, with linguistic units such as *go somewhere, come here*, or as a near-far schema with units such as *far back in the past, near colleague*.

The container-schema with concepts such as in/out/over/inside/outside involve linguistic units such as *in discussion*, or *in town*. In terms of this schema, time can be viewed as a container with three different compartments for what we experience *in* the present, what we remember *out of* the past, and our thoughts *on* what the future holds. The image schema A PERSON MOVES ALONG A ROUTE TO A GOAL can similarly act anaphorically as basis to make time concepts more understandable—one can reflect *back* into the past or look *forward* to the future.

Turner (1995:150) compares time with a story that is projected in time and space, by means of which the experiencer will have a specific temporal focus and a specific temporal viewpoint. Because space has a visual basis and is experienced as concrete, it is commonly used to make time, which is more abstract, understandable. When we say that cycles in life repeat themselves, we use the visual image schema of a circle for our experience of time. We easily think of events connected in time by using intrinsic spatial image schemas such as continuity, extension, discreteness, completeness, cycles or whole and part relations.

The theory of mental spaces (Turner 1996:87) makes it possible to make time as an abstract concept understandable by projecting it as a concrete spatial concept onto an abstract concept. By projecting the concrete concept from a source mental space onto a more abstract concept in a target mental space (the input spaces), a synthesis of the concepts is made in a third space, the blended space, as set out in Figure 1, which is a combination of both the target space and source space. The conceptual contents of the blended space represents the new insights that one acquires of the abstract concepts in the target space. However, it is not possible to blend two concepts without some

counterpart connections between them to guide the blending, and the abstract structure shared by input spaces resides in a generic space which indicates the counterpart connections between the input spaces.

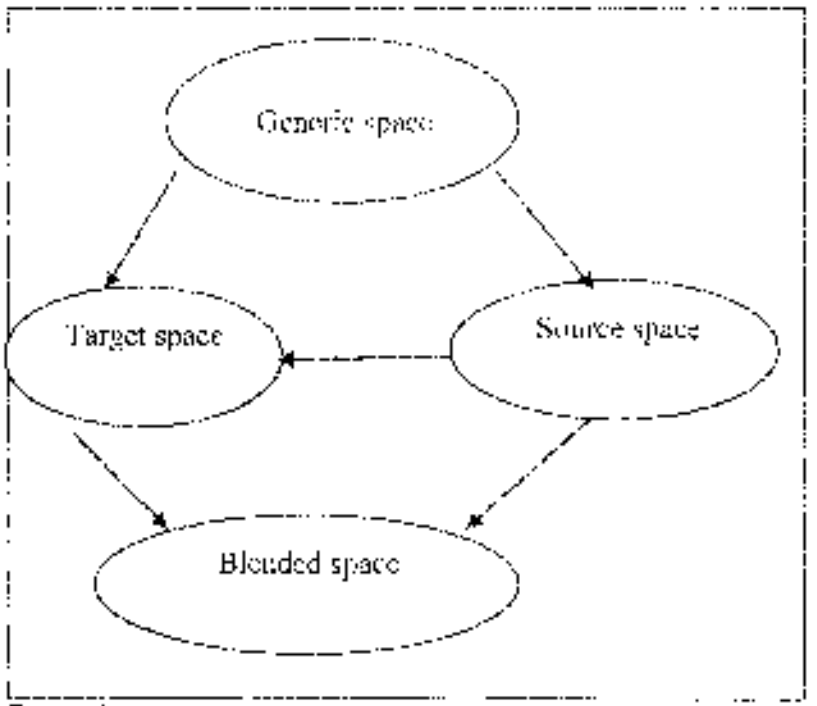


Figure 1: The Source, Target, Generic and Blended Spaces in Conceptual Blending

The Construction of Mental Spaces

A basic point of departure is the theory on the construction of mental spaces (Fauconnier 1985; Lakoff 1987; Langacker 1987a; 1987; 1991) which corresponds with understanding of a sentence within a context. These spaces can be pictures, beliefs, anticipations, stories, propositional realities, and thematic or quantifier domains of situations in time and space. Each space is a version of a logical, coherent situation or potential reality, where it is accepted that several propositions are true, that objects exist and that there are relations between objects (i.e. thematic roles like agent, patient, route, goal, etc.). The

interpretation of a verse depends on the construction of a complex configuration of hierarchically related spaces—as each sentence is processed in the discourse, the configuration of spaces are adapted, based on lexical and grammatical triggers in the sentence. This adapted configuration of spaces is extended pragmatically from background knowledge held in the form of image schemas within event frames.

Event Frames

The notion of event frames, as shown in Figure 2, is used in cognitive grammar to refer to conceptual elements which are combined when we refer to an event. Briefly put, we construct conceptual mental spaces about reality or potential reality; these mental spaces are extended by means of image schemas within event frames; within the event frames the participants in discourse share the same theme on aspects like time, place or occasion, so that effective communication can take place.

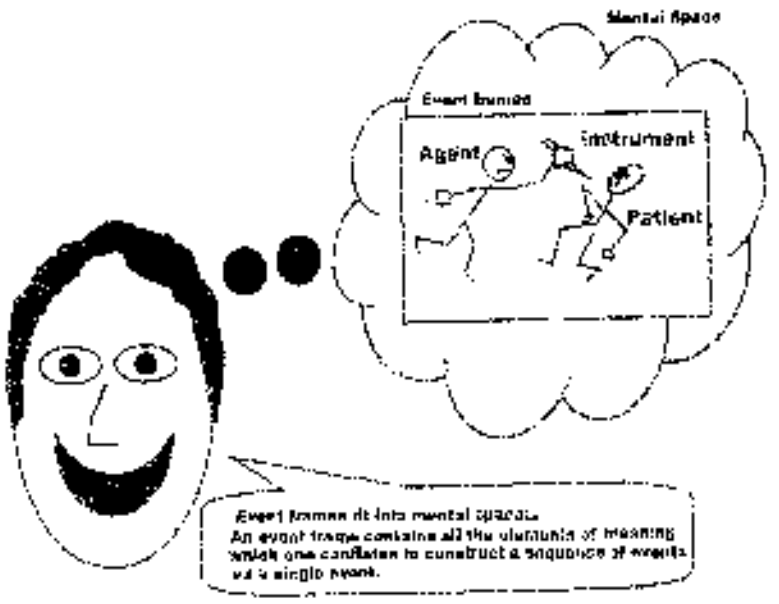


Figure 2: An event frame interpreting thematic roles within a mental space.

Within the event frame certain conceptual elements are on the foreground, while others are present in the background. The thematic roles that are allocated to entities when we conceptualise an event are prominent elements in the event frame. According to Langacker (1987a) the thematic roles of agent and patient are the archetypal, most basic or fundamental members of such a frame. Background elements in the event frame would be the time and locus of the event as well as the instruments used during the event – this view of event frames correlates with the figure-ground gestalt of cognitive psychology (Koffka 1935/6).

In syntactic terms the foreground elements of the event frame constitutes the essential elements of the sentence such as subject, transitive verb and direct object ([$_{subject}$John] [$_{verb}$hits] [$_{object}$the ball]). Other elements may be prepositional phrases which indicate instruments, place, direction and time: (John hits the ball [$_{instrument}$with a bat] [$_{place}$in the backyard] [$_{direction}$at a wall] [$_{time}$every morning/before school]).

Time figures in the background as a fundamental mechanism to organize the sequence of events within a time frame. Because it is present in the background, it is an abstract and subconscious mechanism which organizes time, state, change, event, cause, modality, goal and method metaphorically by structuring time as space. Time, metaphorically conceived as a container in which past, present and future are held in still smaller containers, implies that we can move backwards and forwards in time. This movement in time is described by means of syntactic patterns and morphological markers which manifest in language.

Every movement in time involves a complex operation over a variety of frames. Each operation is conceptualised as a cognitive event frame which includes an internal, causal and modal structure.

- An internal structure: the type of event which expresses the grammatical aspect of the event.

- A causal structure which, according to Goldberg (1995), can be understood as conceptual schemas such as the movement of a body in space, manipulation of objects and force dynamic systems, i.e. our experience of forces which we apply, forces which are applied to us and forces in our environment which influence each other (Sweetser 1990).

- A neural structure which relates to speech acts so that forms which function as pragmatic-illocutionary and event-like can be referred to our experience of bodily possibility and necessity (Sweetser 1990). In terms hereof the different meanings of modal verbs are connected metaphorically so that the physical becomes a metaphor for the non-physical (i.e. mental patterns and social conventions of experience).

According to Sweetser's cognitive projection of physical movement and manipulation of objects rests on the conceptual metaphor THE BRAIN IS A BODY MOVING THROUGH SPACE. This concept which functions as a generic space (Turner 1996:54), is often used to illustrate cognitive events, but basically it is Runyon's *The Brighter Prospect* which completes 'the mental journey to eternity' to a physical journey. The mental state is a locality, and a change from one state to another one is a change in spatial locality.

The manipulation of objects implies that there must be an actor who performs the actions which brings modality into play: the possibility, necessity or obligation to perform a certain action. The actor can act as agent by performing an action which influences a second party, the patient. In some way 'to break through a barrier' can be construed as a physical barrier being broken down or as a thought process during which new insights are obtained. Movement and manipulation can therefore be regarded as a natural combination with which we conceptualize a situation or a state.

Stories in Time: Analogy and Metaphor

Turner (1996:154–158) refers to narrative structure, or stories, as a thought process which coordinates a number of grammatical structures on a number of levels. Stories and grammar have the same structure because grammar developed from stories by means of analogy. Time as a grammatical system develops from narrative structures such as 'she bakes a cake every day', with the grammatical constructions Verb, Present Tense, Transference and Iterative Aspect, elements that arise from the projection of basic categories for stories, namely 'events' and 'action'.

Lakoff (1990:39) agrees that there is an inherent relation between language, metaphor and image schemas, and says that because image schematic mental patterns are transferred to abstract mental patterns by means of metaphoric mapping, certain abstract mental patterns are metaphoric examples of image schematic thought expressed in language.

A metaphor is not only used as a literary figure of speech, but is common in all forms of language use – in spoken as well written language in all genres whether fiction, historiography, scientific formulation or legal discourse (Klopper 1997d). The word 'metaphor' originates from the Greek verb *meta-feridn*, which means 'to transfer', for metaphor thus links two domains of experience with one another, or two subdomains of the same domain of experience (Botha 1996). The distribution of metaphor in everyday language was pointed out for the first time by Lakoff and Johnson in *Metaphors We Live By*, and since then research in the field of human cognition has confirmed the central role of metaphor in our conceptual systems – our perception of things around us, how we handle the physical world and especially our interaction with people in general. This basic, concept structuring metaphors became known as conceptual metaphors.

Metaphor establishes connections between language domains: our knowledge is organized associatively in domains of meaning, in subdomains and subsubdomains. A concrete domain of meaning can be entirely associated metaphorically with an abstract domain to make the abstract domain more understandable. Hence Turner's (1991:76) metaphoric description of time as movement along a route: the present tense is metaphorically experienced as a point on the route where 'we' are situated, the past is metaphorical by that section of the route which lies behind us and the future metaphorically the section in front of us. The point which corresponds with the present tense moves forward in a linear fashion—linearity therefore is an intrinsic part of time and a period of time a metaphoric line segment.

Change in time is understood by projection of physical action onto a non physical state or situation so that we can experience time as a mover or a moving object, as shown in Figure 3. (One can perform an action as a healer by healing a wound, it can move forward into the future or back into past. A story, which implies change in time, carries meaning by integrating at least two mental spaces, a source space and a target space, which can be illustrated by comparing a mental journey with a physical journey. The physical journey is contained the generic space. A journeyer is taking a single journey towards a destination which is not specified; no direction is given as to date, nor the internal form of the journey. The degree of inspecificity allows the generic structure to be projected equally well onto the input spaces:

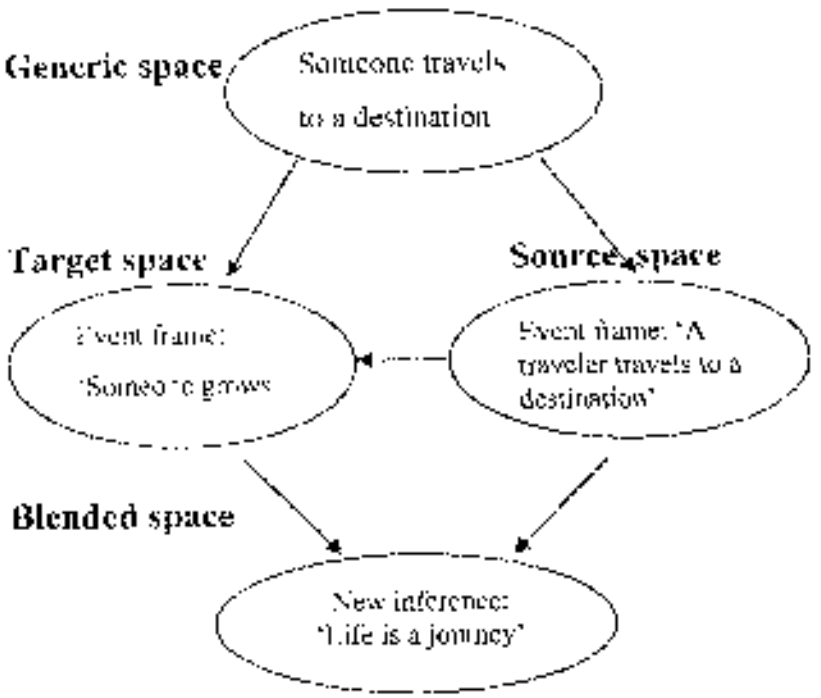


Figure 3: Integration of a source space and a target space into a blended space.

The source and target spaces each tell their own story on the basis of information contained in the event frames. The story in the source space is '*a traveler travels to a destination*'. The story in the target space is '*someone grows older*'. In the third space, the generic space, the stories in the source and target spaces share the schematic concepts '*a mover moves along a route to a destination*'. The shared concepts in the generic space make it possible to project the event frame elements in the source space onto the event frame elements in the target space and based on this association, to make a new

inference. By projecting '*a traveler travels*' in the source space onto '*someone grows older*', and by projecting '*someone arrives at a destination*' in the source space onto '*someone eventually dies*' in the target space, the conclusion '*life is a journey*' is reached in the blended space.

From a topographical point of view, the source space information structures the information in the target space via the metaphor 'life is a journey'. In the case of this metaphor the details of death or dying in the target space block or inhibit the conceptual projection in the blended space, where there is no mention of a destination being reached, or of someone dying.

In summary one can say that the shared information in the source and target spaces generically integrates these two frames during conceptual blending on the basis of shared semantic information and that the information in the source space topographically structures the information in the target space, as set out in the diagram.

Through conceptual blending time as abstract, non-sensory concept is made understandable by projecting more concrete space-like conceptualizations onto it. The container image schema forms the generic basis for the conceptual metaphor TIME IS A BOUNDED SPACE and her two daughter metaphors A SHORT PERIOD OF TIME IS THE OUTER SURFACE OF A SMALL BOUNDED SPACE and A PERIOD OF TIME IS A TIME-FILLED INNER SPACE, which Klepper (1998) illustrates by referring to a short period of time like a *morning* or a *day* which is metaphorically construed as the outer surface of a small bounded space, so that one can travel to Johannesburg *on a certain day*, or *on a morning [when transport will be available]*. Should one construe *a day* metaphorically as a time-filled inner space, one can travel to Johannesburg *in a day* or *in a morning*. On the basis of the same metaphor a good athlete can run the 100 meters *in the wink of an eye* or *within 10 seconds*, or a speculator can become a millionaire on the stock exchange *in three months* or *within three months*.

Time as a Deictic Category

Time as a deictic category and temporal reference point is anchored to a deictic centre, which represents the here and now of the actual time of utterance—this deictic centre is always present, at times in a highly abstract capacity, and functions as a basis space from where access is granted to time differentiation. As there is no such thing as a sentence without a speaker

(Langacker 1991; Cutrer 1994), each sentence has a conceptualizer which is always present in discourse, but which is not always marked grammatically or lexically. In written discourse a basis space may be created for a speaker, hearer, author, fictorial character or narrator, so that the conceptualized perspective will be the initial anchor point for access to and interpretation of time. The speaker uncouples the deictic centre from the here and now and moves to another locality which is specifically identified with the here and now of the reported event.

It is 1652. Jan van Riebeeck arrives at the Cape of Good Hope.

In a real world situation the present tense indicates that the event is aligned with the time of utterance. With the historical present each event can be interpreted as if aligned with the time of utterance so that an imagined replay of events occurs (Langacker 1991), or an 'eye witness' version—by using the present tense, the speaker/author describes the event as if he is 'seeing' it happen (Turner 1996:170-172). The consciousness 'sees' ideas or concepts or observes it from a certain locality and from a certain point of view. The consciousness may then 'move closer' to the concept or 'move away' from it, it can 'move' and idea or 'hunt it', and so on.

In accordance with the above, a mental space for the present tense is viewed as the primary domain of experience, and the cognitive point of departure from where one projects into the past or the future. During discourse the speaker locates himself in the here and now by opening a cognitive mental space for the present tense, so real deictic anchoring takes place. From this mental space he projects metaphorically to a past tense mental space, or to a future tense mental space. A specific mental space contains a sequence of events which is marked for the same temporal distinction: as soon as a shift in time occurs, a new mental space is opened for the new time frame. In this time frame events occur chronologically as a series of causal related events, and can be viewed as a completed event, an anticipated event, a conditional, or as historical present. During projection to another mental space the deictic centre shifts to the time and locality of the event frame contained within the mental space.

By applying the same principles used in actual discourse to fictional narrative discourse, the same projections are made for a story. The story is told by means of association mental spaces are opened, and within the mental spaces event frames are constructed that consist of sequences of events which

organize the chronological sequence of causal related events in the form of time frames. As in everyday discourse, the present tense within a narrative is regarded as the primary time frame for a cognitive point of departure from where projections are made into the past or the future.

Summary

The viewpoints put forward in this article are presented within the framework of Cognitive Rhetoric (Lakoff 1987; Johnson 1987; Langacker 1990; Turner 1991; 1996), which links human concepts with the theory of image schemas to account for the symbolic nature of human thought. Within this framework the mental spaces approach is shown to act as a background-organizing mechanism for the analysis and description of tense in everyday discourse as well as in fictional narrative texts. The interpretation of both types of discourse rests on the construction of a complex configuration of hierarchically related mental spaces—as each sentence in the discourse is processed, the configuration of spaces is adapted, based on lexical and grammatical triggers in the sentence (cf. Spruyt 2000a; 2000b). By applying the principles, which have been developed for time differentiation in actual language use, the same projections are also made for a story.

Time as an abstract, non-sensory concept is understood by the projection thereon of more concrete spatial concepts, so that time can metaphorically be construed as space. Time is thus viewed as a container in which past, present and future are locked up in still smaller containers as cognitive event frames, each with its own internal, causal and modal structures.

Department of Communication Science
University of Zululand

References

Botha, WJ 1990. Hoe beteken taal? 'n Kognitiewe perspektief. In *Suid-Afrikaanse Tydskrif vir Taalkunde Supplement* 28:5-25.

Cutrer, LM 1994. *Time and Tense in Narrative and Everyday Language*. Unpublished doctoral thesis. University of California, San Diego.

Edelman, G 1989. *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books, Inc.

Fauconnier, C. 1985. Generalized Union. In *Communication and Cognition* 165-37.

Fauconnier, G. 1985. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge, Mass.: MIT Press.

Fillmore, C. 1968. The Case for Case. In Bach, E. & R. Harms (eds): *Universals in Linguistic Theory*. New York: Holt.

Givon, T. 1979. *On Understanding Grammar*. New York: Academica Press.

Givon, T. 1982. Evidentiality and Epistemic Space. In *Studies in Language* 6:1:23-49.

Goldberg, Adele 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

Johnson, M. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: Chicago University Press.

Klopper, RM. 1997a. *Embodied Consciousness*. Unpublished manuscript. University of Zululand, Durban-Umlazi Campus.

Klopper, RM. 1997b. *Die Figuur-en-grond-gestalt en die Afrikaanse Werkwoorde van Beweging*. Ongepubliseerde manuskrip. Universiteit van Zululand, Durban-Umlazi Kampus.

Klopper, RM. 1997c. *The Demise of Certainty in Interpretation*. Unpublished manuscript. University of Zululand, Durban-Umlazi Campus.

Klopper, R.M. 1999. Teruglag op Vertrekpunt. In Gouws, Ruins en Lise Fernaner Steyn-Hellens: *Sintuiters op die Voorgrond*. Pretoria: Van Schaik Akademies.

Lakoff, George & Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, G. 1987. *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.

Lakoff, G. 1990. The Invariance Hypothesis: Is Abstract Reason Based on Image-schemas? In: *Cognitive Linguistics* 1,1:39-74.

Langacker, RW. 1987a. *Foundations of Cognitive Grammar: Volume I*. Stanford, California: Stanford University Press.

Langacker, RW. 1987b. *Grammatical Ramifications of the Setting/Participant Distinction*. BLS 13:383-394.

Langacker, RW. 1988. An Overview of Cognitive Grammar. In Rudzka-Ostyn, B. (ed): *Topics in Cognitive Linguistics*. 127-161.

Langacker, Ronald W. 1990. *Concept, Image and Symbol: The Cognitive Basis of Grammar*. New York: Mouton de Gruyter.

Rosch, Eli. 1973. Natural Categories. In *Cognitive Psychology* 4:328-350.

Rosch, Eli. 1975. Cognitive Reference Points. In *Cognitive Psychology* 7:532-547.

Rosch, Eli. 1978. Principles of Categorization. In Lloyd, B. & E. Rosch (eds): *Cognition and Categorization*. Hillsdale, New York: Lawrence Erlbaum.

Spruyt, MM. 2000. 'n Kognitiewe Benadering tot Tydsbeskrywing in Diskoers. *Southern African Linguistics and Applied Language Studies* 18:31-43.

Spruyt, MM. 2001. Kognitiewe Raamteorie: 'n Onderbeginraamwerk vir Tydsbelewing in Diskoers. In *Southern African Linguistics and Applied Language Studies* 19:81-93.

Sweetser, Eve. 1990. *From Etymology to Pragmatics: The Mind as Body Metaphor in Semantic Structure and Semantic Change*. Cambridge: Cambridge University Press.

Talmy, L. 1988. Force Dynamics in Language and Cognition. *Cognitive Science* 12: 49-100.

Turner, Mark. 1991. *Reading Minds: The Study of English in the Age of Cognitive Science*. Princeton: New Jersey: Princeton University Press.

Turner, Mark. 1996. *The Literary Mind*. New York: Oxford University Press.

Dual Coding and Conjoint Retention: Past, Present, and Future

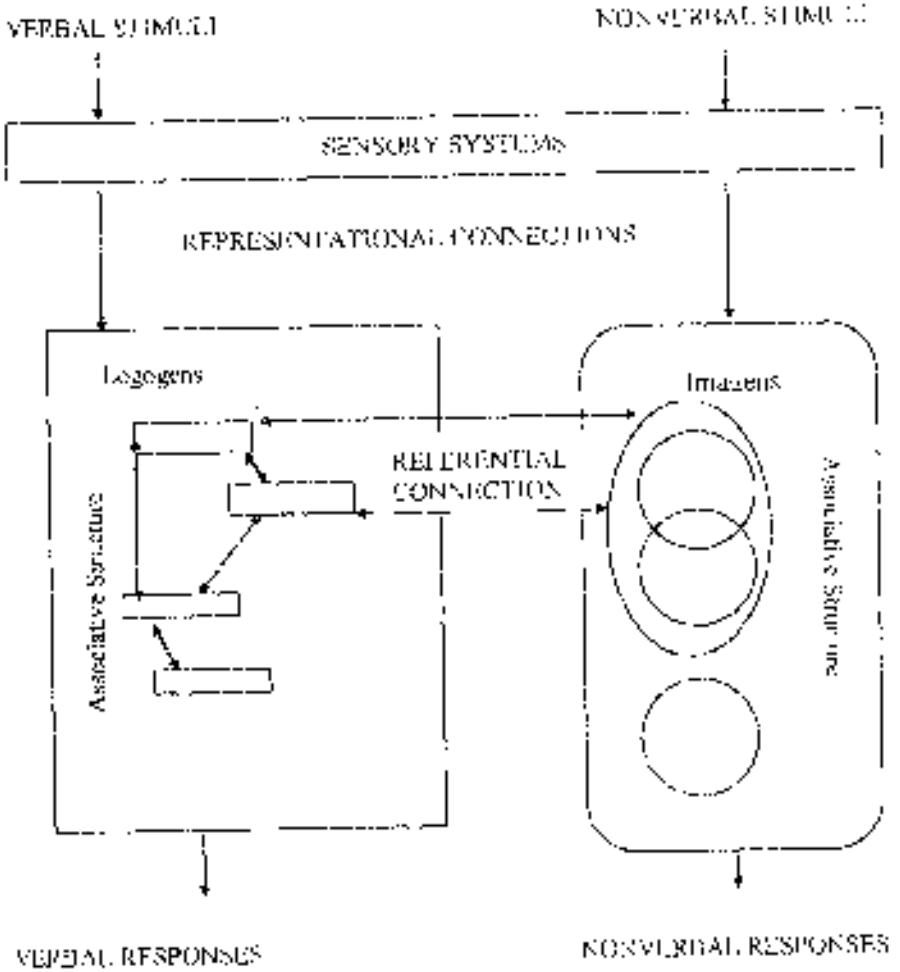
James B. Schreiber and
Michael Verdi

Introduction

Educators and psychologists have researched theoretical models of text or image representation and recall for decades (Anderson 1976; Kulasky 1980, McClelland, Rumelhart & PDP Research Group 1986; Paivio 1971). In the early 1980's, Raymond W. Kulhavy and his colleagues (Dean & Kulhavy 1981; Kulhavy, Lee & Caterino 1985) embarked on a research agenda concerning whether the use of organized spatial displays in conjunction with textual information increases recall. The purpose of this paper is to review the theoretical framework of Kulhavy's Conjoint Retention hypothesis (CR) and provide a review of the research based on this hypothesis over the past two decades. Finally, a discussion of current and future directions of the research is provided.

Conjoint Retention Hypothesis: Extending Paivio's Dual Coding Theory

Paivio's (1971, 1986) Dual Coding model of knowledge representation in long-term memory (LTM) is based on the idea that information is stored in two distinct codes, the verbal code and the nonverbal code (Figure 1). Many models of cognition include both verbal and nonverbal representations (Anderson 1983; Kosslyn 1986; Simon 1989). In Paivio's model, the two codes are functionally separate and distinct with one representing verbal information (propositions) and the other representing spatial information (images). The verbal code is concerned with the processing and representation of verbal information and the nonverbal code is concerned with the generation and representation of images in memory.



Adapted from Paivio (1986)

Figure 1

When stimuli enter the sensory system, Paivio contends that verbal stimuli in the form of propositions are stored in the verbal code and visual stimuli are stored in the nonverbal code. For example, when a person forms the image of the word dog, the word creates both a visual and verbal representation. The

two representations improve the chance of the word being recalled later, because there are two representations to assist retrieval. This memory improvement that results when a visual image and a word are learned together and is an instance of 'two codes are better than one' (Paivio 1986). Further, the codes may be activated in a singular or parallel operation and referential connections exist between the two codes.

Within each code, associative links relate information together. For example, questions about my desk evoke the image of my desk and all of the immediate surroundings in a probabilistic fashion. I can also survey my desk without the continuity being broken (Paivio 1986). More importantly, associative and functionally information that is stored in the verbal code can be linked to related information in the non-verbal code and vice versa. These referential connections (between code connections) make it possible for information in one code to aid activation and retrieval of information in the other code. Finally, both codes can have 'additive effects' on recall (Paivio 1986:57). In addition, referential cross code connections provide additional retrieval cues which increase learning (Kulhavy & Stock 1996; Verdi *et al.* 1997).

Kulhavy and his colleagues (Kulhavy & Stock 1996; Kulhavy, Stock, Verdi, Rittschoff & Savenye 1993) have extended Paivio's dual code theory to explain the acquisition of text material using organized spatial displays, specifically maps. Kulhavy, Lee and Caterino (1985) state that the conjoint retention hypothesis is an expansion of the dual coding approach in the sense that linguistic/verbal and perceptual/spatial representations in the cognitive milieu are treated as separable coding processes, at least portions of which are 'nonoverlapping' in memory. They believe that much of discourse content is stored in a linguistic/verbal mode with few perceptual referents (Anderson 1978), whereas, the geography of a map is stored primarily as a perceptual/spatial representation which possesses pictorial characteristics (Kosslyn & Pomerantz 1977). When we view an image (e.g., map) we encode the components of the image (features and their location). If we then read textual material about the image, we form cognitive links (based on the referential connections) between the information in the image, which is stored in the non-verbal code, and the matching information in the text, which is stored in the verbal code. These cognitive links facilitate the retention of information. Kulhavy and colleagues argue that due to these links formed between the two codes, verbal and non-verbal information can be utilized independently or in conjunction with one another. Then at recall, the person is

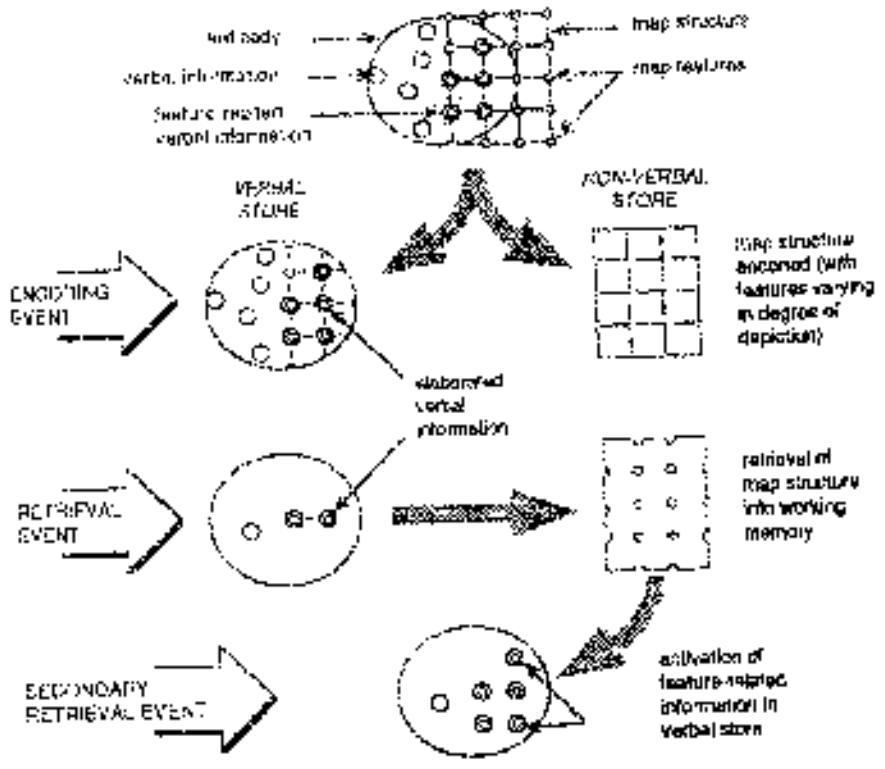
able to draw on both types of conjointly retained cognitive representations, thereby increasing the probability of remembering elements from each. The dual pictorial/linguistic representation provides a richer cuing base and retrieval base for the learner to draw from during recall. Finally, they believe that the possession of the conjoint representations may work to reduce the amount of interpassage interference at recall, by acting in some fashion to modify the functional characteristics of the incoming stimuli, and consequently increasing their availability in memory (Kulhavy *et al.* 1985).

A second extension of the dual code theory states that non-verbal information (images) has two characteristics that contribute to the improvement of memory for textual information. These characteristics can be generically labeled feature and structural information (Kulhavy, Stock, Verdi, Rittschoff & Savenye 1993). In the specific case of a map, the feature information may be landmarks, objects, paths, verbal labels, and other point designations. Feature information also includes what Bertin (1983) calls the 'retinal variables' such as size, shape, and color, which are used to further delineate individual features. Features may be depicted by different types of marks, including icons, boxes, etc. and Honma and Viera (1988) observed that memory for these individual entities is fairly accurate. Finally, Schwartz (1997) argues that it is the feature information found in maps that activate prior knowledge and allows the learner to process and interpret the map. Kulhavy and colleagues define structural information as the spatial characteristics across an image space. Structural information deals with the metric and spatial relations found between the individual features on a map including direction, borders that can be used as reference points for learners, and distance. It is the structural information that allows the learner to create intact images. For a neurological explanation of this phenomenon, see Tippett (1992).

Images within Kulhavy's model enjoy a privileged status when compared to propositions. Propositions are processed and stored serially (Van Dijk & Kintsch 1983), whereas images are processed and stored as a whole unit (Reynolds 1968). Because images are stored in a whole or intact form, the information contained within the image is simultaneously available, and therefore, it is possible to switch attention from one feature to another feature within the intact image without using all the resources available in working memory (Larkin & Simon 1987). Propositions do not share this advantage due to the fact they are stored and retrieved sequentially (Anderson 1983).

Therefore, images have a computational advantage over propositions especially in work environments where there is limited capacity (Miller 1956).

Figure 2 provides a graphical representation of the CR model. The figure shows the model in terms of encoding, retrieval and secondary retrieval of map text information. The top view of the figure shows the relationship between map and text and illustrates the overlap between feature related verbal information and the map features contained within their structural framework. At the encoding stage the two representations are incorporated into separate memory stores, with the feature-text associations denoted as dotted connecting units. At retrieval, both the verbal and image codes are pulled into working memory, and map information is used in the secondary retrieval of associated verbal words.



Also, based on Thorndyke and Stasz (1980), Kirasic and colleagues have described how people process maps based on their purposes or attitudes, encoding, and retrieval strategies (cognitive participation and sampling). Participation is used to retrieve and then to reduce or enlarge information. When an image is too large to learn all at once, participation is coordinated with spatial or conceptual categories. Sampling, the pattern of attention shifts between image elements, may be conducted in one of several ways. Systematic sampling shifts between processes using a decision rule criterion. Stochastic sampling shifts attention to adjacent elements. Random sampling involves haphazard shifts. Memory directed sampling involves shifting attention to elements not recalled during previous recall attempts.

Encoding processes serve to maintain information in working memory. These processes include rehearsal, grouping elements that share a specific property, mnemonics, and association or elaboration. Rehearsal is used to learn both spatial and verbal information while the other processes are used with verbal information only. Spatial information is encoded using four procedures: (a) visual imagery which forms a mental picture of the map or a portion of it, (b) labeling which generates a verbal cue in relation to a spatial configuration, (c) pattern or coding which specifies shapes, and (d) relation encoding which refers to spatial relation between map elements. In addition to these four procedures, schema application may be applied to either spatial or verbal information. Schema application involves associating map information with a prototypical configuration. Within encoding, prior knowledge becomes important. According to Schwartz (1997), 'Maps are not simply encoded into long term memory as images, verbatim'. Indeed, one of the things learners do when processing a map is to relate what they know about the area's geographic space to the stimulus map. In short, a person's prior knowledge about the area or subject depicted on a map plays a large role in how the learner processes the map. Lowe (1993) observed that those who had experience with meteorological maps had superior recall of both the number of markings and their accuracy when compared to a control group. Schwartz (personal communication) has recently observed that when the learner's prior knowledge about the area depicted was inconsistent with what was displayed on the map, map learning was dramatically reduced. Allen, Schwartz, Graham, Knight and McLaughlin (1998) found that when subjects have a prior knowledge of the geographic content discussed in the passage, the presence of a geographic map played a small part in the students' ability to recall the

passage. In short, when given a choice, students choose to study what is most familiar and therefore easier to learn.

> ...ation processes are inferred rather than observed, and direct skills between the other processes, including selecting and terminating which of the procedures are to be used. Thorndyke and Stasz (1980) identified differences in the procedures used by good and poor map learning. Good learners used systematic attentional procedures, a variety of encoding procedures, and frequently and accurately evaluated their performance. Poor learners did not use partitioning, and were limited in the variety of encoding procedures used with spatial information, and were less accurate than good learners when evaluating what they had learned.

Research Observations: What We Have Witnessed

This section provides observations based on research completed using the conjoint retention hypothesis developed by Kulhavy and his colleagues. There are seven themes across the research: Stimulus Characteristics, Edge Effect, Stimulus Order, Accuracy/Prediction, Thematic Maps, Individual Differences, and Task Demands. Within each theme, relevant research studies are provided.

Stimulus Characteristics

Several studies by Kulhavy and his colleagues have focused on manipulating the characteristics of the spatial display and to a lesser degree the textual material. Early studies focused on the iconic concreteness of map features and how that influences how maps are remembered. Kulhavy and Schwartz (1980) reasoned that the degree of iconicity of a feature (varying from a full miniature to a label only) would impact how the map was recalled. They argued that figural characteristics work to maintain relational context and this helps in the case of which a recall task demands performance that is spatially isomorphic with the original stimulus. In their study, only participants who were provided maps with miniatures were able to retain the spatial characteristics of the maps. This study helped provide evidence for the advantage of structural information of maps for learning related prose.

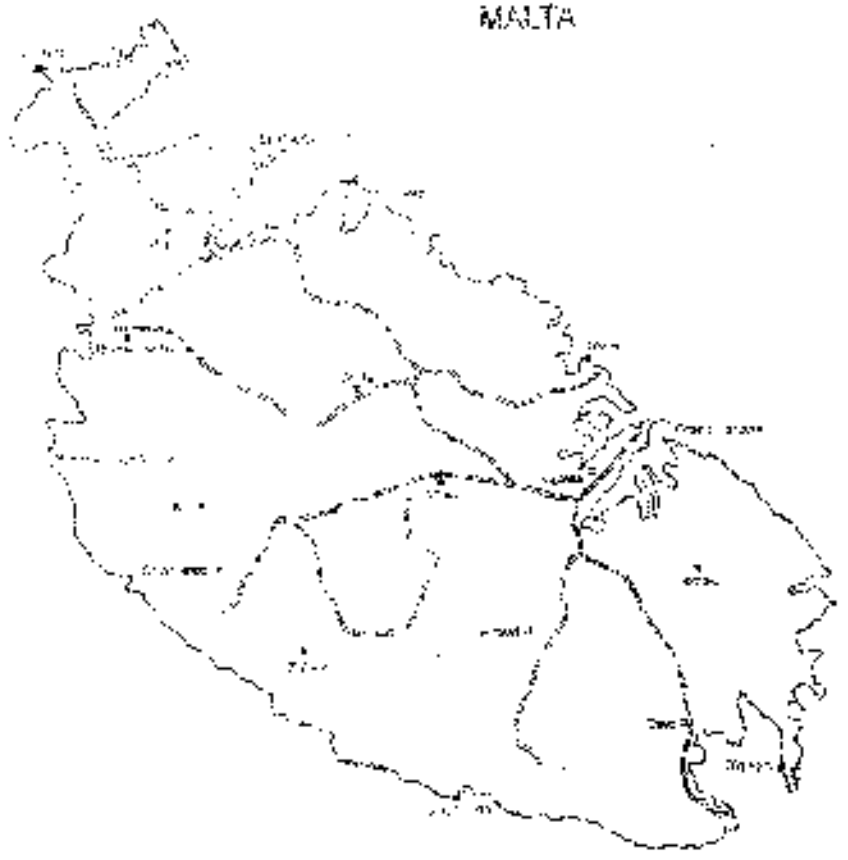
When students are presented with an intact map rather than a list map or segmented map they recall more information. Students who were given an intact map recalled more text information than those who were given a list map (the features were listed on the side of the map; the intact map had features distributed across the map surface) (Schwartz & Kulhavy 1981). Kulhavy, Stock, Verot, Rittschof and Sievenye (1993) used a map of Rome where subjects viewed either the intact map of Rome with twenty features or viewed twenty maps each with one feature and found that those who viewed the intact map had superior recall of information. More recently, subjects in one group viewed a map that contained the five landmasses and features of the Moluccas islands and another group viewed each island individually (Schreiber, Vordi, Patrick-Peckham, Johnson & Kealy in press). Those subjects who viewed the five islands intact recalled more text information than those who saw the islands individually. The key component of that study was encoding the map as a whole unit and processing it as a whole, thereby giving the intact group a computational advantage.

Text fact recall is enhanced when text information is directly related to map information. Subjects read a narrative that contained features which were located on the map and features that were not located on the map and then were asked to remember information on a cued and free recall test (Schwartz & Kulhavy 1981). For both cued and free recall tasks, subjects remembered more features that were located both on the map and in the text than features only located in the text, thus providing more evidence for the argument that 'two codes are better than one'.

Adding perspective to a map may enhance recall of features of the map. Traditionally, the research had focused on plan view (flat map) maps but in the mid1990's the research branched out into perspective maps. In cartographic terms a plan view represents three degrees of freedom in terms of symmetry: the plane is symmetrical about the X, Y, and Z axes. Tipping one edge of the plane, in this case the Z plane, results in a one-point perspective view. Therefore, a one-point perspective has two degrees of freedom in terms of symmetry: the Y and Z axes (Figures 3 and 4).

Subjects in one experiment viewed either a plan map or a one-point perspective map of the ancient city of Rome (Johnson, Vordi, Kealy, Stock & Haygood 1995). Those participants who viewed the perspective map recalled significantly more features and located more features correctly on the map reconstruction task than those who viewed the flat map. Johnson and colleagues argued that the increased dimensionality of the perspective map

provided a rich encoding base for the structural relations among features and rote memorization. Schreiber, Vordi, Parock-Peckham, Johnson and Kealy (in press) tried to replicate this observation by having participants view a flat, one point perspective, or two point perspective map of the island of Malta.



They observed that those who viewed the flat map recalled more information than those who viewed the perspective maps. The key for the disparity may actually be the clustering of items on the maps. On the Malta map, the features

appear to be more evenly spread across the entire area on the Rome map. Nine of twenty-three labelled features are located in the bottom quarter of the Rome map, which may have given the illusion of a richer encoding base.

With regard to textual information, early research was concerned with the difference between hearing or reading the text and recall. Wiel and Kulhavy (1986) observed no significant differences between participants who heard the text or read the text. A part of the study was replicated in the mid 1990's by Johnson, Schreiber, Vordi and Ayel (1996) with similar results. Simply, it appears that hearing or reading the text does not create a difference within the model.

Because of incredible work by B.J.F. Meyer and R.R. Armbruster and others concerning structural aspects of text material, the manipulation of the physical text has also been researched. Schreiber, Vordi, Parock-Peckham, Johnson, and Kealy (in press) had subjects read different text organizations after seeing a flat or perspective map (discussed above). No differences were observed between text types regardless of map type. They also had students read a randomly ordered text or an ordered text after viewing an intact or segmented map. By ordered, the authors meant that the text information would match the scanning pattern of the map. In western civilization it is recognized that literate individuals organize and read text from a top to bottom and left to right format. According to Winn (1991), items on a map would also be encoded in this fashion. Students who read a top-to-bottom text after viewing an intact map recalled more map and text information than those who viewed the same intact map but read a random text.

Edge Effect

The study of map perspective has enabled researchers to begin to examine what elements in the map are used to process and learn the map. This research has resulted in what is called in the literature, "The Edge Effect". The Edge Effect refers to the phenomenon that subjects learn more features found at the edge of a map than in the interior. Several studies have examined why and how this occurs. Rossano and Morrison (1996) ran a series of five experiments that examined how information is acquired from viewing a map. In a typical experiment, subjects were asked to study an experimental map for 10 minutes. At the end of the study period, subjects were given a blank piece of paper and asked to redraw the map they had just studied. One week later, subjects

remained and were asked to draw the map again. The results of the studies showed that border features, those close to the edge, were remembered more often and more accurately than interior features. In addition, the researchers found that subjects learned and remembered the maps in an orderly fashion, namely, features located along the border of the map were learned and remembered before features in the interior. One exception was found when interior lines were present. In these cases, these interior lines served as a border and these features were learned along with the edge features.

Jenston, Kulhavy, Stock, Stamm and Verdi (in submission) found that the stronger the edge of a spatial display (such as a map) is enhanced, the better these features located near those enhancements are learned. In addition, learning these features enables students to make appropriate connections with the facts associated with these features within a related text. Finally, Verdi, Stamm, Johnson and Jamison (2001, Experiment 2) had subjects view one of four experimental maps of Tasmania. Two levels of border (present vs. absent) were crossed with two levels of feature location (edge vs. interior) to form four experimental groups: border present with edge features only, border present with interior features only, no border present with edge features only, and no border present with interior features only. The results of the study showed that students were able to learn more features, accurately recall more facts from the text, and were better able to reconstruct the map when they viewed maps with features located along the edge than within the interior. There was little difference in those students viewing a map with a border and those viewing a map without a border. It should be noted that this border vs. no border finding is contradictory to most other findings in similar situations (e.g. Johnson 1996). This was due to the fact that the arrangement of features in this study allowed for students to create borders when they were not present, and, therefore, the missing border did not affect the subjects' ability to create images (a complete explanation of this phenomenon can be found in Verdi *et al.*, 2001). Recently, Verdi, Crook and White (in submission) have demonstrated that the 'edge effect' is also clearly visible when the map is presented in a computer format as well as traditional paper and pencil format. Schwartz (1997) asserted that it is the intactness of an image that enables students to use the image as a vehicle to aid in learning related text. The research presented here supports the position that it is the 'edge effect' that makes maps viable for forming images.

Stimulus Order

One of the most important observations of our research concerned concerns the order of the stimulus. Originally, studying both a map and a text was important to the recall of information. However, the computational advantage aspect of the conjoint retention model predicts that when structural information is encoded from the map it should produce an asymmetry in retrieval success, depending on which type of stimulus is first learned. This asymmetry suggests greater recall when the map is learned before the text. When students learn a map before a text, they encode both feature and structural information into an image. Because the map is represented as a structurally intact unit, it can be held in working memory without exceeding the capacity of the system. When reading a related text presented later, students bring the map image into working memory and use it to form associations with the facts contained in the text. These map-text associations can then be used to increase retrieval success during recall. Stock, Kulhavy, Webb, Pridemore and Verdi (1991), Verdi, Peterson, Webb, Stock and Kulhavy (1993), and Verdi, Johnson, Stock, Kulhavy and Ahern (1997) observed that participants who viewed a map before reading the text recalled significantly more information than those who read the text before viewing the map. Essentially, those who see the map first form an image that retains spatial information found in the map that is stored in long term memory as a single unit. When a subject views a text, it is encoded as a series of propositions (Van Dijk & Kintsch, 1983). Each time the student reads about a fact or feature in the text, an image of the map can be brought into working memory and associations between the map and text are formed. Those that read the text first process it as a series of propositions so when they are presented with the map a laborious process begins. They expend effort searching long-term memory for each individual proposition previously encoded and then bring them into working memory.

Accuracy/Prediction

Given the conjoint retention model predicts that structural characteristics of a person's map image will be closely tied to the number of text facts recalled correctly, there should be a technique to test this prediction. One way to test this is by the calculation of conditional probabilities. The resulting form is P(correct fact recall| accurate feature location) or P(fact|feature) (Table 1

provides mean values of P(fact|feature) for seven published experiments and one unpublished. The values range from .53 to .85 indicating a close relationship between location accuracy of map reconstruction features (an indicator of structural organization) and memory for associated textual facts. The more accurately a subject reconstructed information on the map the more information the subject recalled. In the Kulhavy, Woodard, Haygood and Webb (1993) and Schwartz and Kulhavy (1981) studies both probabilities were over .75 and specifically fact recall fell sharply when the feature was not located correctly or was absent from the map construction - an average of .37 as compared to an average of .80 when accurately located (cited in Kulhavy & Stock 1996). Ariel and Schreiber (1997) observed similar results with a probability of .71. Kulhavy, Stock, Woodard and Haygood (1993), had students learn a map of a city over three trials prior to reading the text. On each of the three map trials, students studied the map for one minute and then attempted to reconstruct it from memory. The number of times that a feature was correctly located on the reconstruction (0-3) was used as an index of the structural accuracy.

Table 1: Summary Table of Conditional Probabilities of Correctly Recalling a Text fact, Given the Associated Feature was Accurately Located on the Map Reconstruction.

Description of Map Used in Study	Probability of Joint Recall of Text Fact and Map Feature
Train Island with 14 features (Ariel & Kulhavy 1986)	0.67
Fictitious countryside with 16 features (Kulhavy, Stock, Caterino & Meschieri 1989)	0.79
Island of Tasmania with 15 features (Stock, Kulhavy, Peterson, Hancock & Verdi, 1995):	
Experiment 1	0.85
Experiment 2	0.72
City of Rome with 20 features (Kulhavy, Stock, Verdi, Rittschof & Savenye 1993)	0.80
Island of Hemare with 15 features (Kulhavy, Woodard, Haygood & Webb 1993)	0.73
Fictitious island with 16 features (Schwartz & Kulhavy 1981)	0.84
Island of Malta with 21 features (Ariel & Schreiber 1997)	0.71

The correlation between the structure index and associated fact recall was significantly different from zero with fact recall for high index features being 18 percent greater than for low order features. Two studies (Kulhavy, Stock, Werner-Bolhuis & Klein 1993, & Kulhavy, Stock, Woodard & Haygood 1993, second experiment) employed a different accuracy approach. Students were asked to learn a map of a city that had a compass rose at the center of the map. At recall, the students were provided with a series of blank sheets of paper, each with a compass rose at the center (in the same position as the one on the map). On each of these sheets subjects were asked to draw a line from the center of the compass rose to the exact location of a feature from the city map. The name of the feature was given as a clue. They were also asked to write down the facts about that map feature from the text after drawing the line. By using the angle and length of the constructed lines, they were able to compute a goodness-of-fit index using the polar coordinates formula for Euclidean distance:

$$\text{Goodness of fit} = \Sigma \frac{(\text{Student distance} - \text{Objective map distance})^2}{\text{Objective map distance}}$$

Average fact recall for subjects with higher goodness-of-fit values was 21 percent higher when compared to those with lower goodness-of fit values. Kulhavy and his colleagues argue that these results further support the notion that the structural characteristics of the map image facilitate retrieval of associated text information.

Thematic Maps

Most of the research completed used maps that contain landmark features (e.g. cities, banks, caves, palaces), but during the past decade thematic maps have also been incorporated into the research and the theoretical model (see Rittschof, Kulhavy, Stock & Hatcher 1993; Rittschof, Stock, Kulhavy, Verdi & Duran 1994). Rittschof *et al.* (1993) define thematic maps as a base reference map with one or more data themes distributed over the portions of the base map. With regard to the previous research and the conjoint retention model, such thematic maps are interesting because the thematic variables can be depicted in a variety of ways that vary in the extent that they depend on structural information. Therefore, thematic maps provide a way to examine the relative influence of structural and feature information on memory for related

text. Ritchie et al. (1993) also argue that themes add meaning to the map in the sense that they allow levels of one or more variables to occupy the map space. In a 1993 study Kitsantas and his colleagues had subjects view one of four maps (3 thematic, 1 non-thematic) of Ceylon (present day Sri Lanka) and read a text passage on information from the three themes. The thematic variables of agriculture, exports, and principal ports depicted using bivariate shaded areas that varied in their level of gray tone. The study demonstrated, in all cases, that persons viewing a particular theme recalled more text facts related to that theme and that the text's distribution worked in a similar fashion as point designations had done in previous maps-text research. Ritchie and Kulhavy (1998) found that college students could reconstruct similar amounts of map information from any of the four types of maps included in the study. More importantly, Ritchie and Kulhavy found that text recall from a fact-based text was facilitated by the themes presented on the thematic maps. It should be noted that in this case the authors used actual facts but placed them randomly with a given location.

Individual Differences

Individual differences have also been examined within the map-text research. Specifically, differences between males and females and cognitive styles have been examined. Schwartz and Philippe (1993) examined cognitive differences using the Group Embedded Figures Test (Witkin, Raskin & Witkin 1971) and recording whether the subject was male or female prior to the experiment. They observed that on map recall and map reconstruction, field-independent subjects outperformed field-dependent subjects. Finally, females remembered more of the map's features than males when asked to list them verbally, and outperformed males on map reconstruction. Kitsantas, Griffin and Ousler (1998) also conducted a study involving cognitive styles. They had students study one of four types of maps and then read a related text. The maps (a choropleth-shaded map, a proportional symbol map, a numerical data map, and a control chart) depicted book reading in Southwestern Oklahoma. The text that students read was related to the displays. Following the study of the materials, students were asked to recall what they could from the text and then to answer inferential questions concerning the map. The results of the study were quite robust and interesting. Students viewing one of the three maps as opposed to the control chart recalled 17% more county names and 45% more

relative information. Moreover, more related text information was recalled than unrelated text information.

Task Demands

The instructions that the subjects are given about the task also influence recall. Text fact recall is enhanced when students are required to process map information. Dean and Kulhavy (1981) observed, in their first experiment that those who were required to create a map based on the text material presented, recalled significantly more information than those who did not. In a second experiment they observed that when a person is forced to encode a spatial organizer prior to reading are more likely to retain the material studied. When Schwartz and Philippe (1993) instructed subjects to learn map features by clustering them together, either by their semantic or spatial relationship, subjects recalled the information accordingly.

Present and Future Directions

Currently, there are five main topical areas of research in the map-text tradition. The first one study is the examination of eye scanning patterns of subjects while they view a map. One study is in preparation and will focus on the discrepancy between the perspective map studies discussed earlier. A second study which examines eye scanning patterns for subjects who see a map first versus those who see a text first will also be completed. There also has been a lack of research in other domains (e.g. biology). Some research has been conducted with science (e.g. Verdi, Johnson, Stock, Kulhavy & Ahern 1997) and math, but most of it is unpublished at this time. We hope that future research will begin to examine other domains. A third area is the manipulation of the textual material in the tradition of text structure researchers. A study is presently underway that manipulates the structure of the text in order to see if it will influence recall. Next, a conditional probability model for recall of each map feature model is currently under development based on all of the previous observations and new considerations. Currently, the model stands as such.

$$P(C_f | M_f, E, L, C, T, R)$$

where,

* C_f is the Correct recall of feature or fact

* M_R is whether the map was viewed first or second
* E was the item near an edge
* L was the item located correctly on the map reconstruction
* C was the item a concrete feature or abstract
* T was the text organized or random, and
* R recognition level of the item (word frequency book).
* Finally, work with computer generated maps is just beginning.

Computers have the

* potential to create layers of maps that are easily accessible where paper maps have to be
* switched. Obviously, this is a potentially rich research area.

Summary

In this article we have reviewed how maps improve the recall of associated text and have presented a review of a theoretical model that we feel explains the processing of map/text information. Twenty plus years of research have provided a great deal of information about how people process and recall information from maps and text within the conjoint retention hypothesis. The findings showed that structural elements in maps, order of presentation in maps and texts, individual differences, accuracy and task demands all influence the recall of information and corroborate our model. Within the cognitive system, both the form of representation and the constraints imposed by resource limitation of working memory are highly influential in determining the level of information acquisition and recall. Yet there is still a pressing need for research aimed at helping geographers and others understand how various map configurations affect the way in which people use maps (Johnson, Verdi, Kealy, Stock & Haygood 1995). In this article, we have reviewed one way in which the content of maps can be assessed in terms of the cognitive relationship and the research observations over the past two decades. Research using the type of map-text stimuli reviewed here allows us to test how a pictorial display is likely to be represented and what sorts of tasks a person is able to perform using the representation. The information that is retained about the space of a map is dependent upon both the properties of the map itself and on the transformation selected by both the map maker and reader. It

is hoped that the past research, the research that is underway, and the research yet to be designed and completed will continue to increase our understanding of how organized spatial displays, such as maps, can be used to increase recall and retention of text material and how we process pictorial and verbal information.

Southern Illinois University Carbondale
California State University-San Bernardino

References

Abel, RR & RW Kulhavy 1986. Maps, Mode of Text Presentations and Children's Prose in Memory. *Contemporary Educational Psychology* 11: 33-48.

Allen, JR, NH Schwartz, L Graham, R Knight & MF McHaney-Dix 1995. Geographic Familiarity and the Retention of Text. Paper presented at the annual meeting of the National Consortium of Interactive Computing, San Diego, CA.

Anderson, JR 1976. *Language, Memory and Thought*. Hillsdale, NJ: Erlbaum.

Anderson, JR 1978. Arguments Concerning Representations Concerning Mental Imagery. *Psychological Review* 85: 249-277.

Anderson, JR 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Arici, AD & JB Schreiber 1997. Unpublished data.

Bertin, J 1983. *Semiology of Graphics*. Berg, William J (trans). Madison: University of Wisconsin Press.

Dean, RS & RW Kulhavy 1981. Influence of Spatial Organization in Prose Learning. *Journal of Educational Psychology* 73: 57-64.

Homa, D & C Viera 1988. Long-term Memory for Pictures under Conditions of Theoretically Related Foils. *Memory and Cognition* 16: 411-421.

Johnson, JT 1996. *Influence of Spatial Structure on Meaningful Text Based Information*. Doctoral Dissertation. Arizona State University.

Johnson, JT, RW Kulhavy, WA Stock, J Stamps & MP Verdi (in submission). The 'Edge Effect' and the Influence it may on the Retention of Meaningful Text-based Information.

Johnson, JT, JB Schreiber, MP Verdi & AD Arici 1996. Unpublished data.

Johnson, D., MP Verdi, WA Kealy, WA Stock & RC Haygood 1995. Map Perspective and the Learning of Text. *Contemporary Educational Psychology* 20(4):457-463.

Kausler RL 1980. *Human Memory* (2nd ed.). New York: Pegasus.

Kosslyn, SM 1980. *Image and Mind* (2nd ed.). Cambridge, Mss. Harvard University Press.

Kosslyn, SM & JR Pomerantz 1977. Imagery, Propositions, and the Form of Internal Representations. *Cognitive Psychology* 9:52-76.

Kulhavy, RW, JB Schreiber & LC Cochran 1986. Consider Retention of Maps and Related Discourse. *Contemporary Educational Psychology* 11:228-237.

Kulhavy, RW & NH Schwartz 1980. Mnemonics and the Spatial Context of a Map. *Bulletin of the Psychonomic Society* 15(6):416-418.

Kulhavy, RW & WA Stock 1996. How Cognitive Maps are Learned and Remembered. *Annals of the Association of American Geographers* 86(1):123-145.

Kulhavy, RW, WA Stock & WA Kealy 1993. How Geographic Maps Increase Recall of Instructional Text. *Educational Technology Research and Development* 41(4):47-62.

Kulhavy, RW, WA Stock, MP Verdi, KA Rittschof & WA Swenson 1993. Why Maps Improve Memory for Text: The Influence of Structural Information on Working Memory Operations. *European Journal of Cognitive Psychology* 5:75-92.

Kulhavy, RW, WA Stock, L Werner DeSanto & J Crook 1993. Properties of Maps that Improve Recall with Associated Text. *Cartography and Geographic Information Systems* 20:155-166.

Kulhavy, RW, WA Stock, KA Woodard & RC Haygood 1993. Comparing Elaboration and Dual Coding Theories: The Case of Maps and Texts. *Journal of Experimental Psychology* 19(4):433-448.

Kulhavy, RW, KA Woodard, RC Haygood & JM Webb 1993. Using Maps to Remember Text: An Experimental Analysis. *British Journal of Educational Psychology* 63:151-169.

Larkin, JH & HA Simon 1987. Why a Diagram is (Sometimes) Worth ten Thousand Words. *Cognitive Science* 11:65-99.

Lowe, RK 1993. Constructing a Mental Representation from an Abstract Technical Diagram. *Learning and Instruction* 3:157-179.

McClelland, JL, DE Rumelhart & the PDP Research Group 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition Volume II*. Cambridge, MA: Bradford Books.

Miller, GA 1956. The Magical Number Seven plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review* 63:81-97.

Orman, PK, F Raskin & HA Witkin 1971. *The Group Embedded Figures Test*. Palo Alto, CA: Consulting Psychologists Press.

Paivio, A 1971. *Imagery and Verbal Processing*. New York: Holt, Rinehart & Winston.

Paivio, A 1986. *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Reynolds, JH 1968. Cognitive Transfer in Verbal Learning: Transfer Effects after Pre-familiarization with Integrated versus Partially Integrated Verbal-perceptual Structures. *Journal of Educational Psychology* 59:153-158.

Rittschof, KA, MM Griffin & WI Custer 1998. Learner Differences Affecting Schemata for Thematic Maps. *International Journal of Instructional Media* 25,2:179-198.

Rittschof, KA & RW Kulhavy 1998. Learning and Remembering from Thematic Maps of Familiar Regions. *Educational Technology Research and Development* 46,1:1042-1629

Rittschof, KA & RW Kulhavy, WA Stock & JW Hatcher 1993. Thematic Maps and Text: Analysis of "What Happened There?" *Cartographica* 30:87-93.

Rittschof, KA, WA Stock, RW Kulhavy, MP Verdi & J Doran 1994. Thematic Maps Improve Memory for Facts and Inferences: A Test of the Stimulus Order Hypothesis. *Contemporary Educational Psychology* 19:129-142.

Rossano, MJ & TF Morrison 1996. Learning from Maps: General Processes and Map-structure Influences. *Cognition and Instruction* 14,1:109-137.

Schreiber, JB, M Verdi, J Patock-Peckham, JJ Johnson & W Kealy (in press). Differing Map Construction and Text Organization and their Effects on Retention. *Journal of Experimental Education*.

Schwartz, NH 1991. Human information processing of maps: A report to the geographic community. A paper presented at The First Assessment: Research in Geographic Education Conference. San Marcos, TX.

Schwartz, NH & RW Kulhavy 1982. Map Features and Recall of Discourse. *Contemporary Educational Psychology* 6:151-158.

Schwartz, NH & AE Phillippe 1991. Individual Differences in the Retention of Maps. *Contemporary Educational Psychology* 16:1-12.

Simon, HA 1989. *Models of Thought, Volume II*. New Haven: Yale University Press.

Stock, WA, RW Kulhavy, JM Webb, DR Pridemore & MP Verdi 1992. Map and Text Learning: The Privilege of Economical Representation. Paper presented at the annual meeting of the American Psychological Society. Chicago, Illinois.

Taylor, HA & B Tversky 1992. Spatial Mental Models derived from Survey and Route Descriptions. *Journal of Memory and Language* 31:261-292.

Thorndyke, PW & C Stasz 1980. Individual Differences in Procedures for Knowledge Acquisition from Maps. *Cognitive Psychology* 12:137-175.

Tippet, LJ 1992. The Generation of Visual Images: A Review of Neurological Research and Theory. *Psychological Bulletin* 112:415-432.

Van Dijk, TA & W Kintsch 1983. *Strategies of Discourse Comprehension*. New York: Academic Press.

Verdi, MP, SM Crook, D White (in submission). Learning Effect of Print and Digital Geographic Maps. *Journal of Research on Technology in Education*.

Verdi, MP, JT Johnson, WA Stock, RW Kulhavy & P Ahern 1997. Organized Spatial Display and Texts: Effects of Presentation Order and Display Type on Learning Outcomes. *Journal of Experimental Education* 64,4:303-317.

Verdi, MP, SK Peterson, JM Webb, WA Stock & RW Kulhavy 1993. Apm. Creating Maps from Perceptual and Verbal Stimuli: Retrieving Text and Reviewing Maps. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA.

Verdi, MP, J Stamen, JT Johnson & M Jamison 2001. Map Edges: Focal Points for Facilitating Text Recall. *Contemporary Educational Psychology* 26:211-226.

Winn, WD 1991. Learning from Maps and Diagrams. *Educational Psychological Review* 2: 211-247.

Consciousness Studies: Research Prospects in the 'Cradle of Human Consciousness'

Michael Mark Pitman

Introduction

Consciousness seems, at one and the same time, to be both the most obvious characteristic of human existence and yet also the most mysterious. If you were not now conscious, you would be asleep or in a coma, and certainly not aware of the meaning conveyed by the words on this page. Yet we live in a world dominated by a belief in physicalism – roughly, the doctrine that the ultimate constituents of the universe, and thus the 'building blocks' of all entities both conscious (us) and non-conscious (this page, your chair), are purely physical particles—the kinds of 'stuff' studied by physicists. From the point of view of physicalism, it seems thoroughly surprising that there are objects (creatures) in this world that experience qualitative states – states that *it feels like something* to be in. We have made great progress in understanding how certain organisations of matter can 'give rise to' what we call life. After all, we can describe the characteristics and behaviour of living systems in the unproblematically objective terms of a third person observer, or as Thomas Nagel has put it, from a 'view from nowhere'. Consciousness, however, seems essentially private and subjective. The only conscious experiences we ever have direct access to are our own, and there seems to be an unbridgeable gap between that first-person knowledge of consciousness and the prospect of ever examining a system or creature from the outside in a manner which will identify and adequately characterise any conscious states that it may have.

Consciousness as a subject of investigation has indeed had something of a torrid time within the interdisciplinary field of cognitive science. Cognitive science roughly straddles the disciplines of psychology, philosophy, linguistics, artificial intelligence and neuroscience (Thagard 1996). Each

discipline has, for various and interconnected reasons, historically neglected the field of consciousness studies until quite recently. In psychology, the most obvious explanation for this neglect lies in the dominance of behaviourism for much of the twentieth century. In philosophy, and specifically the philosophy of mind, the seemingly intractability of the 'puzzles' surrounding consciousness has seen decades of work on the so-called mind-body problem pay little attention to what, for most non-philosophers (and philosophers when they are off duty), is the most striking characteristic of all things mental. Artificial intelligence (AI) is also central to understanding the mind by modelling it (or, for some, replicating it) in machines that most would agree are not conscious at all. And neuroscience, with its 'outsider' view of the brain and its functions, has seemed equally determined to explain human behaviour with minimal reference to subjective 'inner' experiences of the brains under investigation. Indeed, some philosophers (e.g. Churchland & Churchland 1997) who are particularly keen on both AI and neuroscience seem determined to see conscious phenomena eliminated from any 'respectable' scientific view of the world.

Consciousness studies have, however, recently exploded onto the academic cognitive science. The reasons for this resurgence of interest are no doubt numerous, but they include the fact that respected, Nobel-prize winning scientists such as Francis Crick (1994) and Gerald Edelman (1989; 1992; Edelman & Tononi, 2000) have turned their expertise towards the study of consciousness, and that journals dedicated to the study of consciousness (*Journal of Consciousness Studies, Consciousness and Cognition, Psyche*) have made a significant impact on the academic scene, providing a place for discussion and publication of research for people of all disciplines interested in this most remarkable mental phenomenon.

It is not clear that the same can be said within the South African academic and research community. A casual analysis of articles published in 2000 in the *South African Journal of Psychology*, the *South African Journal of Philosophy* and *Philosophical Papers*, for example, reveals only one article falling in the field of consciousness studies out of 68 articles examined. This article (Gennaro 2000) is, further, the work of an American academic. Of course, these journals do not exhaust the interdisciplinary scope of consciousness studies, nor must local authors submit work in this area to these local journals. On the other hand, the result does suggest that consciousness studies represent a relatively small 'blip' on the local academic radar, at least within psychology and philosophy.

It could be speculated that this is a justifiable reflection of a certain theory-application orientation within academia more appropriate to the pressing social challenges and needs that a developing, African country like South Africa faces. However, if the anecdotal evidence is correct, it does suggest that there are opportunities for research and theorising that are gone amiss.

It was therefore with both excitement and sadness that I encountered a paper on consciousness by the late David Brooks (2000) entitled 'How to Solve the Hard Problem: A Predictable Incomputability' published posthumously in the e-journal, *Psyche*. My excitement was prompted by the interest and challenges that his paper raised. My sadness arose from the fact that I never met Brooks, and that he is no longer here to pursue what, as will be suggested below, is a promising attempt to grapple with the puzzle of consciousness.

The purpose of the remainder of this paper is to briefly outline Brooks' account of consciousness and provide a critical but constructive response to its central claims. The goal will be to delineate a project that is similar (in significant ways) to Brooks' own and which can be linked to other contemporary research within the field of consciousness studies, thus outlining a possible course for further contributions from the South African academy.

Evolutionary Predictability and For-the-organism Properties

Brooks (2000)[1] sets up his account of consciousness as an attempt to address what Chalmers (1995) has called the 'hard problem' of consciousness:

> It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does (Chalmers 1995:201).

This characterisation of the 'hard problem' highlights two features of debates over consciousness. First, there is a conviction amongst many contributors to the debate (Brooks 2000; Chalmers 1995; Searle 1992, 1997;

[1] All further references to Brooks refer to this article.

that the problem arises precisely because we are searching for an explanation of certain *facts* about consciousness, not for accounts of human behaviour that unsurprisingly *require* them. The existence and/or importance of conscious states (e.g. Churchland & Churchland 1997; Dennett 1991). Second, as Brooks notes, the so called 'hard problem' involves two kinds of questions about this 'extra' bit of 'why' questions about the need for qualitative states in intelligent physical creatures like ourselves, and a 'how' question about the way in which the these qualitative states in fact come into existence[7]. Brooks' account is concerned with the 'why' question, as is the balance of this paper.

The importance of the 'why' question can be brought into focus by considering the logical or conceptual possibility of Zombies – creatures physically and behaviourally identical to ourselves except for the crucial fact that they lack qualitative states. That is, there is nothing that it is like to be a zombie, and zombies are able to negotiate the world around them in much the same way as we do despite the absence of qualitative experiences. If the logical possibility of zombies is granted, then we appear to be faced with a choice: either physicalism is false, since we can have atom-for-atom identical creatures without identity in terms of their mental (especially conscious) states, or we accept that the qualitative aspect of mental states is epiphenomenal (causally redundant) since creatures very much like ourselves can achieve all that we can achieve in the absence of such qualitative phenomena.

Brooks' strategy is to grasp the nettle and argue against the logical possibility of zombies so as to preserve both physicalism and the place of qualitative states within the causal order of things. What is offered is an evolutionary engineering account of qualitative states (or qualia) committed to the following claim: any device engineered in accordance with the laws of physics which could perform all the tasks of which a human is capable would also *have to have* qualia. In short, zombies are not possible because qualia are *necessary* for the possibility of certain modes of interacting with and behaving in the world that characterise human cognition and behaviour.

One ingenious example offered by Brooks (2000) against the plausibility of zombies involves behaviour in which the role of the sensuous nature of our qualitative states seems incontrovertibly central to our

[7] For an extended debate and exploration of the 'hard problem', see contributions to Shear (1997).

explanations of such behaviour[8]. Consider John and his zombie twin, Zohn, each returning home from work to relax after a stressful day. Each makes a stop at their favourite Chinese takeaway on the way home to order a tangy Shanghai Steak supper (the prospect of another frozen convenience dinner not being commensurate with their need to relax). Before sitting down to scan mail, each chooses a CD (say Schubert's 'Death and the Maiden') and turns up the volume on their music system. Subtleties and variations aside, Brooks' contention is that whilst John's behaviour is clearly explicable in terms of his preferences between different sensuous, qualitative states (in case over hand TV dinner; Schubert over techno-rave), Zohn's behaviour is thoroughly *puzzling*. With no qualitative states to refer to in explaining Zohn's choices, we are left to marvel at why Zohn did not take the most quick and easy route to relaxation (straight home, quick TV dinner, flick on the radio) rather than go to all the effort that he did. In general then, zombies seem implausible as behavioural 'twins' of humans because we cannot imagine them behaving in ways that we do when the sensuous nature of our qualitative states – specifically, the character of certain sensuous states rather than others – plays a distinctive role in the explanation of our behaviour. In so far as we can imagine zombies, these would be creatures that go through life making choices like our own (what taste to savour, what music to enjoy, what work of art to place on the wall) for no apparent good reason.

This tactic of finding an irreducible causal role for qualitative states in the explanation of human behaviour is the core of Brooks' strategy. However, Brooks rightly wants to argue that qualitative states play an even more significant role in the causation and explanation of human behaviour than merely determining our more luxurious wants and preferences. Specifically, Brooks wants to argue that, from an engineering or design perspective consistent with the laws of physics, creatures with our *intellectual, cognitive* and *behavioural* capacities should be expected (or predicted) to *have* qualitative states like our own. We should not, that is, be too quick to assume that the task of dealing intelligently with the world (as we do) can be achieved in the absence of qualitative states. The way to solve the hard problem is to thus produce an account of the predictability of qualitative states in a world governed by the laws of physics.

[8] The example has been developed and embellished, but is based on Brooks (2000) original example.

The core of Brooks' account involves the idea that humans interact intelligently with the world on the basis of what have been called 'mental models' (see Johnson-Laird 1983; McGinn 1989). The idea is that we interact with the world intelligently on the basis of isomorphic representational models of objects, properties and relations in the world. Brooks' claim is that any system capable of building mental models of the world as sophisticated as ours, but with comparable discriminatory powers, will have to have, as a matter of what he calls *engineering necessity* (1) perceptual equipment similar to ours, such as edge enhancers, and (2) *quality spaces* distinguished by the distinctive qualitative states (the qualia of colour, sound, taste, etc.) associated with different sensory systems.

Philosophers (at least since Locke) have distinguished between the primary and secondary qualities of the objects of our experience. In the visual modality, shape is a primary quality—a perceived feature *of* an object that inheres objectively in the object independently of any perception thereof—whereas colour is a secondary quality—a perceived feature of an object that, although it correlates with objective features of the object (wavelengths of light reflected by the object's surface), is subjective in so far as (1) our colour categorisations depend on objects affecting human consciousness in some way (e.g. producing the experience of red) and (2) correlations between colours reflect peculiarities in our own perceptual equipment (which wavelengths we are sensitive to and can discriminate between) rather than differences in objects 'out there' in the world.

Brooks' first contention is that for systems like us to transform two-dimensional retinal images into (accurate) three-dimensional representational models conveying information about primary qualities such as spatial distance and relative position *requires* representations of objects in a quality space. (The 'space' of colour experience with a range of variations (different colour qualia, such as red, blue, yellow, shades of more rarefied grey to black) that can be mapped onto distance and positions of objects in the world by way of relative similarities and dissimilarities within that quality space. More simply, to model 3-D shape (and position) requires edge enhancers of some kind, and this is to be achieved by differences in quality (distinct colour qualia and shades thereof) within a quality space. Of course, Brooks does not need to argue that our particular colour qualia are necessary to achieve this edge-enhancing function (echolocation seems to serve bats pretty well when it comes to 3-D modelling), but only that *some* distinctive quality space with a requisite range of variations is necessary to model the world in 3-D as well as

we do (echolocation may be fine for navigation, and it works in the absence of light, but colour modelling perhaps enables modelling and behavioural repertoires that echolocation does not).

There is then an adaptive advantage to be gained by modelling important primary properties of the world around us by way of qualitative variations within an experienced quality space. Brooks' second contention is that there is further adaptive advantage to be gained by the association of distinctive quality spaces with different perceptual modalities—colour with sight, sound with hearing, proprioception with the olfactory sense, 'feel' with touch and proprioception. The advantage to be gained is simply that this allows for (a) easy discrimination between sources of perceptual input, and (b) easy direction of sensory and attentional resources to the sources of salient input (e.g. turning one's ear towards a suspicious sound rather than one's eyes, nose or hand).

Systems with edge enhancers and quality spaces do not, however, strike Brooks as sufficient to secure the necessity of qualitative states. He uses the example of the filming and broadcast of a television show at every stage of this process, from the light entering the camera lenses through the radio waves sent by the transmitting broadcaster to the reception and transformation of these waves into patterns of pixel activation on a receiving television screen, there is an isomorphic model of the original 'scene' that preserves or models all the requisite spatial relations. Yet the system is completely 'blind' to this accurate modelling. What is needed, argues Brooks, is to add the self-awareness that characterises our own perceptual and intellectual activities. For a system with edge enhancers and quality spaces to behave as we do requires awareness *within the system* of the representational medium *as being medium of representation* (and this, needless to say, the 'procedural system' lacks). That is, the system requires an awareness that certain properties of its internal states are properties of its own subjective perceptual states (rather than of the world itself) that act as labels, codes or tracking tags within a representational medium whose function is to model the world outside. These 'tags' are such that they are regularly associated with a distinctive class of external things, and such that the system is readily able to distinguish one tag from another. It is only when the system becomes aware of these tags *as being* tags that these qualitative features acquire all the features we associate with qualia. These properties of such systems Brooks calls 'for-the-organism' properties.

To clarify this argument, let us briefly consider Brooks' (2000) reference to research on image rotation (Kosslyn & Pomerantz 1981). This research has found that the response time of a subject asked to describe (or decide) how an object would look from a different angle or perspective *depends* on the angle through which the object would have to be rotated to see it from that point of view. This finding has been used to support the idea that the response times vary because the subjects in fact rotate a mental image or model of the object in producing their response. The research not only supports the 'mental modeling' approach mentioned earlier, but helps illustrate Brooks' point about awareness of tags. The research subject, it seems, is able to mentally rotate the requisite object 'image (in part) *because* they are aware of the different tagging of the object and its features and that of the background against which it is rotated. It is the awareness of these discriminating For-the-organism properties and their mapping onto the object in question that enables the intelligent response of the subject to the task. The subject, in interaction with the object, picks out qualia as the requisite For-the-organism properties, and the resultant manipulations/calculations are carried out in virtue of these For the organism properties.

Brooks, however, leaves the task of establishing the necessity of these For the organism properties as an open-ended strategy to be pursued rather than a thesis already established. Specifically, Brooks leaves us with the task of pin-pointing some intellectuo-perceptual capacity *C* such that (1) we (humans) have *C* and (2) the possession of *C* underpins our ability to report on and take pleasure in our qualitative states. Furthermore, any creature (or system) without qualitative states would lack *C*, thus establishing the necessity of qualitative states. That is, any device *D* engineered, by evolution or otherwise, in accordance with the laws of physics *must* have qualitative states if it is to have the capacity *C*. Identifying *C* will thus secure the place of qualitative states (and specifically, of the *qualitative aspect* of qualitative states) as having a distinctive causal and functional role within the cognitive and behavioural activities of creatures like ourselves—and secure this in a way entirely compatible with physicalism.

If this is an accurate characterisation of the project and challenge with which Brooks has left us, then it is to the critical assessment of this project that we must turn before charting, if possible, a way in which to take it forward.

The Limits of Physicalism, Science and Evolutionary Explanations

In critically evaluating Brooks' (2000) suggested strategy for dealing with the hard problem of consciousness, three areas of contention will be highlighted: (1) the place of physicalist assumptions in articulating the puzzle(s) of consciousness; (2) the usefulness and the limits of evolutionary explanations; and (3) the role of self-awareness with in Brooks' account.

As noted at the outset, the assumption of physicalism is widespread, and nowhere more so than within the philosophy of mind, where much (if not most) of the last three decades of work in this field has been centred on the project of 'naturalising the mind', in the sense of providing an account of mind that secures its respectability and place within a monistic, physical universe. Yet physicalism is an *assumption*—or rather a set of assumptions—that is as regularly left undefined or underspecified as it is assumed (Crane & Mellor 1991). There are those (of whom Crane & Mellor 1991; 1995) are a most notable example) who are inclined to argue that when the assumption of physicalism is unpacked, and specifically unpacked in such a way as to render the mental domain uniquely distinct, mysterious and/or problematic in some way, the plausibility of physicalism comes seriously under question (see also Proton 1998).

Although highly pertinent to the current debate, the substance of these arguments against the 'obviousness' of physicalism cannot be dealt with within the scope of this paper. What should be noted is that one need not be a committed physicalist to find consciousness puzzling. Consciousness is puzzling and mysterious (and fascinating) because it is *the* quintessentially subjective phenomenon in the universe that, for this reason, seems to resist description and explanation from 'the outside'. It is further puzzling because we have a tendency to think that, when neuroscience fills in the gaps in its understanding of the brain mechanisms associated with consciousness, there will be no causal or functional role left for the qualitative or phenomenal aspects of these conscious states: each brain state will follow smoothly on from the preceding one with no apparent need to refer to other causal factors, not least to qualia and the like. There is a strain of physicalism in this tendency, but it is not central to our present concerns.

One last point is that physicalism is a particular brand of what were referred to earlier as 'naturalisation projects' within the philosophy of mind. The guiding motivation behind such naturalisation projects is that if the mind,

or various aspects of the mind cannot be naturalised in some way (usually by way of providing at least sufficient non-mental conditions for the occurrence of mental phenomena), then some dire consequence will follow (Pitman 1998; Stich & Warfield 1994). The usual candidates for such dire consequences are that the mind is really illusory or unreal, or more often that the mind (or rather its features, such as consciousness) is epiphenomenal. This approach seems to me mistaken. The brain is the most complex object that we know of in the universe (Edelman & Tononi 2000), and to blithely assume that we will come to understand all its mysteries, and worse, that failure to do so would render some of its most remarkable characteristics (such as the capacity to 'produce' consciousness) epiphenomenal, seems both arrogant and defeatist.

The point of these observations is not to castigate Brooks or anyone else for their seeing a pressing need in pursuing physicalist or naturalising agendas. Rather, they are presented as a means to suggesting that philosophers and cognitive scientists should be more cautious in how they frame their puzzles and their criteria of success, especially when it comes to the mind. *Trying to find* necessary and/or sufficient conditions of a certain kind for consciousness is a worthwhile project, then, as long as we don't attach too high a cost to the possible failure of our endeavour. Physicalist naturalisation projects are worthy paths of inquiry, but they should not be viewed as the only game in town.

The fault, therefore, in Brooks' account that we might endeavour to avoid is his evident commitment to physicalism. Whether or not physicalists can sleep easy at night (Brooks' phrase) over the problem of consciousness, we can simply concern ourselves head on with the attempt to answer the hard question as to why (and, if possible, how) consciousness arose in creatures like ourselves. And in pursuit of this goal, we should be as willing to question, at any time, the background assumptions within which we are operating (e.g. physicalism) as we are to critically question the plausibility of any account of consciousness on offer.

The second point of contention concerns the limits of evolutionary explanations, both in general and with regard to consciousness in particular. Evolutionary explanations are, of course, perfectly suited and appropriate to answering 'why' questions about the origins and functions of features that we find in living creatures like ourselves. To assign a function to consciousness that both explains human behaviour and outlines the adaptive advantages conferred on creatures endowed with such functional capacities would indeed go a long way (if not all the way) to explaining why consciousness emerged in

our universe in creatures like ourselves. Brooks' positive account is an admirable instance of just such an attempted explanation, and provides considerable 'ammunition' to be used against those inclined to explain consciousness, or its causal efficacy, away.

However, evolutionary explanations are ill suited to the task of establishing the *necessity* of some or other function being achieved in a certain way (see Flanagan & Polger 1995; Polger & Flanagan 1996). The easiest way to see this is by considering the nature of adaptive advantage and the role of chance within evolution. Assigning and explaining function by way of adaptive advantage depends on the idea that characteristics emerging by descent-with-modification owing to influences of relative environmental fitness, survival and procreative success, from the point of view of evolution it matters not so much *how* a function is performed as simply that it does manifest in a breeding population. For this reason, evolutionary accounts of function cannot in general rule out the possibility of other means of performing that function. Evolutionary necessity, if there is such a thing, is thus a highly contingent affair tied to the peculiarities of particular species in particular environments with particular histories of descent. It is not well suited to the task of being used in arguments against the (logical or metaphysical) possibility of a certain function being carried out in different ways by different or even similar creatures that have, for example, radically different histories.

It might be countered that evolutionary arguments *are* at least suited to providing statistical arguments in favour of the necessity of a function being carried out in a certain way, given the time (and genetic and environmental diversity) 'available' for evolution to have come up with suitable alternatives. No doubt there is something to this point, but it brings us to the issue of chance in evolution. The difficulty of providing evolutionary explanations of anything derives not only from lack of the definitive evidence required to establish such explanations, but also because chance has such a significant possible impact on the course of evolutionary history. Many would be familiar with the common speculation that the 'age of the mammals' might never have come about had it not been for the catastrophic (meteoritic) extinction of the dinosaurs. But for that cosmic interference, the phylogenetic scale might look rather different today. Similarly, in relation to our case of the possibility of zombies, Polger and Flanagan (1996) ask us to imagine the scenario in which our ancestors inhabited the earth with zombie counterparts who were functionally identical. By way of a freak volcanic eruption, all those zombies

acts of narrated, I am to our own conscious forebears alone, with their now distinctive but only accidentally unique functional capabilities. Given such potential for chance interference in the otherwise 'neat' evolution of function by selective descent with modification, evolutionary explanations are again unsuited to producing accounts of the necessity of performing functions in particular ways.

Perhaps Dennett's (1996) claim is thus that evolution alone cannot put out the possibility of conscious zombiedom (i.e. the claim that consciousness is not essential to our human cognitive and behavioural capacities) that the case of zombies was designed to illustrate. To this extent, we might see as many of Brooks' attempts to establish the necessity of consciousness to the performance of certain intellectuo-perceptual operations that we humans are capable of. To ask for consciousness be necessary is to ask for more of our evolutionary account of human behaviour.

Does this mean that evolutionary explanations cannot be used to answer the hard 'why' question? Certainly not. First, it simply means that our sights should be set lower—we should attempt to establish that, as a matter of contingent fact, consciousness evolved in humans because it enabled certain intellectuo-perceptual capacities that conferred selective and adaptive advantages on our ancestors. That would surely be an answer to the 'why' question that is as informative as one could desire.

But there is a second point to be held in mind—namely that we might be able to establish (or at least argue for) the 'contingent necessity' of consciousness relative to certain functions of which we are capable as organisms. That is, whatever the logical or metaphysical necessity of consciousness to certain capacities we humans happen to have, we might still hope to eliminate the possibility of zombies in our world by arguing for the necessity of consciousness to certain intellectuo-perceptual capacities in creatures with a biological/physiological makeup like ours. Our ideal would thus be a scenario in which we can plausibly assert that creatures with brains like ours could not achieve what we achieve except in so far as such creatures have conscious, qualitative states. We might not attain this ideal of 'contingent necessity', but it is worth aiming for, and it still leaves us with the possibility of obtaining a 'purely' contingent evolutionary explanation that is still sufficient to our needs. Brooks' account thus essentially points us in the right direction, only with slightly too much confidence in what we will be able to achieve by following his strategy.

An example of this kind of approach to understanding the significance of consciousness can be found in what Baars (1997) calls 'contrastive phenomenology'. Focussing on humans (both normal and abnormal), the idea is that

> by contrasting similar conscious and unconscious mental representations in beliefs, perception, selective attention, imagery and the like, *we can gather a set of sound empirical constraints on the distinctive properties of consciousness as such* (Baars 1997:187, e.o.).

Baars' strategy is thus less concerned with establishing the necessity of consciousness for any given task or capacity as it is with better understanding the nature and role of consciousness in human performances of a range of tasks and operations. Because the analysis is contrastive in nature the outcomes of such investigations are supposed to provide a clear account of what it is that consciousness adds to our mental and behavioural capacities that is not there when we operate with similar capacities that do not involve consciousness. Clearly, a contrastive account of what consciousness *adds* to our mental and behavioural capabilities would go a long way, in evolutionary terms, to explaining *why* creatures like ourselves are conscious (for more on contrastive analyses and Baars' Global Workspace Theory see Baars 1996).

The third point of contention is more concerned with the details of Brooks' (2003) argument. As we have seen, Brooks introduces self-awareness at a seemingly crucial point in his argument where he introduces the notion of For-the-organism properties. It is the awareness of qualia as tags within a representational medium, or awareness of these tags as tags for the organism, that Brooks thinks is the key to such cognitive capacities as image rotation and, by generalisation, mental manipulations, operations and calculations generally. Whatever the virtues of these insights, one might legitimately raise the concern that Brooks has too quickly introduced a form of conscious awareness that is itself distinctively human[4]. Cognisant of the possibility of 'blind' systems equipped with edge-enhancers and quality spaces (e.g. the TV

[4] In Edelman's terms, as will be discussed, the kind of self-awareness that Brooks introduces requires an ability to explicitly recollect one's conscious mental scene, and this capacity only comes with the kind of higher-order consciousness found in humans and some of the higher apes.

broadcast). Brooks introduces a feature of self-awareness which itself remains unexplained within the account, and which leaves 'why' questions about the significance of phenomenal states in non-human conscious animals beyond the scope of the account. If the accusation is correct, then we need to seriously question whether Brooks has pointed to a way in which the hard problem can be resolved with sufficient generality to deal with these non-human varieties of consciousness.

The criticism highlights a concern that Brooks may have tried to achieve too much too quickly. Conscious states are characteristic of many creatures other than humans, and from the perspective of evolution, there are thus many potential consciousness function leaks that we could explore to establish the importance and predictability of qualitative conscious states before we arrive at the relatively sophisticated functions of the human mind. We can, that is, distinguish between primary consciousness (that we share with many other creatures) and higher-order consciousness (of which we share some aspects with other primates, and other aspects of which we seem to possess uniquely) (Edelman 1989; 1992; Edelman & Tononi 2000), which brings us to an examination of some contemporary research that holds promise for taking Brooks' strategy forward.

Primary and Higher-order Consciousness

That Brooks may have moved too quickly to a focus on human consciousness is most evident in the concern expressed in his account for being able to fully account for qualia as these are understood within the traditional philosophical debates over consciousness. Possible reasons for this error are not difficult to find. Edelman (1989), in the process of developing a theory of consciousness grounded in neuroscientific research rather than philosophical argumentation, nevertheless reflects on important philosophical issues raised in the study of consciousness. One of these issues is precisely that it is only humans who have qualia, not because other creatures do not possess primary consciousness and its attendant phenomenal states, but because they do not possess the requisite higher-order conscious capacities, and specifically linguistic capacities, to be able to *report* on these phenomenal states. In Edelman's (1989:169f) terms, a creature may have phenomenal experience,

> but it is not *subjective* experience—there is no subject or person to make discriminations or reports of that phenomenal experience over time. ... [W]e cannot call such states qualia in the sense that we can call our own states qualia[5].

In this sense, Brooks (2006) is correct to place emphasis on self-awareness when it comes to explaining qualia. But the argument stands, because to move too quickly to the focus on distinctively human consciousness is to gloss over other manifestations of conscious states with qualitative aspects that, from an evolutionary perspective, would presumably have preceded human consciousness, and indeed form the basis for human higher-order consciousness. It, therefore, the task is to generate an evolutionary account of qualitative states (even an account of qualia *per se*), it seems far more suitable to start by explaining the causal and functional role of qualitative phenomenal states first before moving into the more complex forms of consciousness of which humans are capable.

This is Edelman's (1989; 1992; Edelman & Tononi 2000) starting point. Edelman and Tononi (2000) describe primary consciousness as (the capacity to form a mental scene, and in most creatures (other than humans) this capacity is accompanied by limited semantic or symbolic capabilities and no true linguistic ability. Higher-order consciousness, they claim, involves capacities that flourish in humans and that *presuppose the existence of primary consciousness*. Higher order consciousness

> is accompanied by a sense of self and the ability in the waking state explicitly to construct past and future [mental] scenes. It requires, at a minimum, a semantic capability and, in its most developed form, a linguistic capability (Edelman & Tononi 2000:102).

The significance of this distinction to how the study of consciousness generally, and to an evaluation of the promise of Brooks' account, becomes evident when one examines Edelman's (1989) theory of the evolution of primary consciousness. Edelman (1989) claims that primary consciousness

[5] It should be noted that Edelman's use of 'subjective' is thus different to the sense of 'subjective' used elsewhere in this paper. Following Nagel (1974) the subjectivity of conscious experience refers to there being something that is it like to be a creature with those experiences. It does not imply, as it does for Edelman, a sense or experience of a self as subject of those experiences.

arose as a result of two developments in the evolution of brain systems: (1) the development of special memory repertoires composed of neuronal groups dedicated to storing past matchings of value and perceptual categorisation; and (2) the development of reentrant signalling between these special memory repertoires and neuronal groups currently devoted to sampling of the sensory environment for perceptual categorisation in all sensory modalities.

Three points of clarification are required before proceeding to Edelman's (1989) explanation of primary consciousness itself. First, 'value' in this context is probably best understood as 'salience for the organism', and is determined by (largely proprioceptive) homeostatic mechanisms in the brain concerned with the immediate needs for survival of the individual creature. Second, 'perceptual categorisation' in this context is taken to involve categorisation at the level of neuronal groups (i.e. patterns of neural activation in relevant sensory cortices), such that environmental stimuli will fall in the same perceptual category if they trigger the same (or a significantly related) neuronal group. Value-perceptual categorisation pairings are thus to be understood as pairings of environmental stimuli (via exteroceptive signals) with 'salience to the organism' values determined by homeostatic (interoceptive) signals from within the organism. To construct a crude example, lion-shape pattern of stimulation would be paired with fear-and-flight salience. Both systems of value-perceptual categorisation memory and of current perceptual categorisation can (and probably do) operate in the absence of primary consciousness.

Third, the concept of reentry and reentrant pathways requires brief explanation. Edelman and Tononi (2000a:48) describe reentry as a process of cyclical signalling:

> ...the ongoing, recursive interchange of parallel signals between reciprocally connected areas of the brain, an interchange that eventually coordinates the activities of these areas' [neuronal] maps to each other in space and time. This interchange, unlike feedback, involves many parallel pathways and has no specific instructive error function associated with it. Instead, it alters selective events and correlations of signals among areas and is essential for the synchronisation and coordination of the areas mutual functions

Reentry thus involves massive numbers of parallel, reciprocal pathways between different areas of the brain which, when operating, synchronise (or

even activate) the functioning of diffuse but related areas and systems within the brain. For example, reentrant signalling between the visual cortex and the part of the motor cortex dealing with eye movement would result in the synchronisation of such movements in relation to what is being processed and attended to within the visual cortex.

The key to Edelman's account of primary consciousness thus lies in reentrant pathways and activity linking the special memory repertoires and the areas associated with each exteroceptive sensory modality charged with perceptual categorisation. This account has two key facets. First, reentrant pathways allow the animal to coordinate activation of past value-category pairings when the respective neural group for this category is activated in the perceptual categorisation area of the brain. This produces the real time interaction between memories of past value-category associations and current perceptual categorisations *before* the value determining parts of the nervous system have a chance to change these memories based on current homeostatic states. That is, the organism is enabled to maintain a degree of synchronisation between acquired memories and current, ongoing perceptual categorisations, such that these memories can influence the relative salience-to-the-organism of particular perceived events, and assist in the choice of goal-directed actions. Such a mechanism is the fundamental basis of learning.

The second facet of the account is that this discrimination takes place in parallel across *all perceptual modalities*. That is, the reentrant signaling enables a unification and synchronisation of 'processing' in all modalities, with each modality area able to simultaneously activate associated value-category pairings within the memory system. Taken together, these two facets reveal the significance of the unified conscious phenomenal experience that thus arises for the organism: in the absence of these mechanisms of primary consciousness, salience in the array of signals arriving simultaneously at the various sensory modalities and perceptual categorisation areas would be determined *almost entirely by* the dominance of one category over another in each of the parallel sensory channels (e.g. the loudness of a sound, size or proximity of an object, strength of a smell), rather than by the adaptive values of the animal. Primary consciousness thus 'allow[s] an animal ... to direct attention to particular events in a selective fashion that serves its own adaptive needs' (Edelman 1989:93).

This supposedly based account of primary consciousness seems to me highly compatible with the spirit of Brooks' (2006) proposed strategy for solving the hard problem, and bodes well for further empirical and

gical work in this area. First, Edelman's (1989) explanation of the significance of the phenomenal or qualitative aspect of primary consciousness also seems to draw on a view of these qualitative features as tags for the organism. Where the explanation seems to differ from Brooks' account is that the significance of these tags lies more at the level of tagging (or representing) stored (and remembered) associations of value or salience for the organism. Edelman's theory does not argue against the significance of quality spaces and edge detection mechanisms, but it suggests that these are not sufficient to explain consciousness. Primary consciousness becomes necessary in linking perceptual categorisations to memories of value-category pairings, and it is this feature of consciousness that has such great adaptive advantage for the organism.

Second, Edelman's theory also makes much of the adaptive value that is gained by having a mechanism that allows for appropriate direction of attention (i.e. awareness). Again, however, his theory differs from Brooks' account by moving beyond mere perceptual sensory considerations (presumably coordination of sensory receptors with perceptual categorisation activity in the cortex could be all that is necessary for achieved 'binding'). Yet the adaptive advantage gained by primary consciousness enables a more adaptive use of attentional resources by (1) integrating and coordinating the perceptual categorisation activities of multiple modalities (i.e. integrating what would otherwise be parallel processes) into a unified mental scene, and (2) by efficiently allocating salience for the-organism-vs-non-categorisations within this integrated mental scene such that attention can be directed, both within and across sensory modalities, to environmental stimuli that have most adaptive significance for the organism. In a sense, primary consciousness is thus a vehicle to liberate categorisations from 'the prison of the present', as well as making the organism less of a slave to the intensity of sensory perceptual input. Brooks was thus correct in his rough outline of the adaptive significance of integrated mental scenes constituted by quality spaces associated with each sensory modality. What was missing was an account of how salience-for-the-organism enters the picture. For Edelman's theory, it is the connection between perceptual categorisation and salience values that holds the key to understanding the evolutionary advantage of primary consciousness.

Unanswered Questions: Future Directions

Where does this leave those of us interested in the interdisciplinary study of

consciousness? In the area of philosophy, it seems clear that above integration requires more detailed development into a systematic attempt to address the hard problem of consciousness. This will, at the very least, involve (a) deciding how much of the hard problem can be solved by way of such an account of the evolutionary significance of primary consciousness (which, it should be remembered, requires basic phenomenal or qualitative states rather than fully fledged qualia) and then (b) developing the theory, in conjunction with an account of higher-order consciousness, to provide an explanation of the evolutionary significance of qualia themselves as well as the other features of higher-order consciousness described by Edelman (1989; Edelman & Tononi 2000). Of key significance will be the evolution of language, an account of its relations to and dependence on primary consciousness, and elaboration of the cognitive, perceptual and behavioural capabilities that language confers on creatures endowed with linguistic higher-order consciousness[1].

In the field of (evolutionary) neuroscience, a wealth of empirical and theoretical work lies in wait for the would-be researcher. Edelman himself is insistent that his model of primary consciousness (Edelman 1989), and his more recent 'dynamic core hypothesis' (Edelman & Tononi 2000) that builds on this model, are testable and *refutable* scientific hypotheses that require investigation and confirmation rather than *a priori* dismissal. Sadly, the research technologies involved in much of this research (PET and functional MRI scanning) are extremely costly for a society and academic environment like our own. But with the appropriate level of interest and international collaboration, South African researchers would do well to contribute to this exciting and boundary pushing area of research. We do, after all, live in the 'Cradle of Humankind', and it would be fitting for us to contribute something to understanding the evolutionary 'birth' of human consciousness so as to complement our considerable contributions to, for example, paleoanthropology and the fossil record.

School of Human and Community Development
University of the Witwatersrand

[1] A current project of my own involves investigating the role of higher-order consciousness and its associated representational and linguistic aspects in generating the behavioural capacities that are typically associated with free will

References

Baars, Bernard J. 1996. *In the Theater of Consciousness: The Workspace of the Mind*. New York: Oxford University Press.

Baars, Bernard J. 1997. Contrastive Phenomenology: A thoroughly Empirical Approach to Consciousness. In N Block, O Flanagan & G Güzeldere (eds): *The Nature of Consciousness: Philosophical Debates*. Cambridge: MIT.

Brooks, David 2000. How to Solve the Hard Problem: A Predictable Inexplicability. *Psyche* 6. http://psyche.cs.monash.edu.au/v6/psyche-6-04-brooks.html

Chalmers, David 1995. Facing up to the Problem of Consciousness. *Journal of Consciousness Studies* 2:200-219.

Churchland, PM & PS Church eds. 1997. *On the Contrary: Critical Essays, 1987-1997*. Cambridge, Mass.: MIT Press.

Crane, T & DH Mellor 1991. There is no Question of Physicalism. In Mellor, DH (ed): *Matters of Metaphysics*. Cambridge: Cambridge University Press.

Crane, T & DH Mellor 1995. 'Postscript' to 'There is no question of physicalism'. In Moser, PK & JD Trout (eds) *Contemporary Metaphysics: A Reader*. London: Routledge.

Crick, Francis 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*. London: Simon & Schuster.

Dennett, Daniel C 1991. *Consciousness Explained*. Boston: Little, Brown & Co.

Edelman, Gerald M 1989. *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.

Edelman, Gerald M 1992. *Bright Air, Brilliant Fire: On the Matter of Mind*. London: Allen Lane/Penguin.

Edelman, Gerald M & Giulio Tononi 2000. *Consciousness: How Matter becomes Imagination*. London: Allen Lane/Penguin.

Flanagan, Owen & Thomas Polger 1996. Zombies and the Function of Consciousness. *Journal of Consciousness Studies* 2:313-321.

Gennaro, RJ 2000. Fiction, Pleasurable Tragedy, and the HOT Theory of Consciousness. *Philosophical Papers* 29:107-130.

Johnson-Laird, Philip 1983. *Mental Models*. Cambridge: Cambridge University Press.

Kosslyn, S & J Pomerantz 1981. Imagery, Propositions and the Form of Internal Representations. In Block, Ned (ed): *Readings in the Philosophy of Psychology 2*. Cambridge, Mass.: Harvard University Press.

McGinn, Colin 1989. *Mental Content*. Oxford, Basil Blackwell.

Nagel, Thomas 1974. What is it Like to be a Bat? *Philosophical Review* 83:435-450.

Pitman, Michael M 1998. Mental Causes, Supervenience and Psychological Explanation: A Defense of the Causal and Explanatory Relevance of Content. Unpublished MPhil dissertation, University of Calgary, Dept. of S

Polger, Thomas & Owen Flanagan 1996. Explaining the evolution of consciousness: the other Hard Problem. Paper presented at the Tucson II Conference: *Towards a Science of Consciousness*. Paper available at http://www.duke.edu/~twp2/tucson2.html

Searle, John R. 1992. *The Rediscovery of Mind*. Cambridge, Mass.: MIT Press.

Searle, John R. 1997. *The Mystery of Consciousness*. London: Granta Books.

Shear, Jonathan (ed) 1997. *Explaining Consciousness: the 'Hard Problem'*. Cambridge, Mass.: MIT Press.

Stich, SP & S Laurence 1994. Intentionality and Naturalism. *Midwest Studies in Philosophy* 19:122-142.

Thagard, Paul 1996. *Mind: Introduction to Cognitive Science*. Cambridge, Mass.: MIT Press.

Why Think that Cognition is Distributed?

David Spurrett

A number of other papers in this issue of *Alternation* (Blair and Cowley, Dellis and Spurrett, Cowley) pursue their particular topics of interest within what is known fairly generically as 'distributed cognition'. The primary purpose of this short paper is to survey some of the case for distributed cognition, or to say why it is that we think it makes sense to regard cognitive processes as extending beyond the brain.

First, though, it will be worth spending a little time considering why it can seem so obvious that cognitive processes are either co-extensive with the central nervous system, or perhaps with some sub-section of it, *i.e.* that cognition takes place 'in the head'. After Cowley and Spurrett (forthcoming), we call this view 'cognitive internalism'. There is, after all, ample evidence for the dependence of cognition on the brain. Persons without brains don't do any thinking. Brains certainly *aren't* there for any of the exotic purposes such as cooling the blood proposed by some ancient authors—they are there to do something with the information transmitted via nerves to the sense organs, and to do *something* else to the nerves connecting the brain with muscles and other engines of bodily activity. It makes a lot of sense, at least *prima facie*, to see the brain as being a bounded system with definite inputs and outputs, and to think that what it does is, somehow, to solve a control problem—finding the best outputs given the inputs.

One large body of evidence, collected especially since the nineteenth century adds to this picture the idea that the brain is regionally specialised for particular functions, and another indicates that it has a particular kind of processing architecture. The first body of work painstakingly correlated variations in performance on particular cognitive problems with haphazardly collected instances of localised brain damage, resulting from either surgery or misadventure. Although the tides in the debates over local versus global processing moved and continue to move, it is clear that particular regions of the brain, whether or not they handle *all* of the work relevant to particular types of performance, are at least crucially important for, *inter alia*, such functions as motor control, speech production, speech comprehension, visual processing and face recognition[1].

The second body of work concerns the finer structure of the brain, one follows from the 'discovery' of neurones and the ways that they are interconnected. The existence of nerves had been established for millennia and the action of nerves, especially motor nerves, studied extensively from the sixteenth century onwards. It was only in 1873 (through the work of Camillo Golgi) that the apparently homogeneous mass of the brain was shown, through a process of treating pieces of brain tissue so that microscopic examination could resolve individual neurones, to have a structure composed of many interconnected neurones, which appeared to operate in similar ways, at least, to nerve cells[2]. The discoveries here fed into and reinforced some of the work on localisation, since they provided a way of thinking about what was happening when some part of the brain was stimulated by an electrical pulse, and by extension when nerves ordinarily stimulated one another. They also enabled more serious thinking about the lower level processing architecture of the brain to begin, just because it was possible now to imagine the brain as being, in some sense, a vast collection of elements which were themselves something like switches.

At another level of abstraction entirely, a significant body of thought was already treating thinking as essentially computational. Central to this was Boole's approach to logic, which formalised propositional reasoning, at least in a way apparently amenable to mechanical treatment, as a series of applications of syntactic rules to representational tokens. Boole himself called his 1853

[1] The founding work on localisation took place in the mid-nineteenth century, by such pioneers as Paul Broca, Carl Wernicke, David Ferrier, Gustav Fritsch and Eduard Hitzig. Glynn (1999) includes an accessible survey with references to the primary literature.

[2] The conjecture that the brain was a network of some sort was much older than the first method of examining the network. In the 1620s Descartes (1985: 107) had written, in his *Treatise on Man*, that one would 'not be able to imagine anything more plausible than that [the brain] is composed of many tiny fibres variously interlaced ...'.

book *An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities*, making quite clear that the serial and symbolic computational vision he articulated was precisely intended to capture what was significant about human thought in general.

Although this 'fit' between a roughly Boolean image of computational thinking, and a neurally inspired image of brain activity is not self-evident, the two ways of thinking naturally converged as a research programme, set on its course by founding contributions from, e.g., Turing and Von Neumann. Given the expectation that thought was computational, and the discovery that brains appeared to be some kind of computing device, the task seemed to be to find out how the brain did whatever it did, so that our own thinking turned out to be computational in the right kinds of ways. That it so say that it was almost obvious, although by no means necessary, to suppose, first, that the 'problem' of optimising outputs on the basis of inputs is a primarily computational problem, and second that what the brain is for is performing those computations in a space between, and bounded by, perception and action[1]. This is pretty much what early, in the words of Haugeland (1985) 'Good Old Fashioned Artificial Intelligence' (GOFAI) did, viewing cognition as a 'computational process that takes and input and produces an output' (1997) and so Simon and Kaplan (1989) could quite fairly say that 'The computer was made in the image of the human'.

GOFAI has produced interesting and important results, and remains a significant research programme. Nonetheless there is a growing sense that real human and animal intelligence is something different entirely. Sketching the reasons for this is the main burden of the remainder of this paper.

The Great Escape

Cognition did not break out of the internalist dungeon all in one go. As is almost typically the case, a variety of voices had been raised against it, from a range of different backgrounds, disciplinary perspectives and theoretical orientations. Prior to, or roughly contemporary with, the beginnings of artificial intelligence, one could mention Vygotsky, Heidegger, Wittgenstein,

[1] Two of the historical figures being given especially short shrift here are Babbage, with his work in the nineteenth century making calculating machines, and Turing (e.g. 1950), who in the mid twentieth century did a great deal to generalise the notion of a computational architecture.

Merleau Ponty, Von Uexkull and others[2]. During the height of early GOFAI, there were Gibson, more Vygotsky, followers of Heidegger and others still. All, in different ways and for different although often overlapping reasons saw fault with cognitive internalism and sought in their work to describe ways in which cognition was bound up with bodily experience and/or the very 'external' environment which cognitive internalists took to be utterly separate from cognition. More than theory, though, the present survey is concerned with evidence. A number of particular empirical studies of different types of cognition show in a range of ways how cognitive processes latch onto, exploit and are intimately connected with the bodily, physical and social environments in which they take place, or that cognition is itself context and role merely related to brain function.

There are, in the literature, two senses of 'distributed' which it will be worth distinguishing. The first, associated with neural networks and some other parallel processing architectures, indicates that cognition is distributed *within* the brain, or computational system. Neural networks, the most brain like type form of parallel distributed processing, or PDP, are now a well established technology. Depending on the particular network, and the system mapping its inputs and outputs onto features of the world or some coding system, a neural network can implement a computational solution to a problem which does not involve internal representational symbols in the way expected by traditional artificial intelligence. Systems which are distributed in this way do not typically involve any form of central processing. Even a system which is highly distributed in this sense, though, need not be at all distributed in the second sense[3].

A system which is distributed in the second sense exploits resources besides those of the brain or any other obviously information processing architecture (which means bodily or environmental resources, or both) for *cognitive purposes*. This is a more radical sense of distributed cognition, since it denies part of what can make it seem so natural and simplistic to think that all cognition is handled by the brain. This concerns the fact that neurones do

[2] See, e.g., Heidegger (1962), Vygotsky (1986), Merleau Ponty (1962), Von Uexkull (1943) Wittgenstein (1953).

[3] See, e.g., Gibson (1979); Dreyfus (1979, 1991)

[4] Clifford (1998) is a recent example of a work forcefully urging a view which is distributed in the first sense, yet remains strikingly cognitive internalist.

not connect up to the world, but that they do connect up to sensory and motor organs.

There is little to do in that cognition is distributed in the first sense. That commitment – of course, leaves open just what the actual computational architecture of the brain is, and in particular in what respects and to what degree it is modular, hence allowing for functional decomposition at levels intermediate between individual neurones and the brain as a whole. It is the case for cognition being distributed in the second sense which is of primary concern here.

One set of examples and evidence can be found with reference to locomotion. Thelen and Smith (1994, 88) point out that traditional cognitive functionalist approaches to locomotion had been encouraged by evidence indicating that locomotion in many vertebrates seemed to be controlled by a pattern of activity generated in the spinal cord. This pattern could be generated and co-ordinated locomotion on a treadmill even in animals where the brain had been removed entirely. Furthermore the pattern could be shown to be produced even when the muscles it 'controlled' had been artificially paralysed. It was widely supposed that the spinal column, considered as a part of the brain, was home to a central pattern generator which was the driving mechanism for locomotive activity.

Let us leave this hypothesis aside for a moment, and look at some recent examples of effective locomotion in robots. The case of robots is illuminating precisely because it is with robots that the most striking examples of artificial or manufactured cases where a computer has access to bodily and environmental resources are to be found.

The hexapod robot built by Ilspenschied (see Quinn & Ilspenschied 1993, and also Nelson *et al.* 1997) has processing capabilities limited to a mere 57 artificial neurones, in an array inspired by biological precedents in nonvertebrate invertebrates.[3] Each leg has its own set of neurones, generating walking rhythms. These could easily be mistaken for the hypothesised central pattern generator, except for the fact that these rhythms are in turn modulated by additional connections between the sets of neurones peculiar to each leg, which in turn allow individual legs to inhibit the activity of neighbouring legs. Further, each leg-specific set of neurones received some 'sensory' feedback from feature detectors built into that leg, which tracked information about joint position and pressure between the foot and the surface beneath. The neural set for each leg controlled the swing (forward and backward) of the leg and, lifting and lowering of the foot.

The overall network of neurones is highly distributed in the first sense noted above, so there is no 'central executive' doing the processing, and, crucially, no stored plan. Incorporating tendencies to some key types of motion, also modelled on biological understanding of nerve functioning, has the result that the robot engages in 'searching' behaviour with individual feet, which enables a single foot to find a secure footing before another is 'permitted' (by the cessation of inhibitory signals from the first leg) to do the same, and 'clambering' behaviour which enables it to navigate obstacles which block its path. The resultant locomotion is capable of dealing with varying terrain and, importantly, exhibits what is known as 'graceful degradation' which is to say that limited and selective 'damage' to isolated parts of the system (i.e. a few neurones, sensory feedback paths, or part of the physical structure of the robot) has relatively little effect on the quality of overall performance. This flexibility and efficiency is achieved by means of exploiting bodily and environmental resources, instead of, for example, time consuming and computationally expensive visual perception followed by modelling of the environment in three dimensions, with the robot placed in that model at a specified position and with a particular orientation, as a method of determining where and how to place the feet.

So even if there is something which works somewhat like a central pattern generator in these cases, it is clearly not doing all of the work. Rather a coalition of resources including non-neural ones *together* account for flexible and responsive locomotion. It turns out to be just the same in vertebrates, as Thelen and Smith found. They argue that there is no evidence for the view that there is some centralised cognitive 'plan' for locomotion, and plenty of evidence against it. They show that varying non-neural factors (by placing the feet of a walking infant on a treadmill where the left foot moves at a rate different to the right, by increasing the effective mass of a limb by means of weights, or by decreasing it by partly placing the infant in a buoyant medium etc.) significant variation in the patterns of locomotion result. Actual performance, as with the robot case, depends not on the activity of the spinal column alone, nor on any purely neural collection of effector processes but rather on a coalition of elements only some of them neural. As a result, they maintain that locomotion is modular in nature, depending on a variety of resources, and where successful performance is an emergent phenomenon arising from the interaction of the various roughly modular elements.

In case the example of locomotion does not seem cognitive enough, one may consider the behaviour of human players of the computer game

Tetris. A classical cognitive internalist account of Tetris playing would separate perception of the current situation of the playing area, and the most recently appearing 'piece' from, on the one hand, computation directed to working out where best to place the new piece, and on the other, initiating a motor plan for placement on the basis of this computation. A full account of such an approach, based on the planning literature, is offered by Kirsh and Maglio (1994), and assumes processing into four phases:

> 1. Create an early, 'lumpy' representation of selected features of the current situation.
> 2. Encode this lumpy representation in a more compact, chunked, symbolic representation.
> 3. Compute the best place to put the piece.
> 4. Compute the sequence of moves to achieve the goal placement.

In their research on Tetris playing, Kirsh and Maglio (e.g. 1994; 1997) used this model as a source of comparisons with the actual measured behaviour of human players. The more human players handle the cognitive demands of the game in line with classical expectations, the more the data should fit that model. The more difficult it is to square the data with the model, the more likely it is that humans are doing something else. It is important to note that in the standard planning model computation and action are radically separated: the placement problem is solved *first* as an internal computation, and only

---

[1] Tetris is a game involving the rotation and translation of 'falling' shapes (moving from the top to the bottom of the playing area) in order to achieve optimal packing of the bottom of the playing area. Completely filled horizontal rows are cleared in comparison, with any incomplete rows above them moving down for the spaces thus created. The game ends when the accumulation of incomplete rows caused by player error prevents any new pieces from entering at the top of the playing area. Additional reference to Kirsh and Maglio on Tetris can be found in Rietveld and Greeley (this volume) and Dolan and Spurrett (this volume).

[2] 'Piece' is the term Kirsh and Maglio use to refer to the individual 'pieces' or 'elements' which fall in the Tetris game.

[3] This would involve translating the raw image into a version that encoded edges, concave corners, convex corners, etc.

once that solution has been worked out is there any role at all for action in the world.

Kirsh and Maglio therefore stated that one consequence of the standard model is that, if it is basically correct, any actions taken before the completion of the internal computation would be, of necessity, non-optimal. This would obtain just to the extent that human action approximates that prediction: there should be a delay between the appearance of a new piece and the commencement of rotation and translation. They state, though, that this 'is patently not what we see in the data' (Kirsh & Maglio 1994). The first gross indication of this mismatch between the predictions of a classical planning model and the actions of real human players is that effective players engage in rotations almost immediately, and that the number of rotations and translations made were significantly in excess of what the internalist model predicted as being optimal.

On the basis of a series of carefully constructed additional experiments Kirsh and Maglio demonstrated convincingly that these 'surplus' actions *all* yielded computational advantages: that they facilitated earlier recognition of particular shapes, simplified the problem of detecting edge fit, reduced the number of vertical placement errors, and so forth. Kirsh and Maglio (1994) say the following in their conclusion.

> This way of thinking treats the agent as having a more cooperative and interactional relation with the world: the agent both adapts to the world as found, and changes the world, not just pragmatically, which is a first order change, but epistemically, so that the world becomes a place that is easier to adapt to. Consequently, we expect that a well-adapted agent ought to know how to strike a balance between external and internal computation. It ought to achieve an appropriate level of cooperation between internal organizing processes and external organizing processes so that, in the long run, less work is performed.

Kirsh and Maglio coined the term 'epistemic action' to describe these actions, which are paradigm examples of distributed cognition. They are to be distinguished from 'pragmatic' actions, where pragmatic actions are oriented towards achieving some physical goal. Epistemic actions on the other hand are real physical actions that are performed in order to modify the computational state of the system performing the action, in this case the person playing the game. It is worth pointing out that this reliance on the external world for

suggests support is given for a developmental phase observed in beginner or intermediate players and less so or not at all in advanced ones. Further work by Maglio and Kirsh (1996) shows that the prevalence of epistemic actions actually increases as skill advances. Epistemic action, or distributed cognition, is not a developmental stage, but an apparently essential property of increasingly optimal solutions. What this research shows, in one carefully selected area (although additional work by Maglio *et al.* (1999) indicates that the type of result found with Tetris is robust with respect to Scrabble playing as well), is that human cognition relies heavily on non-neural resources, that is the manipulative capabilities of the hands, and the structure observable in the environment, in order to solve essentially computational problems.

Part of what makes Kirsh and Maglio's research so compelling is that by using the game of Tetris, they were able to collect fine-grained data concerning human behaviour which could be evaluated in a rigorous manner, given the computational properties of the game. Good players are those who achieve high scores, and good rapid placements are ones which don't create gaps. This feature of the game meant that the specific patterns of epistemic action characteristic of good play could be and were discovered and further studied. Kirsh and Maglio could tell relatively easily who the good players were and focus on making sense of what they did.

In a similar way, good navigation teams are ones whose ships don't go off course. The documented plan concerning where a ship is *supposed* to be at any given time permits evaluation of the practices which lead to it being either there, or somewhere else. In his extended study of the navigation problem, Hutchins (1995, see also Hutchins and Klausen 1996) brought the skills of a cognitive anthropologist to bear on a particular type of situation in which humans co-operate in the attempt to solve a computational problem. Hutchins (1995:12) glosses navigation as "a collection of techniques for answering a small number of questions, perhaps the most central of which is 'where are we?'" In the case of a large ship, answering this question typically involves fixing the position and orientation of the ship on a chart by a specific operation in time, or reconciling information that can be gathered from on board the ship with that represented on the chart.

The particular ship on which Hutchins gathered most of his data, like many others, could not possibly be navigated by a single person[9]. The physical dimensions of the ship, and the arrangement of the navigation area

[9] That is, except when the ship is in the open sea.

and observation positions for taking bearings on landmarks were such that a co-ordinated team was essential. Hutchins' study focussed on the behaviour of the people taking the bearings, the people using them with bearings to take, the people recording those bearings, plotting them on the chart, and integrating them with other sources of information concerning the speed of the ship, the depth of water beneath it and so forth. Contrary to what a cognitive internalist might have expected, that is a collection of cognitively self-contained agents sharing information about elements of the navigation problem, so that some central executive agent could make decisions, he found that the team as a whole functioned as a kind of extended cognitive agent.

Hutchins argued that not only material labour, but also cognitive labour, can be socially divided, and that the ways in which cognitive division of labour is possible depend partly on cultural and material conditions. The particular social conventions of military navigation that Hutchins studied enabled a collection of people and artefacts (including charts, measuring instruments, and other navigation tools, often used in ways which is not one Maglio would describe in terms of epistemic action) to have properties paradigmatic of human (with computer systems, including function-specific "routines" (1995: 198) and protocols for "buffering" data prior to processing (1995: 197). Two important points need making here. The first is that the type of computational architecture Hutchins ascribes to a navigation team has some of the same general properties as the systems of robustness and flexibility (using what Tooby and Smith call "soft assembly" (Hutchins 1995: 185, 228, Tooby and Smith 1994: 60) or the absence of fixed control connections. The second point is that the "level" at which the system being studied most closely approximates a classical computer is that of the distributed system including the world. This is a point which has been made in many different ways recently. For his part Hutchins quotes Dennett (1991: 213, in Hutchins 1995, 361) speculating about how Turing could have considered the formal architecture now known as a Turing machine to be similar to human thinking, by taking a certain view of the activity of a working mathematician as paradigmatic. Hutchins himself points out that a danger here is mistaking the cognitive properties of a larger system in which a mathematician manipulates symbols in some medium such as pencil and paper for the cognitive properties of the brain of the mathematician. Him as Russell (see Rumelhart, Smolensky, McClelland, and Hinton 1986: 46). An encultured brain in a body with access to cognitive prostheses like writing media and symbol systems has powers

different to those of a 'naked' brain in virtue of those external resources. Distributed approaches also enables some sense to be made of the ways in which external products such as measuring instruments and systems of notation can fast-track different types of computation, and permit different forms of division of cognitive labour. Its examples concern navigation systems, but for the sake of brevity of explanation I will refer to number systems. The ways in which it is much easier for users of Arabic (or any other place-value system of numerals) to perform arithmetic, than it is for the users of Roman (or any other system of numerals not using place-values) has nothing to do with the different properties of the *brains* of individuals from different ancient civilisations. Rather 'hooking up' with a particular external system of signs and artifacts confers particular advantages and disadvantages.

On the subject of signs, Dennett (e.g. 1991) and Clark (e.g. 1997), among others, both make a great deal of the ways in which language itself can be regarded as a kind of cognitive prosthesis, permitting different and vastly more powerful forms of social co-ordination, and sharing of epistemological and other resources. In his analyses Clark's list of cognitive advantages includes the capacity for self-stimulation to improve control and performance of tasks (Clark 1997: 203), being able to use symbolic systems to augment memory by using new memory storage modes (Clark 1997: 201), using labels and symbols to simplify categorisation tasks and learning processes (Clark 1997: 201), and satisficing various other types of problem-solving. (See also Bloor and Cowley at this volume.) Clark tends to emphasise the ways in which our native cognitive powers can dovetail with language to extend our cognitive abilities, while Dennett favours the view that in learning language we engage in a process which reprograms our brains, permitting them in some ways to approximate the functioning of a von Neumann machine. What they agree on, though, is that our cognitive powers *with* language, where language is seen as a largely external, public resource, are vastly different to those without.

Rather than accumulate further individual fragments of evidence [1], I want to draw this survey to a close with a sketch of an argument for the *likelihood* of distributed cognition. This argument is by no means a proof that distributed cognition is the only possible type, although a developed form of the argument could perhaps show that distributed solutions to cognitive

[1] There is, at present, no better general survey than Clark (1997) and most of the examples here are also discussed in that work.

problems were at least homologically necessary (cf. solutions designed by natural selection).

In the realm of biology there is a certain temptation to adopt what could call 'functional internalism'. The view that the activities of any organism will be performed entirely by 'on board' components. Consider, in this light, the sponge (Wilkinson; Vogel 1988). Sponges feed by filtering water's bread in water. It has been known since 1766 that the force driving the unidirectional current internal to a sponge was flagellar. Vogel notes that

> ... an important related question was not asked for another hundred years. If the flagella were inoperative would water pass through a sponge anyway? Or, to put the matter in more realistic terms, does flagellar action account for all of the water pressure drop across a sponge, or can ambient water currents make a contribution to filtration? It appears now ... that not only do ambient currents help but that the structure of sponges is most exquisitely adapted to take advantage of such currents, with clear functions attaching to a number of previously functionless features. Dynamic pressure on the incurrent openings [being upstream, valves closing minute pores and downstream, oscillatory suction from the large fistal or apical excurrent openings] serves to give advantage from even relatively slow currents ... Why [it so much time elapse before someone made a crude model of a sponge, placed it in a current, and watched a stream of dye pass through it?] (Vogel 1981:190)

The answer to Vogel's question is, as he notes, that biology's been wedded to what I have just called functional internalism. Vogel's project is to reject this dogma, and for purposes a principle which should be adopted to help guide us away from that view: 'Do not develop explanations requiring expenditure of metabolic energy until simple physical effects are ruled out' (Vogel 1981:192). Andy Clark, who refers to Vogel (in 1989 and 1997) draws a moral for cognition, which tends in the direction of a famous fictional opportunistic exploiter of resources: he didn't bring with him, the 'Quiff Principle':

> In general, evolved creatures will neither store nor process information in costly ways when they can use the structure of an environment and their operations upon it as a convenient stand-in for

> [...] that not only processing operations connected. That is, know only as much as you need to know to get the job done (Clark 1989: 64).

The more general argument which it is possible to see here gives something like the process of natural selection is blind to the boundaries between brains and bodies, and between bodies and world. Not only that, the process of adaptation always takes place in a particular environment, without the particular security and competition problems of which there would be no problem at all. If well-tailored solutions to problems in general, whether the latter fitting challenge facing species or something more cognitive, are more efficient or effective then they will be favoured. The examples above are all more efficient or effective than internalist counterparts. Whether the argument can be made can more than a sketch, enough, would depend on showing that these more solutions have to be more effective, or are or less so widely over likely to do so. In that case we should expect them, and so forced back to constructing internalist options only when we are unable to make sense of some phenomenon as distributed terms. The papers in this collection by Blair and Cowley, Dellis and Spurrett, Cowley and Duvall are not motivated by the notion that cognition *has* to be distributed, but they do all show it in different ways as a working hypothesis that it is more likely to be than not. If that is true, then the investigation of human cognitive capabilities, and the project of maximising them, has far more to do with analysing the cognitive properties of artefacts, environments, bodies and cultures than traditionally expected.

Department of Philosophy
University of Natal

References

Boole, G 1853. *An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities*. London: Macmillan.

Blair, G & S Cowley (this volume). Language in Iterating Activity: Metacognition Re-remembered.

Brooks, RA 1997. From Earwigs to Humans. *Robotics and Autonomous Systems* 20(3-4):291-304.

Cilliers, P. 1998. *Complexity and Postmodernism: Understanding Complex Systems*. London: Routledge.

Clark, A 1989. *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge, Mass.: MIT Press.

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

Cowley, S & D Spurrett, D (forthcoming). Putting Apes (Body and Language) Together Again. To appear in *Language Sciences*.

Dellis, A & D Spurrett (this volume). Real Patterns and Distributed Cognition.

Dennett, D 1991. *Consciousness Explained*. Boston: Little, Brown.

Descartes, R 1985. *The Philosophical Writings of Descartes*, Volume 1. Cottingham, JC, R Stoothoff & D Murdoch (trans & eds). Cambridge: Cambridge University Press.

Dreyfus, H 1979. *What Computers Can't Do*. New York: Harper and Row.

Dreyfus, H 1991. *Being-in-the-world: A Commentary on Heidegger's* Being and Time, *Division I*. Cambridge, Mass.: MIT Press.

Gibson, JJ 1979. *The Ecological Approach to Visual Perception*. New York: Houghton Mifflin.

Glynn, I 1999. *An Anatomy of Thought: The Origin and Machinery of Mind*. London: Phoenix.

Haugeland, J 1985. *Artificial Intelligence: the Very Idea*. Cambridge, Mass.: MIT Press.

Heidegger, M 1962. *Being and Time*. Macquarrie, John & Edward Robinson (trans). Oxford: Blackwell.

Hutchins, E 1995. *Cognition in the Wild*. Cambridge, Mass.: MIT Press.

Hutchins, E & T Klausen 1996. Distributed Cognition in an Airline Cockpit. In Engeström, Y & D Middleton (eds): *Cognition and Communication at Work*. New York: Cambridge University Press.

Kirsh, D & P Maglio 1994. On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science* 18:513-549.

Kirsh, D & P Maglio 1997. Some Epistemic Benefits of Action: Tetris, a Case Study. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. Hillsdale, New Jersey.

Merleau-Ponty, M 1962. *The Phenomenology of Perception*. London: Routledge and Kegan Paul.

Meglio, PP & D Kirsh 1990. Epistemic Action Increases with Skill, *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.

Maylie, PF, T Matlock, D Kaphaezy, B Chernicky & D Kirsh 1999. Interactive Skill in Scrabble. In *Proceedings of Twenty first Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.

Nelson, GM, RD Quinn, A Bachmann, WC Flannigan, RE Ritzmann, JT Watson 1997. *Design and Simulation of a Cockroach like Hexapod Robot*. Proceedings of the 1997 IEEE International Conference on Robotics and Automation (ICRA '97). Albuquerque, NM. April 22-24, 1997.

Simon, HA & CA Kaplan 1989. Foundations of Cognitive Science. In Posner, M (ed) *Foundations of Cognitive Science*. Cambridge, Mass. MIT Press.

Quinn, R & K Espenschied 1992. Control of a Hexapod Robot using a Biologically Inspired Neural Network. In Beer, R, R Ritzman & T McKenna (eds). *Biological Neural Networks in Invertebrate Neuroethology and Robotics*. New York: Academic Press.

Rumelhart, DE, P Smolensky, JL McClelland & GE Hinton 1986. Schemata and Sequential Thought Process in PDP Models. In Rumelhart, DE & JL McClelland (eds): *Parallel Distributed Processing*. Volume 2. Cambridge, Mass: MIT Press.

Thelen, E & LB Smith 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, Mass.: MIT Press.

Turing, AM 1950. Computing Machinery and Intelligence. *Mind* 49:433-460.

Von Uexkull, J 1934. A Stroll through the Worlds of Animals and Men. In Lashley, K (ed): *Instinctive Behavior*. New York: International Universities Press.

Vygotsky, LS 1986. *Thought and Language*. Cambridge, Mass.: MIT Press.

Wittgenstein, L 1953. *Philosophical Investigations*. Anscombe, GEM (trans). Oxford: Blackwell

Book Review Article

The Contemporary Verse Novel: An Explosion of Poetic Narrative

Book Review Article

The Golden Gate
by Vikram Seth
New York: Faber & Faber,
1987, 308 pp.
ISBN: 0571148271

Omeros
by Derek Walcott
London: Faber & Faber,
1990. 325 pp.
ISBN: 057.144594

Love
by Bernadine Evaristo
Tunbridge Wells: Angela Royal
Publishing. 1997. 147 pp.
ISBN: 1899860452

The Emperor's Babe
by Bernadine Evaristo
London: Penguin,
2002. 255 pp
ISBN: 0140297812

Bloodlines
by Fred D'Aguiar
London: Vintage,
2001, 161 pp.
ISBN: 0099284421

Whylah Falls
by George Elliott Clarke
Vancouver. Polestar Book
Publishers, 2000, 206 pp.
ISBN: 1896095808

Sinking
by Michael Conrad Green
London: Penguin.
1997, 164 pp
ISBN: 014058790X

Reviewed by Catherine Addison
Department of English
University of Zululand

1. Introduction and Background

Ever since the appearance in 1986 of Vikram Seth's *The Golden Gate*, novels in verse have been the *dernier cri* of literary fashion in the English-speaking world. From Australia to the USA, from the *avant garde* to the establishment, writers of both poetry and fiction have taken up this hybrid form with enthusiasm. Now that it has appeared recently over the millennium, often as a daring first publication, 'A Verse Novel' or 'A Novel in Verse' must be among the commonest subtitles of the twenty-first century.

This paper will briefly review several recent texts in the process of charting the origins and growth of the contemporary verse novel explosion. An introduction begins with a short defence of the genre, and then provides a sampling of its history and development up until the late twentieth century. Later sections will focus on the six individual works, sketching at the same time the genre's evolution up to the present. The selection of texts covered was governed by a consideration of the verse novel's relevance to Atlantic and South African contexts.

Purists, of course, might be tempted to deny this genre's existence, since narrative poems have always existed and the novel is a rather recent twentieth century prose phenomenon. Certainly, the new texts calling themselves verse novels are lengthy narrative poems, many of them clearly related to older poetry of this type, particularly epics and romances.

But genres are not closed, mutually exclusive categories. This is especially true of the novel, which, as Mikhail Bakhtin (1981:5) claims, has a marked tendency to incorporate other genres into itself. Why not then, poetry? There has been a noticeable tendency among prose novels in the past few decades – predating the outburst of verse novels – to embed poetry into their narratives. Memorable examples occur in Vladimir Nabokov's *Pale Fire* (1962), in D.M. Thomas's *The White Hotel* (1981), in A.S. Byatt's *Possession* (1990) and even Salman Rushdie's recent *The Ground Beneath Her Feet* (1999), in which the lyrics of rock songs written and performed by the main characters feature as an integral part of the text.

This is not to say that including a little poetry in a mainly prose narrative is the same as writing a novel entirely in verse. But it does demonstrate the novel's versatility, its potential for encompassing different voices and styles. And no reason absolutely forbids those more self-reflexive, more stylized, sometimes more emotionally charged voices expressible in verse from embodying all of a novel; for, as Bakhtin (1981:4 ff.) also claims, the

novel is essentially a dialogic, experimental genre, defying strict definition by its continual protean self-transformations. If the novel could, by the beginning of the nineteenth century, already include such extremes as the works of Jane Austen and of Laurence Sterne, then it is no surprise that it could tolerate the invention of the verse novel not long afterwards. To add formal verse to its possibilities was not to catalyse the most extreme metamorphoses of which the genre is capable.

So much for the novel – but what about poetry? If it is easy enough to distinguish a verse novel from a prose novel, it is not so simple to distinguish a verse novel from a narrative poem that has no novelistic pretensions. Here, we find, the defining features are mostly less objective than subjective, the intentions of author or reader playing a decisive role. As long as a poem is conceivably narrative, its parts in some way continuous one with another and the whole thing at least the length of a short book, an argument can surely be made for its status as a verse novel. Obviously, however, partisan readers' ways of seeing novels will influence their categorization of specific poems as verse novels. Readers in the line of Bakhtin (1981:3-40), who defines novels in opposition to epics, and those like Northrop Frye (1957:304-314), who distinguishes them from confessions, anatomies and romances, will develop quite different systems of inclusion and exclusion.

But the novel in verse is after all not a very new phenomenon, its age perhaps lending it some respectability. The first famous work to display the defining subtitle, 'A Novel in Verse', may have been Alexander Pushkin's *Eugene Onegin*, which appeared in Russian in 1825. Apart from its concession to stanzas, it is relatively conventional as nineteenth century novels go. Its story is contemporaneous with its writing, it deals with well-to-do people in society, focuses on love and love's relationship to marriage, includes both country-house and city settings and is intimately sympathetic to a female viewpoint. What makes it delightful is its narrator's playful wit, which, though made integral to the structure of the verse, is nevertheless at least as novelistic as it is a poetic feature.

Having accepted *Eugene Onegin*, thus designated by its author, as a verse novel, we may be constrained to label at least one earlier work in the same way. Byron's *Don Juan*, produced between 1818 and 1824, is so clearly the model not only for *Eugene Onegin* but for other verse novels by Pushkin that it should surely be categorized as a verse novel too. Though Byron (1926:648) described his own poem as 'epic', its comic, rambling, shaggy-dog narrative clearly resembles an eighteenth century novelistic romance such as

Henry Fielding's *The History of Tom Jones* (1749) more than a 'mere' epic. This is a point that Pushkin clearly appreciated.

Probably the first English work described by its author as a 'novel-poem' was Elizabeth Barrett Browning's *Aurora Leigh*, a gently feminist story of a woman poet published in 1856. A handful of other nineteenth-century works were also recognised as verse novels by their authors, including Owen Meredith's *Lucile* (1860) and Violet Fane's *Denzil Place* (1875). Professors of literature offering courses on the Victorian Verse Novel tend to add other titles to this list, for example, Clough's *The Bothie of Tober-na-Vuolich* (1848) and *Amours de Voyage* (1858), George Meredith's *Modern Love* (1862), Robert Browning's *The Ring and the Book* (1868–9) and even, amazingly, Tennyson's *Idylls of the King* (1859). Though his less popular *The Princess* (1847), having a more comic subject-matter, might fit the category better.

Clearly the genre, once accepted, may be cut to fit almost any long poem that we may like to put in it – but this does tend to be true of genres. And some family resemblances can be discerned by the believer. What all these long narrative nineteenth century texts have in common—with the possible exception of *Idylls of the King*—is that they are realistic and that they pay close attention to human behaviour in society. Indeed for the Arthurian *Idylls of the King* and the Italian Renaissance *The Ring and the Book*, they all have contemporary settings. However, their use of poetic verse form is totally variable. Aspects of theme, setting and treatment, in other words, are what mainly determine categorization as 'verse novels'. Those poems that resemble the 'typical' Victorian prose novel most closely in these aspects are of course most likely to be thus classified. Whether they be in blank verse (as is *Aurora Leigh*), *ottava rima* stanzas (*Don Juan*) or even anapaestic tetrameter couplets (*Lucile*) is not the issue.

Less than two decades ago, the verse novel was regarded as extinct. A largely nineteenth century phenomenon, it could be examined in its entirety in a graduate seminar. Term papers, like epitaphs, dispassionately accounted for its demise. During the great part of the twentieth century, the writing of narrative poetry had come almost to a standstill and long poems, when they do appear, focussed mainly on the philosophic moment and consisted of a series of lyrics. Two recently and somewhat unexpected exceptions to this rule are Joseph Moncure March's *The Wild Party* (1928) and Vikram Seth's *The Golden Gate* (1986), both of which are obviously verse novels. *The Wild Party* is a hard-boiled tale about the down-and-out girl Queenie and her amours, written in jazzy rhythms: "Jack's Cigarette – I've hurt my shin – "/ The door swung open / And the cops rushed

in! (1928:stanza). *Lucile Bulwer* belongs to the tradition of *Aurora Leigh*, being composed in blank verse and uttered by the voice of a respectable young woman during the First World War in London. But apart from another work by March (*The Set-Up*, 1928), little else that could be called a verse novel was to appear until the last decades of the century.

II. *The Golden Gate* and its Alternatives

Vikram Seth's *The Golden Gate*, appearing in 1986, was the great catalyst of the present verse novel explosion. This culturally quite local book, concerned with an Indian, set in late twentieth century San Francisco and composed in the complex Russian stanzas of Pushkin's *Negrate Onegin*, is unmistakably a verse novel. The term would have had to be coined for it had it not already existed. The narrative's focus is the everyday life of a group of young Americans who go to work, fall in love and live with each other, discuss about sex, politics and religion, protest about the environment, consume coffee, use drugs, have car accidents and worry about their parents. The plot is well structured and ends with a marriage and a death. Despite the difficult stanza, dialogue is frequent and sounds quite 'natural'. Due to of course the great virtues of *The Golden Gate* (including literary technical ones), it tells a fairly mundane tale well – and realistically—using a form so complex to have been invented in Turkish. Pushkin's stanza was designed for a language more rich with rhymes than is the case. The basic pattern is ABABCCDDEFFeghi. Hard enough as it stands, but the real snag is an alternation between masculine and feminine rhyme. The A-rhymes are all compulsorily feminine (involving a repetition of two syllables) as are C, E and G. Masculine rhymes (repeating one syllable only) are mandatory for B, D and F. To manage this stanza successfully in English—which has been done before only in large translations of *Negrate Onegin* so far as I know – is a feat of curiosity almost unprecedented in our stress poor language. And *The Golden Gate* consists of over 600 of these stanzas. Thus, reading the text is an experience of dizzying fluctuations between consciousness of technique and consciousness of story. The fictional tale is so much of 'really' real that we become sucked in and yet the style of writing is so 'artificial' that we must gasp regularly at its daring, dragging ourselves outside the illusion to admire.

Lights flash. 'Oh God! This means a ticket'

> The siren wails. John frowns. 'Well, I'm
> Not going to tell them they can stick it ...
> A cop comes. 'Licence please ... this time
> It's not a fine. The speed you're going
> Is hazardous. How about slowing
> Your pace a little, Mr. Brewer?'
> John fears him with a restive frown.
> With more cupidity'n then repentance
> Officer, work begins at eight'.
> The cop says, 'Better late than "late",'
> Signing an ticket on first sentence.
> 'Have a good day, now!' But John, cross
> Can't wait, and quarrels with the boss (2.9 - p78).

This colloquial effect with the simultaneous formal density is the Byronic experience *par excellence* of course, and Seth's inheritance via Pushkin of the mode is perfectly legitimate. However, Byron and Pushkin both chose stanzas that were better adapted to the languages they were writing in. *The Golden Gate* at its best is very good, but in the long haul one encounters many stanzas for which the epithet 'dogged' hovers at the tip of one's tongue.

Now although, as mentioned, *The Golden Gate* has operated as a catalyst for the production of scores of other verse novels, it spawned only one close stylistic copy. This novel, a whodunnit by the veteran British crime writer H.R.F. Keating, is set in 1929 and is called *Jack the Lady Killer* (1999), a title actually invented by Seth for an imaginary book that is read by one of his characters in *The Golden Gate*. But Keating, though he accepts the octameter and basic rhyme pattern of Pushkin's stanza, draws the line at the masculine-feminine alternation in the rhyme scheme—and sensibly so, in a language in which it has never been a convention. Keating justifies his foray into verse after a long career in prose in two interesting ways. He claims on the one hand that 'writing in verse, far from being an obstacle to easy reading, is a positive inducement to continue turning the pages' (Keating 2002:n.p.) and on the other that the verse form 'allows one to say extra things which might never have risen from the subconscious if one were writing straightforward prose' (Keating 'Preface' 1999:n.p.). How much this is! Only a short time ago poetry was one of the great turnoffs for the middlebrow reader and, for the critic at least, the capacity of formal verse to make an author say something that he or she hadn't intended was seen as kind of secret shame (Lodge 1977:89).

Jack the Lady Killer appeared as late as 1999 and was already not a single copy of an original but part of a definite trend. By then scores of new verse novels were being printed, as in the nineteenth century in a great variety of different styles and verse forms. They were also being written in many different parts of the world by people from a variety of ethnic origins. Australia may be a world leader in the form with the groundbreaking work of figures such as Dorothy Porter (*Akhenaten* [1992], *The Monkey's Mask* [1997]; *Wild Surmise* [2002]), Les Murray (*The Boys Who Stole the Funeral* [1980], *Dog Neptune* [1999]) and Alan Wearne (*The Nightmarkets* [1985], *The Lovemakers* [2002]) being imitated or universal upon by numerous other writers. But most English language countries have produced at least one verse novel and it may be more interesting to classify these texts by content than by country of origin. Subject matter includes science fiction (Frederick Turner, *The New World* [1985], Frederick Pollack, *Happiness* [1998], John Barnie, *Ice* [2001]), stories based on classical myth (Anne Carson, *Autobiography of Red* [1998]; *Akhenaten*), community and family histories (Craig Raine, *History: The Home Movie* [1994]; Brad Leithauser, *Darlington's Fall* [2001]), fictional community and family histories (David Budbill, *Judevine* [1991], Anthony Burgess, *Byrne* [1995]; Siddharth Kaviraj, *Dark Rooms* [2002]), mysteries (*The Monkey's Mask*; *Jack the Lady Killer*; Martha Grimes, *Send Reindeer* [1998]), love (Marilyn Hacker, *Love, Death, and the Changing of the Seasons* [1986]; Anne Carson, *The Beauty of the Husband* [2002]), documentary (Jordie Albiston, *The Hanging of Jean Lee* [1998]; Mary Jo Bang, *Louise in Love* [2001]) and fiction for children and young adults (Virginia Bower Wolff, *Make Lemonade* [1993]; Juan Felipe Herrera, *Crashboomlove* [1999])[1].

If there is any one feature that links all—or most of—these texts together it is probably their relationship to popular literature. Even though a significant proportion of them use strict verse forms, these are not works composed for a university audience only—unlike a good many of the more 'difficult' poems of the twentieth century (Lake 1999:39). The explicit aim of many of these poets, as Dana Gioia (1999:34, 39) among several commentators to a book on contemporary 'formal' poetry points out, is to appeal to the broad readership of the novel. It seems that Keating is not the only writer to see a potential for page-turners in long verse narratives. And even in their poetic

[1] It will be noticed that one or two of these verse novels actually appeared before Seth's 1986 *The Golden Gate*, suggesting that *The Golden Gate* itself may have been just a part—albeit an important part—of a trend already begun.

forebears, many recent verse novels show more influence from populist forms such as jazz songs and folk tales than any highbrow academic tradition.

III. *Omeros* and the African Connection

There is, of course, an African element to this new trend, though it is less dominant in Africa itself than in the diaspora. For probably the best poetic narrative to appear in print since the trend's beginnings is the Caribbean Derek Walcott's *Omeros*, published in 1990. This work is usually categorized not as a verse novel but as an epic, which is clearly what its author intended, considering that its title is the Greek name of Homer, that it contains characters called Hector, Achille and Helen, and that its stanza is an adaptation of *terza rima*, the form invented by Dante for his *Divine Comedy*. Certainly the poem has epic dimensions, for it pulls together history and traditions from Europe, Africa and the Caribbean, while still creating and transmitting the national myth of a particular place, St. Lucia, the island of Walcott's birth. But, as previously mentioned, genres are not exclusive categories, especially not the novel. *Omeros* is as much novel as epic. Its characters are nearly as realistically—and certainly as sympathetically—drawn as in a nineteenth novel like George Eliot's *Middlemarch* (1871–1872). For example, the elderly Plunketts, a British colonial couple, are finely observed in their frailties, disappointments and crudenesses, just as are the West Indian characters in their love triangles, their poverty, their heroic struggles with the sea, their postcolonial plight. If the style displays the dreaming, contemplative diffuseness of much lyric poetry, this should not exclude the work from the novel genre. Virginia Woolf's *The Waves* (1931) is no less 'prosaic' despite being, in prose, and *The Waves* is generally accepted as a novel.

Omeros succeeds resoundingly where *The Golden Gate* and other recent novels in verse do not because it achieves exactly the right combination of the traditional versification. *Terza rima*, unlike Pushkin's stanza, is well-suited to the English language and has been used in it successfully before (for example, by Shelley in 'Ode to the West Wind' and 'The Triumph of Life'). This invention of Dante's is not exactly a stanza so much as a running design, with the enclosed end syllable of one tercet reappearing as the enclosing rhyme of the next: ABA BCB CDC DED... etc. What Walcott does in *Omeros* is to apply the verse's rules only loosely, approaching and withdrawing from the pattern according to the exigencies of his discourse, so that at times he appears to be writing a kind of free verse:

Plunkett never thought he would ask the next question.
'Heaven?' He smiled.
'Yes. If heaven is a green place'.
And her shut eyes watered while his, as a wave open.

That moment bound him for good to another race.
Then the Major said: 'Tell her something for me, please'.
'She can hear you', the *gardeuse* said. 'Just like in life'

'Tell her', said the Major, clearing his throat, 'that's the ... that time when I slammed them, I'm sorry that I cursed her all that pain. Tell her'—he stopped—'that in the next world

would have been so much'. He lifted the small saucer where the candle had shrunk to a stub, and he edged a twenty-dollar bill under it, next the Bible. (307)

But this moving little passage in which Plunkett consults the obeah-woman, Ma Kilman, in order to communicate with his recently-dead wife is not actually 'free' at all. The line chosen is very regular but more subtle than the traditional English iambic pentameter, being measured in syllables rather than in stresses or beats per line. Nearly every line composes exactly twelve syllables and yet, like most syllabic effects in English, this regularity remains subliminal for the average reader. As for the *terza rima* rhymes, some are imperfect, as is 'question/open' in the enclosing lines of the first tercet, and some are non-existent, as in the enclosing lines of all the other tercets. But the pattern does not disappear completely, for the rhyme linking the enclosed line of one tercet with the opening line of the next is characteristically present, as in all cases here: 'pleat/heat', 'pleasekeys', 'caused/harsanneer'. This haunting but insistent formal presence comprises the very flexible basis of the work's brilliance. Walcott's apparent compromise is an implied choice for an age in which a repeated formal pattern can seem overly self-conscious.

IV. *Love* and *The Emperor's Babe*

Omeros, being epic in scope, covers both the colonial and the present dispersal of the West Indian people who are its chief focus. In addition to sections set in

America and in several of the great cosmopolitan cities of the world. It includes one that follows Achille on a visionary journey across the ocean and back to return to West Africa in the age of slavery, where he meets his own ancestors and lives with them for a while. This fictional return seems to have struck a chord with other poets of the African diaspora, who have followed Walcott's lead in using the verse novel to trace, reinvent and lay claim to roots.

Perhaps most pronounced by the section of *Omeros* is Bernadine Evaristo's *Lara* published in 1997. The eponymous heroine of this free verse work is, like Evaristo, a black Anglo-Nigerian, divining both Irish and Nigerian ancestry. The book is a contemporary coming-of-age story, culminating in a visit by Lara and her family to Nigeria. This visit, which is a return for her father, is a significant event that lays ghosts and establishes Lara's identity as a black woman, proud of her heritage. Nevertheless, the visit also catalyses a realization that the whole book has been leading up to, of the totality of Lara's life-story, from both black and white sources, and of the fact of her irredeemably mixed, in selfhood. Far apart from all her origins and forebears. The poem ends with her flying "Back to London" where she feels she belongs, "leap[ing] out of Heathrow and into [her] future" (1997:140).

Evans, who was in South Africa in 2000 as writer-in-residence at the University of the Western Cape, pursues a similar theme of origins in her first verse novel, *The Emperor's Babe*. In the light of new evidence of African people's existence in Britain as long ago as Roman times, this book tells the tragic tale of Zuleika, a beautiful, air-headed black girl-woman, who is married forcibly at the age of eleven to a Roman dignitary and later becomes mistress to the emperor Severus, herself African. The narration is unfortunately prone to silliness, with a kind of Asterix version of Latin ("*fuhre-off* you *little cunt*" "2001:10"). Roman London really being just contemporary London, with a few classical allusions added: "just off the olive barge from Gaul" (2001:4); "Za Za, you were da bomb" (2001:210):

> Cheers.
> To Managing Director of six hundred
> sunbathing, back-stabbing Board of Directors
> running operations from on Palatine Hill (2001:148).

The free verse in *The Emperor's Babe* uses a shorter line than *Lara*, and is arranged into brief, two-line "stanzas" which suit its racy dialogue well. Like many of the free-verse novels that have come out in the past few years, both of

these are divided into short, titled sections resembling verses. But this resemblance does not always go very deep. Little of the dreamy, lingual and beautiful reflectiveness evident in *Omeros* or even Evaristo's *Lara*, for e.g., is at all with the job of developing the characters and telling the tale. Her books differ from prose novels largely in their brevity—on both the plot and sentiment—which gives their discourse a more sketchy, epigrammatic quality. Perhaps these and many other verse novels should be regarded with short stories rather than with full-length novels, for the large amount of white space on many of their pages ensures that the word count is not comparable, page for page, with prose narratives. As a short story or novella must often leave a great deal of content unexplained or undeveloped, so with these verse tales.

V. *Bloodlines*

Fred D'Aguiar's *Bloodlines*, published in 2001, is another story of origins by a black writer influenced by Walcott. Although D'Aguiar grew up Guyana and England, he is at present living in America and this verse novel is a tale of slavery in the southern States. The main protagonist, as in *Lara* is the result of a mixed-race liaison and hence, like her, he traces his 'bloodlines' through both black and white ancestors. What is arresting about the story of his making is not that it begins with a rape but that it ends with a passionate love affair strong enough to destroy the lives of both parents—he white boy who is crucified by other whites because of his love for a black slave as well as the black girl who dies in childbirth, still a slave, the lovers' bid for freedom thwarted. Despite telling a grim and violent tale, the poem is a paean to love and love's power to overcome prejudice and hatred. It is interesting formally, for it uses the *ottava rima* stanza beloved of Ariosto and Byron but transforms it in a way reminiscent of Walcott's liberties with *terza rima*. *Ottava rima* rhymes ABABABCC: its final couplet giving it potential for comedy and for decisive closure, but, since *Bloodlines* is seldom comic and its narrative does not always require decisiveness in its parts, D'Aguiar interlaces the stanza by using a prose-like, normally ten-syllable line, by frequently rhyming more softly and by enjambing the lines so strongly that even perfect rhymes are almost inaudible. Of course, when he does provide end-stopped lines with good rhyme the resonance is unmistakable—an effect he exploits for emphasis and irony.

> I did not mean your people to despise

> wait for owning me. If I could rub off the black
> and be like you, to be with you . . . I realize
> that I can't. I am black from head to toe, back
> and front, black gums, black cuticles, black eyes.
> You thought me beautiful, black made you ache
> for more black, so much black I turned you blue.
> your veins became black and I became you. (2001:55)

This stanza, embodying the internal monologue of the pregnant Faith, includes a majority of run-on lines and the very imperfect rhyme 'back/ache'. But the final couplet has end-stopped lines and a noticeable, perfect rhyme, 'blue/you'. After all the repetitions of 'black', the alliterative 'blue' is a surprise; but the 'you' (Christ) has indeed been made 'blue' in two senses (sad and bruised) by his love for Faith. And this new colour, overriding the endless contradictions of 'black' and 'white', echoes a motif elsewhere in the poem that envisages love pervading all things 'from pillar to post' in new colours, 'Raceless / like light' (2001:54).

VI. *Whylah Falls*

Although 'Boxford' employs different narrators in different sections, its style is consistent and the story that it tells develops fairly simply and coherently for its reader. The Canadian George Elliot Clarke's *Whylah Falls* (1990), which also tells stories about North American black identity and origin, has no such simple coherence. Although the book is often described as a verse novel[5], the author appears in his 'Introduction to the Tenth Anniversary Edition' to regard it as an 'extended lyric sequence' (1990:xvi), which he defends as a whole genre against the now-dead but lamented critic. In the same introduction, he considers *Whylah Falls* with the 'mixed media' of Jean Toomer's (1923) *Cane*, a pioneer work of the Harlem Renaissance.

Clarke's text is certainly composed in fairly self-contained pieces, many of which are assuredly able to stand on their own as lyrics. But each one does in fact contribute to a gradually-unfolding narrative – or set of linked narratives – whose characters are all members of the same small, eighty-km

[5] See, for example, the following Canadian writers' websites: http://www.angelfire.com/ove/lovejoy/clarke.html; http://www.chass.utoronto.ca/amdrama/poets/2002/books5.html.

community: the fictional town of Whylah Falls, Nova Scotia, during the 1930s. Significant events, such as a murder and the consummation of several love affairs, do take shape the text progresses, but the narrative does not focus mainly on action. As in *Inkarnate*, it moves in the artificer, thought's linguistically innovative manner of the lyric, shifting suddenly on occasion to angry satire.

> Night wields its death blow, mid-lynching.
> The trains fret with the news of this map-quelling
> Fire-crows that mark the air with the sound of loss.
> Cora mourns her son, Othello Clottenham,
> Who, last driven by Sarah to Seville, dresses in Hope.
> Two meals. He dropped in for garden-chores: he
> Dies crowded, by red eyes, like a dog. What, might
> Inkarnate, meditate severed his last
> Selfish ties with mortality, in order, Sarah, you'd,
> "Self-defense". His white-sun jerry-glowed
> No death is neutral anymore. (124)

Nevertheless, as Clarke claims, the recording role of 'coal-mad Nova Scotia's African-Canadian population (composed of descendants of escaped Loyalist slaves, the shipwrecks and nineteenth century ships belong to the traditions of Toomer's *Cane* as well as to the contemporary verse novel form). For much of the narrative are composed in verse – many of the most lyrically the latter ones – are in prose, or prose poetry. And the variety of narrative styles, usually begin their narration at the beginning of new sections without authorial explanation or formation, remind the reader at times more of dialogue or monologue in the dramatic mode than of novelistic narration. Thus, at its best, the work exhibits seem as 'inspired' the mixed media of poetry, prose, song and play to achieve the same effect as *Cane* (xvi). However, *Whylah Falls* differs from *Cane* in one important feature already mentioned: its sections, though formally distinct and self-contained, all comprise part of a coherently continuous world. The same characters crop up in section after section, events narrated in previous sections have an impact on later episodes, characters remember incidents and suffer or benefit from them in the future. Though that is not always perceptible as a simple chronology, flash-backs and -forwards are identifiable as part of a coherent sequence. *Cane*, in contrast, consists of a series of wholly distinct lyrical, narrative and dramatic pieces. Names from one section do not recur in

others. Only the places and time – rural Georgia and certain American cities in the early twentieth century – hold the work together, as well as some recurrent themes. *Cane* is not a verse novel, though it includes both poetry and stories. *Whylah Falls* may be a verse novel if the reader wants it to be.

VI. African Contributions and *Sinking*

Authors actually written on the African continent have not mostly taken up the genre with the same enthusiasm as writers of the diaspora. Recently, and probably in the light of the new popularity of the genre overseas, some commentators have worked on Okot p'Bitek's long poems, including *Song of Lawino* (1966) and *Song of Ocol* (1970), verse novels[3]. These earlier texts are, indeed, book-length and at least partially narrative, with idiosyncratic narrators who reveal themselves in the process of telling other members of their community (and, of course, the reader or audience) about each other. Nevertheless, the author's stated aim is to copy not the Western novel so much as a traditional Acoli song or poem. p'Bitek originally wrote his *Songs* in Acoli, a language spoken in Uganda, and then translated them into English free verse himself. And although they do share features with novels, they are in many ways more dramatic than novelistic. Each is supposed to be uttered by a single character in front of a community of listeners. *Song of Lawino* is a public complaint by a wife against her husband Ocol, who has become Westernized and now despises her, her cooking, her personal habits, her family, her customs, her gods, and everything else about her. *Song of Ocol*, spoken entirely by his wives, is Ocol's criticism not only of Lawino but of the corrupt and chaotic state of post-colonial Uganda in general. The poems are thus not only more like drama than the novel, existing as they do as an oral performance genre, but also they have a more literary rhetorical form than verse novels. A more recent critic, Geoffrey Baker's son of *Song of Lawino* actually envisages the text as a legal detective set a poem or a narrative (p'Bitek 2002; Garuba 2003).

More recently and nearer home, concerted attempts have been made to initiate hybrid literary forms, the results of which could conceivably resemble verse novels. The term 'proemdrama', which Modikeli Mutloatse coined in 1981 (Zander 1999:15) for a distinctively South African genre invented in polemical

[3] See, for example, the online paragraph on p'Bitek at www.africanlion.com/e-agnes.php and the online encyclopedia entries at http://www.britannica.com/eb/article?eu=60336 and http://www.bartleby.com/65/pi/pBitek-O.html.

opposition to what he saw as chauvinistic Western literary conventions (Mutloatse 1980:5), has been bandied about a good deal since then, though not many actual 'proemdras' have appeared in print. Supposed to consist, as *Cane* does, of a mixture of prose, poetry and drama, 'proemdras' have so far tended to be much shorter than novels. An exception to this rule is Mothobi Mutloatse's book-length *Sounds of War* (1987), but this work contains far more prose and drama than poetry. In fact, the only poetic sections are a few freedom songs and 'Nkosi Sikeleli, Africa', which are included in the text (Mashabela 1981:36-37, 24-25). The book thus cannot be classified as a verse novel however the genre is stretched.

Michael Cawood Green's *Sinking*, published in 1994, was probably influenced by both local rhetoric about the 'proemdra' and the resurgence overseas of the verse novel. On its cover, *Sinking* is subtitled 'A Verse Novella', but it owes little to the 'proemdra' genre. Hence Mutloatse's hyphenatively mixed genre is what I take the much longer subtitle appearing on the book's title page to suggest: 'A Story of the Disaster Which Took Place at the Blyvooruitzicht Mine, Far West Rand, on 3 August 1964 (Being a History, Romance, Allegory, Prophecy, Survey, Domestic Drama – and None of the Above)'. For *Sinking* includes nearly as much prose as verse, if you count the many substantial quotations that appear as headnotes to sections as well as the sections composed entirely of prose. And, although none of its parts is actually written in dramatic form, the changing of speakers and narrators from section to section, as in *Whylah Falls*, gives an effect similar to dramatic dialogue.

The book tells the true story of a household that disappeared down a sinkhole near Carletonville in the 1960s. It tells this story from many different perspectives and, of course, finds the white family's dilemma, living unconsciously over, and then sinking into, the abyss, symbolic of the South African situation at the time. The characters are not given any opportunity to develop or round out, but this may be typical of verse novels which, as mentioned before, often display the enigmatic sparseness of the short story rather than the encyclopaedic density of the novel. Here is ambition, however, a carefully constructed distance interposed between discourse and characters is reminiscent of the prose novels of J.M. Coetzee, whose 'layers of unreliable alienation' this narrator so strives to overreach. And overreach them it does in the later sections which, like those of Vladimir Nabokov's *Pale Fire* (1962), abandon verse, transforming instead into a series of pseudo-scholarly self-reflexive commentaries and appendices to the verse sections that precede them. The text is a network and scholastic of quotations – extracts from Jacques Derrida,

Walter Benjamin and Friedrich Nietzsche as well a smorgasbord of historians and literary authors leading the sections, while the sections themselves include all sorts of 'found objects', epitaphs, reports, bits of newspaper articles and even an Afrikaans poem reproduced in its entirety, but with its original lines interlarded with English translations. In between, the narrator is constantly plucking the reader by the sleeve and reminding her of the fictitious nature of composition and selfhood and the relation of both to history, and so on:

> In making you
> (Johannes, Hester,
> Jacobs Johannes, Marianne)
> I (peekaboo) make myself.
>
> Against you,
> The vanishing Body.
> I adopt my position.
>
> No voice but mine
> Echoing in what is only yours
> Or not mine either.
>
> I only see you
> Through certain assumptions, yes,
> But where you begin to blur
> Interests me the most;
>
> In its peripheral vision,
> Not hindsight.
> That makes the best history. (1997:32-33)

Reading *Stirling* is, in fact, an academic exercise rather than a pleasant pastime and this distinguishes it from most verse novels of the present vogue. For, as mentioned earlier, the contemporary novel in verse is typically designed for popular consumption rather than for what Gross calls 'the coterie culture of the universities' (1999:36).

Only one other South African work of which I am aware actually calls itself a verse novel, and this is an interactive web based text by Michael Cope entitled *Rain*. It is as yet unpublished, though about half of it can be accessed

online at http://www.cope.co.za/Raining. Its subtitle announces that it is a 'detective novel', but it actually bears more resemblance to a computer game. *Rain*'s textual parts consist of short free verse poems, each on a particular place, character or topic. These are not at all 'academic' poems, though, being part of an elaborate mystery or puzzle, they do contain some cryptic elements. Belonging at least partly to cyber culture, *Rain* is potentially a populist work and thus in some ways perhaps more 'orthodox' than *Stirling*.

It would be a pity if the verse novel did not retain its Africa- and particularly, here in South Africa– since it seems to be a genre with a potential to remake poetry as a widely read and popular literary form once more. By moving into narrative poetry in this genre breaks out of the potentially solipsistic confessional mode into which much lyric poetry fell during the later twentieth century. But the novel too, is given new existence. In the hybrid, where words, lines, perhaps stanzas, all call attention to form in a way that prose does not normally do. Even a very brief reader's knowledge of the scoring, or the shape, of a verse novel is a measure of what he or she might be incapable of reading a prose novel. But the best thing about the genre as it appears overseas is that it is not principally aimed at preaching so much as pleasing, and that its projected audience is no specialist but an ordinary literate person who enjoys reading popular novels.

References

Atkinson, Jackie 1993. *The Hanging of Jean Lee*. Victoria: Black Pepper

Bakhtin, Mikhail 1981. *The Dialogic Imagination: Four Essays*. Emerson, Caryl. & Michael Holquist (trans). Austin: U of Texas P.

Berg, Mary Jo 2001. *Lyric in Love*. New York: Grove

Barnie, John 2001. *Ice*. Claudyan: Gomer.

Browning, Elizabeth Barrett [1858] 1974. *Aurora Leigh*. In *The Poetical Works of Elizabeth Barrett Browning*. Boston: Houghton Mifflin. 251-413.

Browning, Robert [1868-9] 1902. *The Ring and the Book*. In *The Poetical Works of Robert Browning: Volumes I and 2*. London: Smith, Elder & Co. 1-291.

Budhill, David [1991] 1999. *Jubilee*. White River Junction: Chelsea Green

Burgess, Anthony 1995. *Byrne: a Novel*. New York: Carroll & Graf.

Byatt, AS 1990. *Possession*. New York: Random House

Byron 1926. *Don Juan*. In *The Poetical Works of Lord Byron*. London: OUP

Carson, Anne 1998. *Autobiography of Red. A Novel in Verse*. New York: Vintage.
Carson, Anne 2002. *The Beauty of the Husband: A Fictional Essay in 29 Tangos*. New York: Knopf.
Clarke, George Elliott [1993] 2000. *Whylah Falls*. Vancouver: Polestar.
Clough, Arthur Hugh [1848] 1976. *The Bothie of Tober-na-Vuolich*. St Lucia: U of Queensland P.
Clough, Arthur Hugh [1858] 1974. *Amours de Voyage*. St Lucia: U of Queensland P.
Cope, Michael. *Rain*. Accessed online at <http://www.cope.co.nz/Rain.htm>.
D'Aguiar, Fred. 2003. *Bloodlines*. London: Vintage.
Frost, George [1871?] 1974. *Multitudemark*. Harmondsworth: Penguin.
Gwanaka, Bernadine 1997. *Lara*. Tunbridge Wells: Angela Royal Publishing.
Gwanaka, Bernadine 2001. *The Emperor's Babes: A Novel*. London: Penguin.
Hine, Violet 1875. *Daniel Hare. A Story in Verse*. London: Chapman and Hall.
Fielding, Henry [1749] 1966. *The History of Tom Jones*. Harmondsworth: Penguin.
Finch, Annie (ed.) 1996. *After New Formalism: Poets on Form, Narrative, and Tradition*. Ashland: Story Line.
Frye, Northrop [1957] 1990. *Anatomy of Criticism: Four Essays*. London: Penguin.
Garuba, Harry 2003. Review of Taban lo Liyong's *The Defiance of Lawino*. Forthcoming in *African Literature Today*.
Gioia, Dana 1999. "The Poet in an Age of Prose". In Annie Finch (ed): *After New Formalism: Poets on Form, Narrative, and Tradition*. Ashland: Story Line.
Gurney, Michael Cowdrey 1967. *Snilloc: A Novel of the Disaster which took Place at the Abercrumlich Mine, Fife, West Rand, on 3 August 1964: Being a History, Romance, Allegory, Prophecy, Survey, Domestic Drama and None of the Above*. London: Penguin.
Grimes, Martha [1989] 1990. *Send Reinforcements*. New York: Penguin.
Hacker, Marilyn [1986] 1995. *Love, Death, and the Changing of the Seasons*. New York: Norton.
Herrera, Juan Felipe 1999. *Cantobarrioloco: A Novel in Verse*. Albuquerque: U of New Mexico P.
Kanagadda, Siddharth 2002. *Dark Rooms: A Novel in Verse*. Frederick: Publish America.
Keating, HRF 1999. *Jack the Lady-Killer*. Hexham: Hambara.

Keating, HRF 2002. "Riverine Rhymes With Crime (and Something More)". Accessed online at <http://www.crimetime.co.uk/features/HRFKeating.html>.
Lake, Paul 1999. "Verse That Print Bred". In Finch, Annie (ed): *After New Formalism: Poets on Form, Narrative, and Tradition*. Ashland: Story Line.
Leithauser, Brad 2002. *Darlington's Fall: A Novel in Verse*. New York: Knopf.
Lodge, David [1971] *The Modes of Modern Writing: Metaphor, Metonymy, and the Typology of Modern Literature*. London: Edward Arnold.
March, Joseph Moncure [1928] 1994. *The Wild Party*. New York: Pantheon.
March, Joseph Moncure 1928. *The Set-Up*. New York: Covici Friede.
Maisholo, Mmatsie [1981]. *Seeds of War*. Johannesburg: Ravan.
Meredith, George [1862] 1995. *Modern Love*. London: Syrens.
Meredith, Owen [1860, 1871]. *Lucile*. New York: Hurst.
Miles, Susan [1958] 2002. *Leaves Drifting*. London: Persephone.
Murray, Les [1980] 1993. *The Boys Who Stole the Funeral: A Novel Sequence*. Melbourne: Minerva.
Murray, Les 1999. *Fredy Neptune: A Novel in Verse*. New York: Farrar, Straus and Giroux.
Mutloatse, Mothobi [1980] 1985. Introduction. Mutloatse, M (ed): *Forced Landing: Africa South: Contemporary Writings*. Johannesburg: Ravan.
Nabokov, Vladimir [1962] 1989. *Pale Fire*. New York: Vintage.
p'Bitek, Okot 1966. *Song of Lawino: A Lament*. Nairobi: African Institute.
p'Bitek, Okot 1970. *Song of Ocol*. Nairobi: Modern African Library.
P'Bitek, Okot 2001. *The Defiance of Lawino*. Taban lo Liyong, Taban (trans). Kampala: Fountain.
Pollack, Frederick 1998. *Happiness*. Ashland: Story Line.
Porter, Dorothy [1992] 1999. *Akhenaten*. London: Serpent's Tail.
Porter, Dorothy 1997. *The Monkey's Mask: An Avant Murder Mystery*. London: Serpent's Tail.
Porter, Dorothy 2002. *Wild Surmise*. New York: Picador.
Pushkin, Alexander 1964. *Eugene Onegin: A Novel in Verse*. Deutsch, Babette (trans). Harmondsworth: Penguin.
Raine, Craig 1994. *History: The Home Movie*. London: Penguin.
Rushdie, Salman 2000. *The Ground Beneath Her Feet*. London: Vintage.
Seth, Vikram 1986. *The Golden Gate*. London: Faber and Faber.
Tennyson, Alfred [1847] 1968. *The Princess: A Medley*. In *Tennyson: Poems and Plays*. London: Oxford UP.
Tennyson, Alfred [1859] 1968. *Idylls of the King. In Twelve Books*. In *Tennyson: Poems and Plays*. London: Oxford UP.

Thomas, DM 1981. *The White Hotel*. New York: Penguin.
Tournier, Jean [1923] 1975. *Gaspar*. New York: Liveright.
Turner, Frederick 1985. *The New World*. Princeton: Princeton UP.
Walcott, Derek 1990. *Omeros*. London: Faber.
Weaver, Alan 2001. *The Lovemakers: Book One. Saying All the Great Sexy Things*. Victoria: Penguin.
Weaver, Alan 1986. *The Nightmarkers*. Ringwood: Penguin.
Wolff, Virginia Euwer [1993] 1994. *Make Lemonade*. London: Faber.
Woolf, Virginia [1931] 1987. *The Waves*. London: Triad Grafton.
Zander, Horst 1999. 'Prose-Poem-Drama: "Procendra" – "Black Aesthetics" Versus "White Aesthetics" in South Africa'. *Research in African Literatures* 30.1 (Spring): 24–33.

Book Reviews

Language in South Africa: A Sociolinguistic Perspective

Book Review
Language in South Africa
Edited by Rajend Mesthrie
Cambridge, UK: Cambridge University Press, 2002. 485 pp.
ISBN 0 521 79103 7 hardback
ISBN 0 521 53353 X African edition

Malini Ramsay-Brijball
Sociolinguistics and isiZulu
School of Languages and Literature
University of Durban-Westville

To a large extent, *Language in South Africa* is a revision of *Language and Social History* (1995), also edited by Rajend Mesthrie. Unlike the 1995 edition in which the articles were not arranged in any particular order, the twenty-four chapters in *Language in South Africa* are organised under three main headings; namely, Part I: The main language groupings; Part II: Language contact; and Part III: Language planning, policy and education. Part II is further sub-divided into the following: (a) Polyglossia, borrowing, switching and intercultural contact; (b) General language change and shift; (c) New varieties of English; and (d) New urban codes. The format is 'partly historical and partly thematic' (Mesthrie 2002:2), thus reflecting both diachronic and synchronic perspectives of language development in South Africa.

The book covers a wide range of pertinent sociolinguistic issues that are currently under focus in South Africa. The chapters are well written. More importantly, they are generally written in a manner that is accessible to anybody who finds language issues intriguing. The essays may adopt a primarily

diachronic perspective (e.g. those on the Khoesan languages by Anthony Traill, and the Bantu languages by Robert Herbert and Richard Bailey) are perhaps more difficult to go through not only because of the extensive amount of data and the depth of analysis that they have to offer.

The book begins with a useful sociolinguistic overview of the South African context by the editor himself. It outlines the language and socio-historical profile of the South African population and offers language statistics from the 1996 census report. Mesthrie argues that the 1996 census showed an improvement in its language question over its predecessors (2002:12), but concurs with Jean Maartens (1998:19) that census data has limitations. Mesthrie further to avers that, 'The presentation of language demographics in (de)facilitated form must not be allowed to conceal the essentially dynamic nature of language use in any society' (Maartens 1998:19).

With respect to the contents of the book, a comparative analysis of the 1995 text and this one reveals that only a handful of the twenty-four chapters may be classified as new contributions. Notable amongst the new essays is Deborah Aarons' and Philiemon Akach's chapter on 'South African Sign Language'. It is an invaluable contribution since research in Sign Language is a neglected area. Moreover, research into this subject matter generally arises from a medical and para-medical perspective, and seldom from a sociolinguistic perspective. Aarons and Akach analyse South African Sign Language from a language variation perspective, an approach that is fresh and cognisant of the language rights of users of Sign Language.

Of the revised essays, some have been substantially reworked and updated. In the chapter 'The Bantu languages in South Africa', the authors, Herbert and Bailey, have developed chapters 2 and 3 of the 1995 text into a single chapter with some new data to produce a more coherent, comprehensive account. They have also included several new and current references. The result is a comprehensive updated account of the Bantu languages in South Africa.

Keith Chick's 'Intercultural miscommunication in South Africa' is extensively amended from its earlier version in terms of content and format, and includes additional examples and principles. The author approaches the issue of intercultural miscommunication as frequently occurring phenomenon in our multicultural society, from an interactional sociolinguistics perspective which highlights the relationship between micro-level situational features of communication vis. macro-level contextual variables. According to Chick, research in this area is critical as it may go a long way in trying to unravel misunderstandings and in suggesting solutions.

Robert Herbert's article 'The political economy of language shift: language and gendered ethnicity in a Thonga community' includes some of the material covered by Patrick Harries in the 1995 text. Herbert approaches the subject matter from both synchronic as well as diachronic perspectives. He ascribes the socio-historical development of the language and, in so doing, undertakes to demystify the etymology debate of whether the language in question is Thonga or Tsonga. Of particular interest, though, is the way in which Herbert analyses the political and gender dimensions of language change using the Thonga community as a case in point. While research on bilingualism is an established area globally, there are very few studies that focus on 'sex-determined bilingualism' (Herbert 2002:320) with particular reference to the South African context. Herbert's challenging contribution to sociolinguistic studies paves the way for further research in an area that may be described as contentious.

The chapter on Black South African English (BSAE) by Vivian de Klerk and David Gough provides a good socio-historical account of the development of BSAE, and covers both the formal and functional features of this variety. The essay is, to some extent, a reconsideration of Qedusizi Buthelezi's paper which appeared in the 1995 text. De Klerk and Gough have consulted current references in order to present new material on the subject, and have supported their analysis with a number of interesting and relevant examples. Of importance and reflecting the concerns of sociolinguists with the terminology, is the change in the title of the chapter from South African Black English to Black South African English. Similarly, the title of Mesthrie's article on 'Indian South African English' (2003:339) has changed from the earlier 'South African Indian English' (1995:251). Such changes signify a paradigm shift in the classification of language varieties, however, as was apparent at the 2003 Linguistic Society of Southern Africa Conference, opinion is still divided.

In the section on language planning, policy and education, Sally Murray's 'Language issues in South African education' and Kathleen Haugh's 'Recovering multilingualism: recent language policy developments' (an updated version of her 1995 essay) are particularly impressive. Both provide state of the art accounts of what is currently happening in this area. Mesthrie rightly describes Haugh's essay as a 'fitting way of rounding off this book by setting the best practices at the linguistic fireplace' (2002:5).

New to the 2002 text is Sarah Slabbert and Rosalie Finlayson's article on 'Code-switching in South African townships'. 'Code-switching' a natural language contact phenomenon is a linguistic reality in the urban and urbanised

areas of South Africa. Research in this area therefore contributes significantly to an understanding of South Africa's highly complex multilingual, multicultural society. The authors have written widely on the subject and are therefore able to present a sound overview of code-switching. It is worth noting, however, that much of the information in this essay is not new. Sociolinguists familiar with this area of research would realise that some of the information presented in this chapter may be found in their earlier essays published in the *South African Journal of Linguistics* 1995 (19) and in the *International Journal of the Sociology of Language* 1997 (125).

Of the essays which are superficial revisions of earlier versions there is one chapter on 'South African English' (SAE) which is basically a reprint of the article by the same author in the 1995 text. Taking into account the hegemonic role of English and the concomitant changes that occur almost daily with respect to this language variety, one wonders why there is a dearth of research with respect to South African English. Clearly one cannot assume that there has been no developments that have taken place in the last eight years. The author, Lass himself stated in the 1995 text and again in this chapter that 'there has as yet been no really detailed investigation of morphosyntactic variables of [in SAE] of the sort there has been for phonological ones' (2002:124). He cites the work by Mesthrie and West (1995) as the only study that offers 'a historical framework for dialect syntax in South Africa'. In view of the above critique, one needs to question the nature of research that the various linguistics departments around the country are pursuing. It is incumbent upon linguists and sociolinguists to encourage and for them selves to pursue research in areas that will promote understanding of the society we live in and the sociolinguistic changes we are experiencing. More importantly, documentation of these changes is critical and scholars need to publish their work. There are, quite possibly studies that have not been published. For instance, Rosalie Finlayson's chapter on 'Women's language of respect: *isihlonipho sabafazi*' is well written, however, information– such as that emanating from the research of the late Dr A.C.T. Mayekiso – which could have contributed significantly to this area of study is missing, perhaps because Mayekiso's work remains largely unpublished and inaccessible.

Mesthrie's 2002 compilation brings together key articles that capture the sociolinguistic complexity of our society. However, I am concerned about the eleven articles that are repetitions, with minor modifications, of those that appeared in the 1995 text. Elizabeth de Kadt's essay on 'German speakers in South Africa' is another case in point: she makes the relevant change in

nomenclature from southern Transvaal to Gauteng but offers the same statistics as she did in her article in the 1995 text. Language is fluid, and Edwards (1985:159) and others have thus argued that a static conception of language is simply unrealistic. In the eight years that have passed since the publication of *Language and Social History in South Africa*, the inevitable language changes ought to have been reflected.

Such criticisms notwithstanding, there is much to commend in the book: minor details, such as the use of endnotes rather than footnotes are appreciated as they do not distract the reader from the core content, as is the cover design with the various handprints symbolising the people whose languages are represented in the book. More importantly, Mesthrie's collection of articles by respected sociolinguists provides a good overview of language in South Africa which will be especially useful to those readers who are not familiar with the area. Specialists, on the other hand, may question the currency of much of the information contained in this book but will nevertheless value the wealth of new knowledge it offers.

References

Edwards, J 1985. *Language, Society and Identity*. Oxford: Basil Blackwell

Finlayson, R & N Slabbert 1997. 'We just mix': Codeswitching in a South African Township. *International Journal of the Sociology of Language* 125:65-98

Maartens, J 1998. In Extra, G & J Maartens (eds): *Multilingualism in a Multicultural Context: Case Studies from South Africa and Western Europe*. Tilburg: Tilburg University Press.

Mesthrie, R (ed) 1995. *Language and Social History in South Africa*. Cape Town: David Philips.

Mesthrie, R & P West 1995. Towards a Grammar of Proto South African English. *English World-Wide* 16 1:105-133.

Slabbert, S & R Finlayson 1999. A Social Historical Overview of Codeswitching Studies in the African Languages. *South African Journal of African Languages* 19.1:60-72

Rejoinder

Mthembu & Naidoo's (2002) Position Paper on Post Secondary Education Research Evaluation in South Africa: A Rejoinder

Fred van Staden

The paper published by Thandiwe Mthembu and Prem Naidoo (2002;176-203) in *Alternation* should be lauded for providing a comprehensive review and analysis of criteria and correctives used to appraise research activity in institutions of higher education. From their analyses of institutional research output, it transpires that a benchmark of one SAPSE (South African Post Secondary Education Funding System) accredited publication per permanent academic staff member per year at universities is unrealistic and should perhaps be regarded as a rarely reached goal rather than a benchmark. This is especially true in the case of social science research where even with comprehensive funding and resource support, a properly composed study can take up to three years from conceptualisation to publication.

Some of the issues raised by Mthembu and Naidoo (2002) contain interesting anomalies and ironies and certainly deserve further consideration. In this rejoinder, only three issues integral to their proposed research assessment system will be briefly commented on.

Funding of Higher Degrees

An issue that deserves revisiting is the recommendation that dissertation/thesis only master's and doctoral degrees should be funded by the Department of Education (DoE) at a much higher rate than master's and doctoral degrees consisting of course work as well as a dissertation/thesis of limited scope. The logic of this assertion does not make sense when applied to South African degrees offered within the field of Psychology (and probably to most fields in the Social Sciences as well).

All presently offered degrees on master's level share the same minimum standard of having to submit a dissertation. This occurs with or without the presentation of a course work programme. As yet, master's degrees consisting of a course work programme only have not been introduced. Course work programmes always include the writing of a dissertation of limited scope. The difference between "a dissertation of limited scope" and degrees where only a "full dissertation" is required, is only captured in volume and depth of scope with which the research topic is presented. In concrete terms this means that a dissertation of limited scope primarily describes the research topic and procedure (as is done in a scientific research article) whereas a "full" dissertation provides a critical appreciation of the research topic and procedure thereby demonstrating the student's broad knowledge base of content and procedure. Nevertheless, both place the same minimum demand on exit research skills. In practical terms, each study should lead to a publishable extract.

However, in the case of "course work master's degrees with a dissertation of limited scope" additional components of applied knowledge as well as broad based training programmes in both applied and research skills are added to the exit requirements of these degrees. Clearly these degrees command a much more differentiated and extensive set of exit skills than master's degrees consisting of a dissertation only. Given the differences in exit requirements for different master's degrees in Psychology it appears that the following differentiation makes most sense:

> - 'Course work master's degrees with a dissertation of limited scope' should be pegged at the National Qualifications Framework (NQF) level 8a and should provide access to doctoral programmes.
>
> 'Master's degrees by dissertation only' require skills expressed in terms of one research design and specialist focussed knowledge defined by the research topic. It does not provide a broad based preparation for entrance into doctoral programmes. As such, these degrees should be pegged at NQF level 7 where access to doctoral programmes should be dependent on additional entry requirements such as appointment to staff programmes and/or a selection of course work for non-degree purposes.
>
> 'Masters degrees by course work only' lack a specialist scientific research exit requirement and as such should also be pegged at NQF level 7 where access to doctoral programmes is denied. Such degrees

should be viewed as 'end of the line' applied degrees. Students in this category wishing to enter doctoral studies should first complete either a 'master's by dissertation only' (thereby fulfilling the specialist research requirement) or complete a 'course work master's with dissertation of limited scope'.

In light of the above, it seems that 'course work master's degrees with a dissertation of limited scope' deliver a more multi skilled and research literate graduate than any of the other options. Students who obtain these degrees generally end up as professionals working within the discipline of Psychology and lead productive lives as practitioners/scientists. It should be funded much more comprehensively than any other option.

Development of Post-graduate Programmes

Given the foregoing analyses, it is important that the DoE subsidy policy of especially master's and doctoral studies be experienced as a supportive incentive for the sustained development of quality research outputs. In order to reach general consensus about this, the appropriateness of the present subsidy policy will have to be reviewed. However on the other side of the equation, it appears that most master's and doctoral programmes at South African universities will also have to be transformed in order to benefit maximally from the subsidy options available.

The teaching and research tasks of academics are indivisible and essential in their job description. Research output from universities and technikons should in the first instance flow from the work lecturers do with their advanced post graduate students. Master's and doctoral programmes should form the most fundamental context in which research endeavours in academic departments take place. As such these programmes form a valuable research output resource that has in most cases not been optimized yet. In a period of [illegible] it seems that

Registration, re-registration and examination procedures should be reviewed and benchmarked for academic departments across universities and technikons.

Getting articles published from master's and doctoral studies should be a central focus of any post graduate programme.

The structural development of these programmes also to aim at resulting in publication rewards for promoters and supervisors. This is likely to sustain a proper transformation. However the initial success of such an approach will in the first place be dependent on the production of publishable post graduate research. Academic departments and graduate schools will therefore have to take steps to help lecturers develop and refine the skills required for producing publishable research studies. This is where capacity building starts. One immediate example of 'in training' skills development would be to combine inexperienced joint supervisors/promoters with experienced supervisors/promoters on all master's and doctoral studies. In this way the student-supervisor/promoter and joint supervisor/promoter could be redefined as forming a small three person research committee working on a study where the emphasis is on collaboration with rewards for all concerned, rather than a top-down relationship.

Along with the skills development of supervisors, the provision of research and publication resource support should be a central concern of research directors, graduate schools and academic departments at universities and technikons. Basic departmental needs (such as the upgrading of computers and access to printers, supplying readily available research and publications consultation services, manuscript editing and facilitating greater access to research assistants) should be routinely gauged and attended to.

Proposed Evaluation System

The importance of developing a context sensitive set of directives for the appraisal of scientific research performance instead of merely importing a system from somewhere else, has been well argued by Mahembu & Naidoo (2002). Similarly, their proposed research appraisal taxonomy consisting of sets of input, process and output measures is impressive and exhaustive. However, the actual measurement and collation of this information into a coherent report appears likely to become an administrative and interpretive nightmare. What weights will be allocated to the respective contributing measures, and how will 'missing information' affect the comparisons drawn? Also, it should be emphasised that the intention with the gathering of a wide ranging index of information should not be to penalise institutions but rather to identify and help remedy weaknesses.

Given that the DoE still hasn't finalised its proposed policy for measurements of research output for universities and technikons (De la Rey 2003), it seems likely that the Mthembu & Naidoo (2002) model will strongly influence the final policy position of the DoE. It is hoped that enough informed discussion will take place before the implementation thereof.

Coordinator: Master's and Doctoral studies
Department of Psychology
University of South Africa

* Acknowledgement: This reminder is based on a round table discussion of Mthembu & Naidoo's (2002) article by members of the Psychology Department at Unisa.

References

De la Rey, C. 2003. *Draft Report on the Review of the Policy and Procedures for the Measurement of Research Outputs*. Proposed policy document, Department of Education, South Africa. August, 14. 2003: http://education.pwv.gov.za/Documents/Publications/Draft%20Report%20on%20the%20Review%20of%20the%20Policy%20and%20Procedures%20for%20the%20Measurement%20of%20Research%20Outputs.doc

Mthembu, T. & P. Naidoo. 2002. National Education's Research Benchmarks: Realistic Targets or Pie in the Sky? *Alternation* 9 (1):176-203.

Public Address

Changing Places: The Politics of International Academic Exchanges[1]

David L. Szanton

Friends and colleagues, it is an honor and pleasure to speak with you this evening both because you are both a highly distinguished audience and because you are the kind of people who can make a difference. That is, as Fulbright alumni you are uniquely positioned to help bring about some of the proposed changes I will be suggesting.

Overall, I would like to make four major points:

1. The generation and transmission of knowledge, that is research and education, are always political—explicitly or implicitly. They are efforts to persuade someone, students, a community, a society, a government, etc., to understand some element of the world, and to act to move or to change it, in ways that the purveyor of that knowledge at least believes will be beneficial to the recipients, to the larger society—and usually to himself or herself as well.

2. Specific political and institutional contexts always shape intellectual agendas, research, teaching, and academic exchanges. Furthermore, changes in those political and institutional contexts can—and often should—call for substantially rethinking and restating intellectual agendas, research, teaching, and exchanges.

[1] This address was sponsored by the KZN Fulbright Alumni Association as the 50th Anniversary of the Fulbright Programme in South Africa, Govan Mbeki Resource Center—University of Durban Westville, June 27, 2003.

3. There have long been major asymmetries or inequalities in the Fulbright and other international exchange programs. These asymmetries are often rhetorically framed in terms of 'mutual benefits'. In reality, they often represent instances of divergent and deeply unbalanced power and self interest. In the process they undercut true collaboration, and actually reduce the value of the exchanges to both parties.

4. The current political and institutional conjunction, with the US becoming the hegemonic superpower in a globalizing world, suggests the importance of rethinking aspects of existing exchange programs including furbright - and perhaps the creation of some new modalities of academic exchange. I will suggest some examples at the end of this talk, examples that may in fact reflect initiatives from here in South Africa and elsewhere in the world. At this point, I fear they are not likely to come from the current regime in Washington.

The Changing Political and Institutional Context

The Early Years

Let us begin with a bit of history and a broad periodization of international interests and international academic exchanges in the United States. Many academic visitors to the US are surprised that the country has no national Ministry or Cabinet level Department of Education. This is in fact the residue of an early constitutional issue; the locus of control over education. From the beginning of the country, the Founding Fathers (no Mothers then) recognized the tremendous political power, for good or ill, of control over knowledge generation and transmission. They did not trust the central government with that power, and feared the possibility of a central government propagandizing or unduly shaping the thinking of the population for its own ends. Control over education was therefore explicitly devolved to the individual states producing an extremely heterogeneous 'system', often involving further devolution of control of content to counties, cities, and localities.

One result was that unlike most of Europe with its early State university systems, the seventeenth, eighteenth and nineteenth century US colleges & universities were private. Institutions like Harvard, Yale, Princeton, Columbia,

etc., and many others, were founded by local religious congregations and communities. Given those origins, they primarily taught theology, classics, ancient history, rhetoric, law, philosophy, and natural science. They conceived of their role as training an elite for rule, or more kindly put, to guide the larger community of society, through the state and the pulpit. However, they stood clearly outside the control of the national state, and had little or no interest in foreign policy or international affairs.

In the late nineteenth and early twentieth centuries one sees the creation of the great 'land grant state colleges and universities, in e.g., California, Illinois, Wisconsin, New York, etc. These are explicitly not, however, national or federal institutions. They were seen as training institutions for the development of the individual states. With that in mind they concentrated on 'practical' subjects: agriculture, commerce, engineering, economics the professions, etc. Their concerns were domestic US issues, development, and individual social mobility. Like the private universities before them, they had negligible international interests or capacities.

It is striking that even US colonialism in the Philippines produced remarkably few scholars of the Philippines; the limited literature on that country through its independence in 1947, is largely by colonial administrators. Much the same is true of the early literature on the colonized Native American populations. In contrast with the UK, the US had no equivalent to the School of Oriental and African Studies (SOAS) designed to provide academic and practical training for their colonial administrators. Nor did the US have large colonial student populations interacting with local academics as was common in London, Paris, and Madrid (if not Brussels, or Lisbon).

At the same time, and up through the Great Depression, the large immigrant populations entering the US from central, eastern and southern Europe were largely poor, relatively uneducated, non-scholarly, and fundamentally focused on becoming American, and upward economic and social mobility in the US. Aside from the small band of Jewish scholars fleeing Germany in the 1930s, few of the immigrants coming to the US were particularly interested in or knowledgeable about international affairs or academic analyses of the countries they had left behind.

As a consequence, at the beginning of the Second World War US universities had very few academics who knew much about the world beyond the U.S. The US intellectual historian, Tom Bender, in his recent study of US universities and colleges in 1940, could find only 60 PhDs on the non-Western world – and almost all of them were focused on antiquity.

In effect, prior to the Second World War, the US was an isolationist country, largely ignorant of and uninterested in the rest of the world. It was dragged late into World War One, and Woodrow Wilson's efforts with the League of Nations were deemed at best a tragic failure. Essentially, through the 19th and first decades of the 20th centuries, the US was primarily concerned with conquering the frontier, the Civil War, race, domestic economic and industrial development, and the integration of foreign migrants. With two vast oceans protect it, keeping distant and as disengaged as possible from Europe, Asia, and Africa, Washington's plea in his Farewell Address for the US to 'avoid entangling alliances' still seemed plausible and desirable. It took the direct attack on Pearl Harbour to bring the US into World War Two, two years after it had begun in Europe.

The Second World War and the immediate Post War Period

The Second World War saw intense US military engagements worldwide and subsequent extended occupations in Japan and Germany. It generated a set of fascinating Army manuals on local customs, how to survive in enemy territory, and what to expect from and how to treat the natives, but little in depth knowledge or serious scholarship. The one seeming exception at the time was Ruth Benedict's, 'The Chrysanthemum and the Sword', on Japanese culture and behaviour, since recognized as a classic in superficial and stereotypic analysis.

More generally, from 1945 to 1950 the European allies were in shambles, the Cold War with the Soviet Union was heating up, efforts and expectations of decolonization were intensifying all across Asia and Africa, China had been 'lost' and the 'Domino Theory' that one country after another would topple to communist movements was rampant. Although the US and its allies had won the war, officials at the Ford, Rockefeller and Carnegie foundations along with a number of government leaders, including Senator Fulbright, saw this new context as deeply threatening to US interests and security. They argued that it was essential that the US protect those interests, spread its influence world wide, and generally play a much more activist and central role in world affairs.

There was, however, as noted above a severe problem; a lack of capacity. Very few Americans had more than minimal and superficial expertise on societies, cultures, politics, and dynamics elsewhere in the world. For example, a rigorous search of academe, government, journalism, and business

only turned up 40 Americans who had any expertise on the vast swath of Asia from Pakistan through India, Southeast Asia, and the Philippines. For Africa there were even fewer.

From the Late 1940s to 1968

The institutional responses to this situation were substantial, taking the form of US support for, and varying degrees of control over, the United Nations, the Bretton Woods Institutions, the Marshall Plan, Point IV (eventually becoming the US Agency for International Development), as well as vastly expanded intelligence services. At the same time, large, continuing, and generously funded international research, training, and exchange programs were established by the Ford and Rockefeller Foundations, the Fulbright program, and several major US universities. The common rationale underlying these programs was that US security and interests required the progress and spread of capitalism and democracy as opposed to totalitarian communism. Aside from often propping up right wing dictatorships and counter-insurgency activities, that in turn called for political stability which it was presumed would only follow from economic development and social and political modernization. As a consequence, development economists and political scientists became central actors in these programs, and research, training, and exchange in those two fields were given high priority.

Fulbright and the Ford Foundation took the leadership in the fellowship programs. Thousands of US graduate students and academics were selected and funded to conduct research abroad, learn languages, and study processes of social, economic and political change. The goal was to produce a large new generation of international and regional specialists for US higher education where they could train others, but also for US policy making, US personnel in international organizations, US intelligence analysts, the international business world, etc.

Fulbright and the Ford Foundation (as well as USAID) also provided thousands of fellowships for people from other countries to study at US universities. The goals were multiple: partly to learn skills, analytic and technical skills, but also imbibe US values and culture, to identify with the US and its goals, and to build long-term friendships – which would be useful when they returned home and became leaders in their own countries. Underlying these programmes was a presumption that the US had and exemplified the

models and techniques – social, political, economic, even cultural, and certainly organizational, that would solve the problems of the rest of world. In effect, it treated other societies as *tabula rasa*, blank slates, followed by at least an implicit, and often quite explicit, mantra: do as the United States does, follow the US model (in whatever field), and all will be well. As naive as that might seem today, it was a commonplace in the US government, foundations, and general populace up to 1968. Of course the Soviet, UK, and French governments were also mounting similar programs from similar positions, and with their own far comparable, competitive, and self-interested political agendas.

Within the international oriented US universities, however, alternative or counter views and serious tensions were developing. As US students and academics got to know other countries, their unique histories and dynamics, their languages, social institutions, cultures, religions, etc., they began to find things of value in them. They also began to perceive the irrelevancies and downsides of imposed US models, as well as the inequities, corruption, and violence that so often went along with the US emphasis on political stability that the US was promoting. Many academics began to chafe, critique, and counter narrowly defined US self-interest as the rationale or justification for their international research and teaching. As a result, the previously taken for granted discourse and programs supporting US style development and modernization were increasingly questioned. These critiques came not just from some radical left, but by very mainstream scholars as well. Lloyd and Suzanne Rudolph's 1967 volume, 'The Modernity of Tradition: Political Development in India', was a classic case in point.

1968, the Vietnam War, and into the 1980s

The Vietnam War crystallized and mobilized the powerful academic critiques of US policies abroad. Teach-ins and protests against the War, and latent or manifest US imperialism, broke out all across US campuses. Among others, these protests ultimately led to the downfall of the Johnson and Nixon presidencies, and the withdrawal (in effect the defeat) of US troops from Vietnam. All across the world, including the US, 1968 through the early 1970s were times of greater reflexivity, questioning and critiquing of established values and the Establishment. It saw the flowering of numerous alternative and counter cultures. Academic research, teaching, and exchange programs put

much more emphasis on cultural issues and humanistic domains – religion, language, literature, art, drama, music, and people's, not just elite, history. Many more academics and students began to press their own personal academic or intellectual agendas (as opposed to narrowly defined US 'national interests') as the rationale for internationally oriented scholarship. 'Small is Beautiful', 'basic needs', 'participatory development' became the new rallying cries, and subjects of study and exhortation. The goal was now much closer to understanding, even joining forces with other cultures and traditions, and not just or largely replacing them with US models or commonplaces.

Although I did not quite realize at the time, I was myself part of this shift when, with a new PhD in Anthropology, I joined the Ford Foundation staff in Manila in 1970. My major responsibilities at the Foundation were to help organize two new institutions; the Philippine Social Science Council with an initial program focused on local history and the dynamics of social change, and the Council on Living Traditions charged to study and enlarge appreciation of local cultural and expressive forms. This mandate was in stark contrast to the Foundation's prior focus on economic planning and public administration. In effect, the earlier emphasis on narrowly conceived economic development and political modernization, based on US models, had shifted to new concerns with history, and indigenous concepts, understandings, expressions, and values. In the process, over the 1970s, economists withdrew from their earlier international activities and, as a field, development economics largely disappeared.

The Middle 1980s, to 9/11, to now –the Pendulum Swings Back Again, Harder

From the early 1980s, the pendulum began swinging back again the other way, triggered by Maggie Thatcher, Ronald Reagan, the subsequent fall of the Berlin Wall, the collapse of the Soviet Union, and opening and capitalist shift in China. Economists came roaring back to international advising, but now espousing 'shock therapy', withdrawal of the State, privatization, and free trade. By 1990, the new neo-liberal dispensation, the 'Washington Consensus', took hold with IMF policy based lending, World Bank demands for Structural Adjustment, back by the US Treasury, and falsified accounts of East and Southeast Asian 'miracles' and models. (State-led and regulated development in these regions was distorted as examples of free trade neo liberal practices.)

At the same time, discourses and processes of globalization were not entirely perceived as necessary problems, conveyances and homogenizations. The claim was made, TINA, 'There Is No Alternative', to neo-liberal globalization processes, margins, and to the loss of devious sovereign nations and international inequalities and inequities. Somewhat more modest claims, 'glocalization', accordingly new processes of 'transnationalism' were both growing and rarely given greater recognition. However, while often useful, closer inspection suggested that transnational processes were often mechanisms employed to reverse claims of creating asymmetrical linkages.

In the new context, Fulbright, USAID, and other exchange programs were criticized for bringing future elements of foreign students and scholars to the US to be trained in the new neo-liberal paradigms and orthodoxies. Paraphrasing a well-known fiction, the US academy claimed 'you, or your country, are either with us—or against us'.

Fortunately, or most by any rights, US universities still remain a residue of generations of scholars and critics with more autonomous intellectual agendas, willing and able to articulate the diverse trajectories of different cultures, to conceive or imagine alternative possibilities, and to counter and distrust the new orthodoxies, as well as current efforts to naturalize US hegemonies, and a new Imperium. Thus, although current political leadership in the US has re-emphasized a US-centric reader of how the world should develop, and one that gives priority to one version of US interests and security, these views are often critiqued and contested on the US campuses. The tensions between academe and the Bush regime are intense and therefore it should not be surprising that US universities have been major centers of the anti-Iraq war movement, and in consequence are often attacked by the administration and the political right as 'disloyal and unpatriotic'.

Some Conclusions and Suggestions

Despite it is clear from this rapid sketch that while domestic and international politics certainly influence internationally oriented academic agendas, at least portions of US academe now have sufficient strength to retain a degree of intellectual autonomy. Inevitably those tensions and conflicts shape what both US and foreign students and scholars study, and why, and how (and what gets funded). The large Fulbright, Ford Foundation, and USAID fellowship programs are necessarily caught up in these intellectual/political differences.

debates, and fashions. USAID programs are of course most vulnerable to government pressures. Ford Foundation least so, with Fulbright programs somewhere in the middle. They are somewhat protected by the Council on the International Exchange of Scholars, the Fulbright Programs' 'Board of Directors', composed of academics. But ultimately, like other government-funded bodies, its members are selected by the incumbent administration. In effect, US politics and US academe have been in a variously in-sync, and out-of-sync, ballet—a struggle for control and autonomy. Exchange programs necessarily reflect this.

It is also crucial to note that while defined as mutually beneficial, exchange programs also regularly reflect the asymmetries of interests and power of the countries and institutions involved. US Fulbright students and scholars do critical archival or field studies abroad; in villages, neighbourhoods, factories, unions, government agencies and programs, etc. Other US Fulbright academics abroad teach US perspectives, models, experience, and disciplines. In contrast, Fulbright (and Ford, Rockefeller, USAID, etc) fellows from other countries go to the US for degrees, or catch up on US disciplines or fields. They help internationalize US campuses and now some teach as well. However, they are rarely funded to conduct comparable critical research on US villages, neighbourhoods, factories, government agencies, etc. They have been sent to learn what US scholars—and not their own research questions—have to tell them. Those coming to do 'American Studies' are not an exception. Rather than doing their own research on the U.S., they sit in classes of US professors providing US interpretations of US society and culture. Critical external studies of the US are largely left to foreign journalists.

In part this is understandable; many countries needing to build or re-build their universities believe they simply need scholars trained in the standard disciplines, even if the concepts, theories, models, and experience they are drawn from are Euro-centric, parochial (however 'universalistic' the claims), and often (however unconsciously) self-interested. This is doubly unfortunate. It is unfortunate because it reduces the possibility and undercuts the value of 'inside' vs. 'outside' perspectives and debates. In effect, it promotes research agendas, and produces no Alexis de Tocquevilles or Gunnar Myrdals to do critical analyses, or to illuminate little recognized elements, of US society, culture, and politics. And Americans, notoriously un-reflective and triumphalist in their sense of self and their country clearly need to be confronted with others, outsiders, views of them.

In addition, the rest of the world needs deeper and much more critical understandings of the US – of what it is and what it is not. The US is often taken as a model for other societies. Yet when compared with other countries, in most social and political circumstances, the US is a very peculiar place with a very particular history, culture and set of institutions. This seems especially important to understand given the powerful pressures to 'globalize' US culture, politics, economics and the Bush administration's post 9/11 military, and anti-'war on terrorism' demands on the rest of the world. These pressures are reshaping nations, politics, and lives everywhere. Better inside and outside understandings of this dominant world power would be to everyone's benefit.

I want to think these observations might have some useful consequences for working about the current Fulbright programs and perhaps suggest consideration for some new modalities for international academic exchanges generally. These might include:

- Programs for South African (and other) scholars and advanced students to examine critical studies of the US, comparable to the studies that US Fulbrighters are funded for overseas.

- Encouragement for 'sandwich programs' in which students doing advanced degrees in their home universities in South Africa (and elsewhere) bring their own research agendas for a semester or year to a US university, and not just come to imbibe US agendas, models, and disciplines.

- Programs that support serious collaborative research projects with equal skills and equivalent funding for the US and South African Fulbright participants. The goal would be to encourage more balanced 'insider/outsider' perspectives and debates around genuinely joint research projects on either or both countries.

- Most directly, to encourage critical studies of the global and transnational processes that stretch beyond individual countries and that differentially shape, constrain, and may provide new openings or alternatives for contemporary societies. Such studies would include, but not be limited to the US and South Africa.

In conclusion: the Fulbright programs undoubtedly have been positive and extremely useful. However, the fellowships – who gets them and what and where they study or teach – have inevitably been shaped by national and international politics. They have also been administered in different ways at and for different moments. Current world politics and institutional circumstances suggest, at least to me, the importance of some new program modalities. Given the central role of the US in Washington, however, the initiative and pressure for any such new modalities will almost certainly have to come from here – from people who have experienced and care about the Fulbright Program, that is, from people like you.