The Dialogical Brain

Contributions of Emotional Neurobiology to Understanding the Dialogical Self

Marc D. Lewis
UNIVERSITY OF TORONTO

ABSTRACT. Using theory and data from emotional neurobiology, I suggest a neurally realistic model of Hermans' dialogical self. The model is premised on Hermans' idea of voicing and its implications for motivation, action and subjectivity. Because states of motivated attention unify brain activity, coexisting I-positions are as problematic for neuroscience as they are for psychology. To overcome this problem, I postulate an internal *monologue* in which the familiar I-position is subserved by an attentional system in the orbitofrontal cortex, linked with nearby affective and premotor areas. This internal monologue is fueled by gist-like perceptual expectancies of another's response, and it perpetuates and adjusts itself by updating these expectancies. A second I-position may be underpinned by an attentional system in the anterior cingulate cortex and its connections. These two attentional systems are partly independent, and they compete for control based on changes in emotional content and intensity. Thus, switching activation between them may account for semi-autonomous, but not coexisting, I-positions.

KEY WORDS: brain, emotion, expectancy, internal dialogue, subjectivity

Mary is alone in the kitchen, cooking for tonight's guests. She goes about the routine tasks of chopping and mixing, with little awareness of anything except a vague pleasure in familiar actions and anticipation of a tasty outcome. Then she notices that she put the stove on too high, and the rice will soon be ruined. There is not enough time to start the rice again if all is to be ready when the guests arrive. She notices a change in her mood, a loss of pleasure and increase in anxiety, but this is expectable under the circumstances. Then, as she rushes around trying to repair the situation, she decides to stop and 'look inward', as the yoga teacher suggests, and notices that there are phrases in her mind: 'What do you expect . . . I can't do everything on my own. It's not my fault.' And she notices that these phrases are directed at someone, but she isn't sure who it is. It seems a bit like her mother, who is about to arrive, or her husband, who does not appreciate how

THEORY & PSYCHOLOGY Copyright © 2002 SAGE Publications. VOL. 12(2): 175–190
[0959-3543(200204)12:2;175–190;022628]

difficult it is to cook a perfect meal. But the person she is addressing is apparently much closer, because that person now responds sarcastically, 'Oh, and whose fault is it?'

How can the self be one yet many? This question has preoccupied philosophers and writers for centuries. It is a question that bridges worlds as different as Buddhist meditation, psychoanalysis and cognitive science (Varela, Thompson, & Rosch, 1991), it underlies our fears about personality changes due to strokes and pharmacology, and it guides theories of psychodiagnostics and psychotherapy that try to make sense of self-induced suffering. Recently, social and personality psychologists have attempted to address the multiplicity of the self by replacing terms such as ego and superego with the more contemporary language of self schemas and narrative structure. However, these conceptualizations essentially exchange one set of metaphors for another, and they propose static mechanisms for explaining what appears to be a very active process (Hermans, 1996). To move beyond this impasse, Hermans models the multiplicity of the self in terms of voiced positions engaged in dialogue. This formulation is clearly action-based, and it captures some of the phenomenology revealed by clinical reports. But is it correct? Does it point toward tangible psychological mechanisms, or merely provide another set of metaphors?

Psychologists are increasingly looking toward the brain in order to ground their modeling in biological reality, and the discipline of cognitive neuroscience has compiled a great deal of data to assist them. Using these data, I suggest a neuroscientific instantiation of Hermans' model of the dialogical self in order to evaluate its core tenets, increase its precision and help extract it from its remaining metaphors. Following Hermans, I see the agentic notion of voicing as an important refinement, and I use it as a bridge to the brain. I discuss the brain's premotor systems and the attentional states that guide them, and I show how these and related systems are rooted in emotion and anticipation. I then speculate on how these systems might be engaged in producing internalized voices and responding to those voices as if they came from someone else. Finally, I address the claim that different voices or positions coexist, occupying the same mental space at the same time. This portrayal of multiplicity creates problems for neurobiology just as it does for psychology. Yet they are creative problems that point toward a fundamentally new perspective.

Voicing in the Dialogical Self: Embodiment, Agency and Position

In order to move beyond the metaphors of previous accounts, including the modern metaphors of information processing and narrative, Hermans (1996) highlights the active voice in the multiple self. The 'agentlike qualities of the

I' (p. 42) give each of the self's various positions a vital, active part to play in the internal dialogue. In fact, *positioning* assumes the status of a verb to convey this activity, and this allows Hermans to break away from the more passive constructs of schema, script and story-line. Hermans frames the dialogical self as an interaction between various voices. These I-positions take turns in an internal dialogue, like interacting characters in a story. They agree or disagree and tell stories from their own perspectives. Moreover, they have the capacity to change or evolve by taking into consideration the perspective of the other. This means that they must hear as well as speak, and indeed the concept of voice is meaningless unless it is related to the perception of other voices. For Hermans, 'voice assumes an embodied actor located in space together with other actors' (p. 44).

Hermans' focus on embodiment places him in league with contemporary cognitive scientists for whom information processing is the most recent in a line of inadequate metaphors. The information-processing account, which has been the standard language of cognitive psychology for years, compares cognition to the sequential analysis of symbols by a digital computer. Despite its enormous contribution, this metaphor has lost utility because it cannot be squared with subjectivity, intentionality and motivation, qualities that set humans apart from machines. In its place, theorists have proposed cognitive processes that are fundamentally embodied. First, they are viewed as emergent, self-organizing, global gestalts that arise from reciprocal interactions among processing units, not linear sequences (Clark, 1996; Varela et al., 1991). Second, they are necessarily affective or motivated, linking biological requirements to the formation and manipulation of societal meanings (Fogel, 1993; Freeman, 1995; Lewis, 1995). Indeed, Hermans and Kempen (1993) provide a central place for emotion in their model, and Hermans and Hermans-Jansen (in press) propose that changes in voiced positions can be explained as phase transitions in non-linear dynamic systems. Thus, their embodied account of the multiple self fits with an emerging zeitgeist in cognitive science.

In cognitive science at large, the move toward embodiment includes a commitment to understand the brain as the basis of cognition. The richly distributed, reciprocally interactive and self-organizing character of neural activity provides a radical alternative to the linear sequences of symbol-processing machines (e.g. Thelen & Smith, 1994; Varela et al., 1991). It follows that Hermans' move toward embodiment should be compatible with neural realism. Moreover, his emphasis on voice provides a useful entry point to test this compatibility. Voicing, construed as action, points toward the brain regions and subsystems directly involved in planning and generating voluntary speech. Voicing, construed as listening, points toward the attentional systems that anticipate others' speech and prepare for one's own. Thus, a good place to look for dialogicity in the brain is in systems where attention and action are integrated. These are generally acknowledged to be

the frontal and prefrontal cortical systems, and they are held responsible for a focused sense of the self acting in the world. By studying the character of these systems, their dependence on emotion and their contribution to learning and memory, we can speculate on how a dialogical self might actually be housed in a dialogical brain.

One of the most difficult and intriguing question for the dialogical self is how one can be both subject and object in the same dialogue, as was the case with Mary. To address this question, this article is restricted to discussion of 'internal' dialogues, that is, dialogues between one's various I-positions (Hermans & Hermans-Jansen, in press). This exploration is guided by a contention of Hermans' that is particularly challenging from a neural perspective. Hermans (1996) emphasizes that the positions in the dialogical self are distributed in an imaginal space. This leads him to propose the 'simultaneous existence' (p. 46) of internalized voices. However, as we shall see, attention, action and motivation appear to synchronize the brain, such that numerous subsystems become highly coordinated in real time. It is difficult to imagine how semi-autonomous positions could coexist at the same moment in such a unitary brain, and this problem indeed recapitulates the classic difficulty of seeing the self as both unitary and multiple. Thus, what is problematic for a neural explanation is problematic for the study of the self more broadly: the reconciliation of unity and multiplicity. We now look to the brain to see how this challenge can be addressed.

Neuropsychological Mechanisms of Attention, Action and Emotion.

The first step to modeling the dialogical brain should be to refine, as much as possible, the psychological reality of multivoicedness. What is the actual subjective experience of internal dialogue? Most people who are not psychotic do not actually hear internal voices most of the time. Unless they conscientiously track their thoughts, as Mary did while cooking, they probably think that the idea of internal voices is bizarre. Yet, people often notice that a word or phrase was on the tip of the mental tongue, as if about to be spoken, and they may even notice their lips move at times. Less often, people report hearing phrases that sound like they come from somebody else, such as a parent or spouse: 'That was dumb!' or 'You're going to eat all that?' After meditation, Gestalt therapy or some other form of introspective learning, people may begin to notice that they are actually the author of these phrases, speaking to an objectified self from a different perspective. But if one attends closely, most of the phrases one hears are one's own. They emerge from a familiar I-position and there is no clear respondent to whom they are directed.

At first glance, this portrayal seems to contradict Hermans' (1996) notion

of multivoicedness, the 'process of dialogical movements in an imaginal space' (p. 44). However, one does not have to be actively speaking or actively listening to be in a dialogical relationship. Consider the following logic: if Mary normally proceeds in a familiar I-position, but then hears a self-directed comment from inside her own head, she should be shocked or baffled—unless, to at least some degree, the comment is expected. In other words, the experience of dialogicity may often be the experience of expecting dialogue, that is, the experience of acting *as if* someone might be listening to us, evaluating us and ready to react verbally. Even in the familiar I-position, there may be a kind of listening going on that presupposes the presence of an other. Such phenomenological and logical assumptions of course need to be supported by research. But if they are even close to being correct, they potentially clarify what self-reports and questionnaires miss: that internal dialogues are real, but they are usually sublingual and inchoate, and the voice of the other is a rare event, but it is not unexpected.

Based on these assumptions, a neural model of the dialogical self that is precise and comprehensive must be able to do at least three things. (1) It must be able to determine what position, either a familiar I-position or some other position, is doing the talking and the listening. In other words, who is the subject and who is the object at each point in the dialogue? (2) It should be able to specify the degree of articulation at which motor (speaking) and perceptual (hearing) events are taking place, on a spectrum from vague, gist-like sensations to articulated words or phrases. (3) It should be able to specify whether a dialogical act is happening in the moment or else expected in the near future, and it should be concerned with the relationship between the present and the anticipated future.

To think about (1) the subjectivity of self or other, it is helpful to utilize a canonical example of self talk. Following McAdams (1985), we can identify many versions of a familiar parent–child relationship being voiced in the internal dialogue. The self can switch from the child, who is subjectively hearing the parent's criticism or who subjectively argues back, to the parent, who criticizes or scorns the child (subjectively speaking from another position). This dialogue must have been the basis for the psychoanalytic notion of the superego (Freud) or internalized object (Melanie Klein). A more user-friendly version comes from developmentalists such as Dore (1989) and Kegan (1982), who observe children enacting highly familiar exchanges with a pretend parent, often switching between the child's and parent's role. While they are enacting the child, the child is subject and the parent object. When they are enacting the parent's role, the parent-self is subjective instead.

In order to think about (2) the articulation issue, it is useful to conceive of actions and perceptions as evolving from a gist-like sense of the world to fully elaborated behaviors and percepts. Much cognitive processing remains at the level of gist, where a supramodal image of an object, action or

situation lacks sensory details yet has emotional significance (Brainerd & Reyna, 1990). Actions begin with global, gist-like intentions that can be rapidly refined into specific motor plans and finally into a sequence of muscle movements (including speech). Perceptions begin with a global sense of the world; they can remain gist-like for some time but then shift to a rapid extraction of the specific features of a situation (e.g. the actual words in the case of dialogue). In fact, for both action and perception, most attentional activity (and neural activity) is engaged below the level of articulated words or sounds. If one were to attend to the details (the actual words) in dialogue, one would not be able to keep track of the global meanings of phrases and sentences. Thus, voices may be actively *intended* in speaking, or actively *attended to* in hearing, without any specification of words or phrases.

Finally, with regard to (3) expectancy, a key principle of cognitive psychology can be applied to internal dialogues. Action is always guided by anticipation, and anticipation always takes place in the context of perception (Neisser, 1978). To complete the cycle, actions change or reinforce aspects of the perceived world, setting the occasion for further anticipation and action. From a neural perspective, the link between perception and action-planning takes place at an executive level of *anticipatory attention and control* subserved by the frontal and prefrontal cortices (Fuster, 1996). Here higher-order processes of volitional attention arise from a gist-like integration across perceptual modalities, providing the foundation for emerging behavioral plans and actions. It has recently been established that perception is guided by anticipation of action as well. Freeman (1995) describes how *preafference* (or, so to speak, pre-reception) travels from incipient motor plans to the sensory cortices, focusing them on particular aspects of the world in anticipation of intended behavior. Putting these ideas together, we can see the brain as generating actions and perceiving events that fit an anticipated future, a model of the world just ahead of us in time. Like a marksman aiming at a moving target, the brain is concerned with what is about to happen, not what is happening now. In terms of dialogue, attention (hearing) would be guided by the expectation of what 'I' am about to say, and action (speaking) would be guided by the expectation of what 'you' are about to say.

As mentioned earlier, cognition in an embodied system cannot be independent of emotion, and the importance of emotion is central to the dialogical self. But how might the interplay between cognition and emotion guide neural modeling? Psychologists have become increasingly aware that negative emotions (e.g. anxiety, anger, sadness) narrow the focus of attention and anchor it to particular aspects of the world (Derryberry & Tucker, 1994; Mathews, 1990). In fact, the biological purpose of emotion is to motivate action, that is, to move attention toward aspects of situations that are highly important, and to urge the production of actions to deal with

them. However, the control of attention by emotion is reciprocally related to the control of emotion by attention (Lewis, 1995, 1996). Attentional states amplify, and then regulate, emotions by appraising their causes and generating plans to resolve them. These plans, though they may not be executed immediately, relieve immediate concerns by taking control of whatever needs to be done to face challenges and alleviate dangers.

How is this reciprocal relation between attention and emotion played out in the brain? Attention, and cognition more generally, is usually ascribed to the cerebral cortex, the outer sheath of the brain, and particularly to the prefrontal cortex, which synchronizes all sensory and motor regions according to a particular plan or goal (Fuster, 1996). Emotion, on the other hand, is mainly the province of the limbic system, an assortment of structures, including the amygdala, hippocampus and septum, that are enfolded within the cortical layers. The limbic system is considered directly responsible for affective feeling, emotional conditioning and action tendencies related to emotion (Buck, 1999; LeDoux, 1996; MacLean, 1993). Communication between these two macrosystems allows for the interplay between attentional and emotional states. Dense bundles of fibers carry signals in both directions between the prefrontal cortex and the limbic system, such that they activate each other simultaneously and reciprocally. Cortical attention regulates emotion through downstream paths, while emotional (limbic) activation feeds up to direct attentional focus. However, once the limbic system and prefrontal cortex recruit each other, they become entrained or coupled, producing a unitary and coherent attentional focus. This global state synchronizes activation throughout the remaining regions of the cortex and controls arousal and other functions housed in more primitive brain systems (Tucker, Derryberry, & Luu, 2000). Most important for the present discussion, this global attentional state, because it is fueled by emotion and directed toward action, is necessarily subjective. Thus, any I-position that is autonomous, intentional and, as Hermans would say, voiced must be subserved by an affectively grounded attentional state that synchronizes the entire brain.

Modeling the Dialogical Brain

We can locate the epicenter of internal dialogues in the network of circuits between frontocortical and limbic regions. These reciprocal connections are where global attentional states emerge in conjunction with emotion and anticipated action. I have characterized these global states as unitary gestalts. But this poses a problem: how can multiple, coexistent voices or positions be supported by a single, unitary attentional stance? Hermans' description of the 'simultaneous existence' of internalized voices appears contradictory in a brain that can only do one thing, at least one subjective thing, at a time, and

this fundamental paradox needs to be resolved in a plausible model of the dialogical brain. One solution would be to model internal dialogues as turn-taking, such that one global attentional state, belonging to one voice, alternates frequently with another. Yet the phenomenology of internal dialogues does not support this idea. To recapitulate, internal dialogues are normally sublingual, and one spends much more time in the familiar I-position than in any other state. Thus, one subjectively speaks one's own voice more readily than one subjectively speaks another's, and most of the time the (objective) voice of the other is expected rather than heard. Given these considerations, switching one's global, subjective stance back and forth in a turn-taking sequence is not very likely. What is more likely is that one remains in a continuous dialogical relation with an *anticipated*, *almost-heard* other, but does so from the subjective stance of the familiar I-perspective.

Let us begin to model Mary's experience in the kitchen using this framework. When she realizes that the rice will be spoiled, anxiety flickers quickly in the limbic system and begins to recruit attentional circuits in the (pre)frontal cortex. A global attentional state coheres, and anxiety is now directed toward an imagined reaction to the late meal (perhaps a critical remark or merely a look of disappointment). Now a gist-like sense of being 'in trouble' prevails as a corticolimbic gestalt, and plans for self-defense begin to arise in the premotor circuits extending backward (posteriorly) and upward (dorsally) through the frontal cortex. What is happening is that a subjective I-position is emerging spontaneously and generating behavioral strategies (i.e. verbal intentions) about what to say to the critical other. However, these speech-like plans need not be articulated in words. They may resonate in the premotor circuits without actually arriving at the motor cortex. Mary may have a general sense of being wronged, feeling grumpy and arguing or defending, without any words appearing on the screen of the mind. This anticipatory, inchoate, dialogical stance may last for a while. Then, perhaps spontaneously or perhaps because she stopped to examine it, this stance gives way to more articulated motor plans, resolving to actual words and phrases (e.g. 'It's not my fault') as the motor cortex is activated. These may not seem like an internal dialogue, because Mary only hears her own voice in passing, through feedback from motor circuits to the auditory cortex. Of particular importance, this *monological* state could go on almost indefinitely, and certainly without much awareness. It is not until Mary notices a second internal voice that a sense of dialogue emerges.

The internal monologue just described is surprisingly self-sustaining, and one way it sustains itself is by anticipating the other's response. Whether inchoate or articulated, the motor plans generated by anticipation of criticism are *in relation to* the imagined other. The construct of preafference suggests that these plans maintain the anticipation of the criticism, the 'reality' of the critic, as an ongoing, but abstract, perceptual gestalt. According to Alan

Fogel (1993), relationality is responsible for the emergence and stabilization of psychological states throughout the lifespan. In fact, Fogel proposes 'consensual frames', stable patterns of interpersonal interaction, as the unit of analysis for all psychological processes. These frames are stable because they lock together anticipation and response between two interacting partners. But the stability of interpersonal states also relies on emotion (Fogel et al., 1992; Lewis, 1995). Emotion keeps both partners committed to the exchange and prevents it from dissipating. In the case of internal monologues, the relational and emotional aspects of frames may serve the same stabilizing function. One imagines the critical stance of the other, but without resolution or detail, and this stance has a particular emotional significance. This emotional significance rivets attention and continues to propagate anticipatory action plans. But there is more to the story. The stance of the imagined other is not static. Rather, it continues (through updating of anticipation) to adjust itself to Mary's evolving action plans. Thus, Mary's monologue is self-perpetuating, not through static anticipation, but through a progression of updated action plans and anticipated responses. For example, Mary's monologue of argumentative self-defense is likely to elicit opposition from the imagined other, and the anticipation of that opposition continues to fuel her anxiety and motivate adjustments to her defensive posture. Here we can observe a full-fledged dialogical process, continuous, fluid and self-perpetuating, without ever leaving the familiar I-position.

But to model this internal monologue with precision, we need to go back to the brain for more detail. So far the (pre)frontal cortex has been portrayed as a single, undifferentiated, attentional system. However, as shown in Figure 1, the frontal/prefrontal cortex can be subdivided into at least two attentional systems, each with access to its own premotor network and each with different limbic connections. One attentional system of particular relevance to emotion is the *orbitofrontal* cortex. This region is at the very base of the frontal lobe, and it is tuned to rewards and punishments in the immediate environment (Depue & Collins, 1999), probably because of its dense connections with the amygdala (the limbic structure responsible for fear, anxiety and some kinds of anger). It is also closely connected to the temporal lobes, where perceptual (mostly auditory) input is processed and integrated, and to Broca's area, which controls speech production and reception. Finally, it is closely related to a premotor system called the arcurate premotor area (Goldberg, 1985). Together, these linked systems have a particular style of functioning, characterized as controlled, responsive and input-driven (Goldberg, 1985). A second attentional system is the *anterior cingulate cortex* (ACC), which is closer to the top and center of the head (more dorsomedial). This system broadly integrates information across spatial and temporal perceptual elements (Fuster, 1996), and it is associated with a second premotor system, the supplementary motor area, where

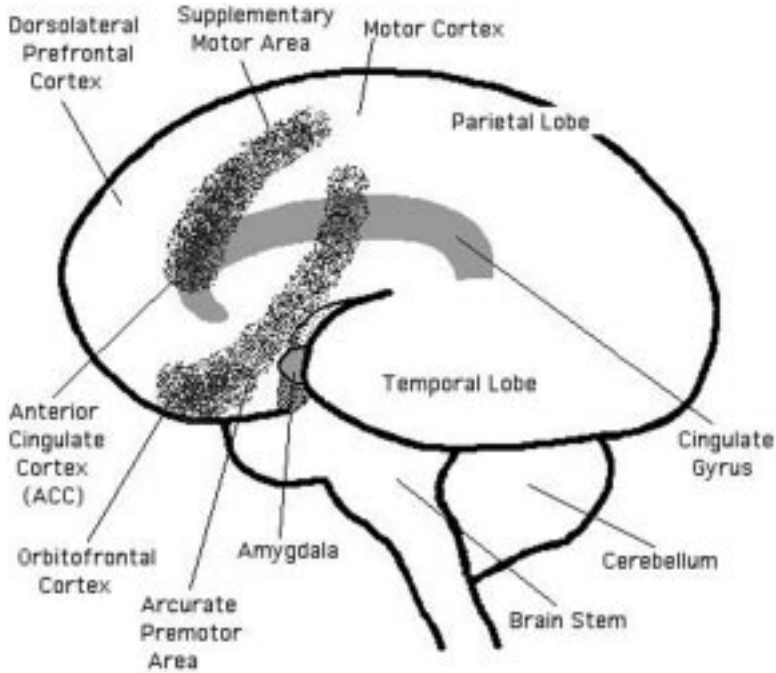

FIGURE 1. Drawing of a medial slice of the brain partially occluded by the temporal lobe and amygdala. Two frontal attentional systems, the orbitofrontal system and the anterior cingulate cortex, are shown along with their associated premotor areas, and their independent paths to the motor cortex are highlighted with 'spray paint'.

spontaneous volitional acts are generated. The outputs of these linked systems are characterized as coherent or packaged motor plans, projectional (or impulsive) rather than responsive in style (Goldberg, 1985; Luu, Tucker, & Derryberry, 1998). The ACC is also connected to the limbic system, but less closely to the amygdala and more closely to the septum and hippocampus. Its motivational base is concerned with moving through familiar situations efficiently rather than carefully confronting challenges and threats. In summary, the orbitofrontal system attends to potential rewards and threats, and it drives behavioral plans, including speech, via careful monitoring of perceptual feedback. The ACC system integrates more diverse information into familiar gestalts, is less 'hot' emotionally and drives behavior plans that are spontaneous, global and intact.

The internal monologues described so far involved as much listening as speaking, and they were fueled by anxiety about anticipated rewards and punishments in Mary's case. These features suggest mediation by orbitofrontal systems. However, there was no actual perceptual input from another

person; thus, the orbitofrontal cortex would have to create a gist-like imaginal figure while formulating its own voice. This should not be problematic for emotionally compelling monologues. According to Luu et al. (1998), the attentional system in this region has a short-term memory capacity sustained by the motivational significance of the recalled evaluations. It may also combine these evaluations into a categorial rather than specific sense of another person, explaining why Mary could not quite discern who was criticizing her. According to Schore (1994, 1997), the orbitofrontal cortex, particularly in the right hemisphere, produces an affectively charged, gist-like sense of an interpersonal respondent, based on expectancies from many past interactions. These interactions begin with attachment figures from infancy and childhood. As a result, this gist-like image is the fundamental arbitrator of emotion regulation, and it sets the rest of the brain in a mode of readiness based on preconscious expectations. An image of a warm, soothing parent permits rapid emotional equilibration, whereas the expectation of rejection or criticism promotes defense or withdrawal. From Schore's perspective, these attachment-based appraisals would be the basis of individual differences in dialogical styles, consistent with case histories presented by Hermans and Hermans-Jansen (in press). Thus, internal monologues mediated by the orbitofrontal cortex would be attuned to a highly predictable response from an imagined person or type of person, and this attunement would resonate with a particular emotional state maintaining a particular style of other-directed speech.

The internal monologue described so far does indeed have a dialogical character, based on the creation of an expectable other in relation to whom one voices one's own position. It also addresses the problem of coexisting positions in a unitary attentional frame, at least partly, because subjectivity and agency reside in a single, coherent I-position that continuously awaits the other's response: no bifurcation of attention is necessary. However, these monologues do not have the vitality of Hermans' polyphonic self. They repeat familiar stances, they are devoid of novelty and they miss any true exchange or confrontation between *autonomous* voiced positions. Let us move on, then, to the final scene of Mary's internal dialogue. After her grumbling monologue, she seemed to hear another voice with an entirely different tone, speaking sarcastically in fact, and saying 'Oh, and whose fault is it?' This implies a direct accusation coming at Mary from another. We know that the voice actually came from her. But does this imply a different I-position, as Hermans proposes? How would such a shift be modeled? I present the following speculative proposition as a direction for future work.

The position of the 'other' may bubble up from a gist-like expectancy to an actual memory of a distinct word or phrase. If this is the case, Mary might hear the other's words yet retain her familiar I-position. But this possibility is not very interesting, because it maintains the same positioning as the

monologue already described. It would be much more interesting, and more congenial with Hermans' polyphonic self, if Mary's subjectivity switched from the familiar I-position to an alternative, autonomous I-position. Like the child who switches roles from the helpless little girl to the powerful mother, perhaps scolding and even punishing her favorite doll, adults also notice occasions when they take themselves as an object and speak subjectively as someone else.

How can this switch of subjectivity be modeled? One possibility is that internal monologues progress through many shades of emotional content as well as gradations of emotional intensity, along with the shifting images of the anticipated other. These changes would be expected to alter the frame and focus of attention, as discussed earlier (e.g. Niedenthal & Kitayama, 1994). When anxiety gives way to sadness, for example, the narrow beam of attention broadens, and when it gives way to anger, that beam switches to an object that can be construed as an obstacle. Such changes in emotion and attention may perturb the coherence of an attentional stance, changing the sweep of perception and anticipation, and this may be the point at which subjectivity switches. In previous work, I have suggested that fluctuating negative emotional states can shift the focus of an appraisal radically without changing its content (Lewis & Junyk, 1997). For example, a blend of anxiety and rage felt toward a parent can shift to rage directed at the self, with only a small piece of the appraisal (i.e. its object) being replaced. Here I want to highlight the neurobiological fault-line that subserves such changes, and the most obvious candidate is the distinction between the two attentional systems described earlier, the orbitofrontal and anterior cingulate systems. Interestingly, the orbitofrontal cortex is one of the only frontal regions whose activity is independent of the ACC (Koski & Paus, 2000). Instead of working together, these two systems actually compete for the control of attention, depending on the type and intensity of emotion (Bush, Luu, & Posner, 2000; Drevets & Raichle, 1998). In the presence of negative emotion, and particularly anxiety, the orbitofrontal region shows heightened activation while the ACC does not; but in more 'cool' or 'cognitive' tasks, activation shifts to the ACC instead. While a scolding parental voice is not emotionally neutral, it is also not particularly anxious. In fact, it may be somewhat of a relief, emotionally, to become the perpetrator rather than the victim of rebuke. Thus, an emotional shift away from anxiety, perhaps to anger or contempt, may be what triggers activation of the ACC, providing the occasion for a shift in subjectivity.

What is most intriguing about this hypothesis is that the rapid, spontaneous and projectional style of motor plans stemming from the ACC seem to match the phenomenology of the alternate I-position. For example, the critical rebuke launched by the 'internal parent' is preformed, of a piece, it emerges without forethought and it is not the least bit attuned to perceptual feedback. In fact the child-self who bears the brunt of this rebuke is almost

entirely ignored by the scolding parent, a point that is frequently made by therapists who encourage their patients to 'be nice to yourself'. The sudden parental tirade pays no attention to any environmental feedback, in sharp contrast to the anxious, preoccupied and attentive style of Mary's familiar I-position. Instead, it is well practised and smooth. In fact, we could suppose that the familiar and intact parental rebuke Mary hears is actually a descendant of her own practising of her mother's role when she was a child. Modeling the scolding parent is a common exercise for children (Dore, 1989), it is coherent and automatic by the age of 3 or 4 and, arguably, it is emotionally cool and uncommitted. Finally, it is fascinating to compare the 'foreign' quality of an internal, parent-like rebuke to the 'alien hand syndrome', a sense that dorsally lesioned patients have that their well-practised actions are not their own. Luu et al. (1998) speculate that, even in normal subjects, a similar kind of dissociation may result from the projectional character of actions mediated by the dorsal ACC and related systems.

If this hypothesis is eventually borne out by research, it will have remarkable implications for internal dialogues. The switching of activation between independent attentional systems provides the neural basis for semi-autonomous I-positions of the sort that Hermans postulates. These positions are not active at the same time. From a neural perspective, subjectivity can only be in one place at a time. However, the capacity to switch subjective positions rapidly and completely does permit autonomous voices, engaged in something like a dialogical exchange, just as Hermans proposes. Furthermore, the switch from orbitofrontal to ACC activation permits a different premotor circuit to take control of motor output. As this motor output is released, it is also picked up perceptually (whether in the mind's ear or out loud). Now the brain finds itself confronted with new auditory information, and that information is bound to have an emotional impact. To actually hear a parental rebuke, regardless of its true author, is likely to trigger an immediate emotional response, whether of fear, shame or anger. In turn, this emotional response should interrupt the present attentional frame, generate plans for new potential words or actions, and thus propel the internal dialogue forward, perhaps in a novel direction. In this way, the vitality and creativity of internal dialogues can be squared with the constraints of biological realism, and we can begin to understand the power they have to color the internal world.

Conclusion

In this account I have attempted to model an internal monologue, hypothesized to be the basis of the dialogical self, subserved by an attentional system in the orbitofrontal cortex and nearby affective and premotor systems. I have

argued that this internal monologue implies the presence of another person, because it is directed toward an imaginal (but unspecific) sense of that person and it adjusts and updates itself through a changing anticipation of how that other person will respond. This model was then extended to account for a switch to another, autonomous, voiced position, underpinned by the ACC and its connections, as exemplified by a spontaneous rebuke directed by a parent-self toward the usual child-self. This simple duality between two positions does not fill out much of the imaginal landscape proposed by Hermans to include many interacting voices. Yet the switch of subjectivities between attentional systems permits the injection of enormous emotional intensity and novelty into internal dialogues, and this may provide the creativity to carve out unique dialogical frames. Different emotional constellations, triggered by self- and other-directed voicings, may recruit distinct cognitive appraisals, each laying down unique synaptic networks that consolidate over development. These distinct networks, reactivated on subsequent occasions, may provide the basis for what appear to be a variety of stances or characters. We are not yet ready to map out these possibilities in detail, but rapid advances in cognitive and affective neuroscience may provide the means for doing so before long. Meanwhile, aspects of the dialogical self appear to be consistent with our knowledge of the brain, and continued research in each of these domains can guide our exploration of the other.

References

Brainerd, C.J., & Reyna, V.F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10,* 3–47.

Buck, R. (1999). The biological affects: A typology. *Psychological Review, 106,* 301–336.

Bush, G., Luu, P., & Posner, M.I. (2000). Cognitivie and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4,* 215–222.

Clark, A. (1996). *Being there: Putting brain, body, and world together again.* Cambridge, MA: MIT Press.

Depue, R.A., & Collins, P.F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences, 22,* 491–569.

Derryberry, D., & Tucker, D.M. (1994). Motivating the focus of attention. In P.M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 167–196). San Diego, CA: Academic Press.

Dore, J. (1989). Monologue as reenvoicement of dialogue. In K. Nelson (Ed.), *Narratives from the crib* (pp. 231–260). Cambridge, MA: Harvard University Press.

Drevets, W.C., & Raichle, M.E. (1998). Reciprocal suppression of regional blood flow during emotional versus higher cognitive processes: Implications for interactions between emotion and cognition. *Cognition and Emotion, 12,* 353–385.

Fogel, A. (1993). *Developing through relationships: Origins of communication, self, and culture*. Chicago, IL: University of Chicago Press.

Fogel, A., Nwokah, E., Dedo, J.Y., Messinger, D., Dickson, K.L., Matusov, E., & Holt, S.A. (1992). Social process theory of emotion: A dynamic systems approach. *Social Development, 1*, 122–142.

Freeman, W.J. (1995). *Societies of brains*. Hillsdale, NJ: Erlbaum.

Fuster, J.M. (1996). *The prefrontal cortex: Anatomy, physiology and neuropsychology of the frontal lobe* (3rd ed.). New York: Raven.

Goldberg, G. (1985). Supplementary motor area structure and function: Review and hypotheses. *Behavioral and Brain Sciences, 8*, 567–616.

Hermans, H.J.M. (1996). Voicing the self: From information processing to dialogical interchange. *Psychological Bulletin, 119*, 31–50.

Hermans, H.J.M., & Hermans-Jansen, E. (in press). Dialogical processes and the development of the self. In J. Valsiner & K.J. Connolly (Eds.), *Handbook of developmental psychology*. London: Sage.

Hermans, H.J.M., & Kempen, H.J.G. (1993). *The dialogical self: Meaning as movement*. San Diego, CA: Academic Press.

Kegan, R. (1982). *The evolving self: Problem and process in human development.* Cambridge, MA: Harvard University Press.

Koski, L., & Paus, T. (2000). Functional connectivity of the anterior cingulate cortex within the human frontal lobe: A brain-mapping meta-analysis. *Experimental Brain Research, 133*, 55–65.

LeDoux, J. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Simon & Shuster.

Lewis, M.D. (1995). Cognition–emotion feedback and the self-organization of developmental paths. *Human Development, 38*, 71–102

Lewis, M.D. (1996). Self-organising cognitive appraisals. *Cognition and Emotion, 10*, 1–25.

Lewis, M.D., & Junyk, N. (1997). The self-organization of psychological defenses. In F. Masterpasqua & P. Perna (Eds.), *The psychological meaning of chaos: Translating theory into practice* (pp. 41–73). Washington, DC: American Psychological Association.

Luu, P., Tucker, D.M., & Derryberry, D. (1998). Anxiety and the motivational basis of working memory. *Cognitive Therapy and Research, 22*, 577–594.

MacLean, P.D. (1993). Perspectives on cingulate cortex in the limbic system. In B.A. Vogt & M. Gabriel (Eds.), *Neurobiology of cingulate cortex and limbic thalamus: A comprehensive handbook* (pp. 1–15). Boston, MA: Brikhauser.

Mathews, A. (1990). Why worry? The cognitive function of anxiety. *Behavior Research and Therapy, 28*, 455–468.

McAdams, D.P. (1985). *Power, intimacy, and the life story: Personological inquiries into identity*. New York: Guilford.

Neisser, U. (1978). Anticipations, images, and introspection. *Cognition, 6*, 169–174.

Niedenthal, P.M., & Kitayama, S. (Eds.). (1994). *The heart's eye: Emotional influences in perception and attention.* San Diego, CA: Academic Press.

Schore, A.N. (1994). *Affect regulation and the origin of the self: The neurobiology of emotional development*. Hillsdale, NJ: Erlbaum.

Schore, A.N. (1997). Early organization of the nonlinear right brain and development of a predisposition to psychiatric disorders. *Development and Psychopathology*, *9*, 595–631.

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

Tucker, D.M., Derryberry, D., & Luu, P. (2000). Anatomy and physiology of human emotion: Vertical integration of brain stem, limbic, and cortical systems. In J.C. Borod (Ed.), *The neuropsychology of emotion* (pp. 56–79). New York: Oxford University Press.

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

MARC LEWIS is Professor of Human Development and Applied Psychology at the University of Toronto, Canada. Over the past decade he has published theory and research in cognition–emotion relations in development and introduced a dynamic systems perspective on emotion and its development. Recently, the dynamic systems approach has led him to integrate neurobiological and psychological accounts of emerging structure, especially mood states and personality formation. This paper is his first piece on dialogicity per se. A key article on the promise of dynamic systems approaches in developmental psychology was featured in the millennial issue of *Child Development*. A recent book, edited with Isabela Granic, is entitled *Emotion, Development, and Self-Organization: Dynamic Systems Approaches to Emotional Development* (Cambridge University Press, 2000). ADDRESS: Department of Human Development and Applied Psychology, OISE/UT, 252 Bloor St West, Toronto, ON, M5S 1V6 Canada. [email: mlewis@oise.utoronto.ca; website: http://home.oise.utoronto.ca/~mlewis/]