The Public Journal of Semiotics

Volume III December 2011 No. 2

CONTENTS

ISSN 1918-9907

Languaging Behaviour as Catalytic Process: Steps Towards a Theory of Living Language[1]

[PART I]

Paul J. Thibault[2],

Faculty of Humanities, University of Agder,
Kristiansand, Norway
Faculty of Education, University of Hong Kong,
Hong Kong SAR, PR China

1. On the Reduction of the Material Dynamics of Languaging Behaviour to 'Text'

In many recent accounts, instances of 'language-in-context' are often thought of as 'text' consisting of verbal patterns and, increasingly, 'multimodal' combinations of, for example, images and verbal text (Kress and Van Leeuwen 2001; Baldry and Thibault 2006; Thibault 1994). The concept of 'text' has been an important one in the language sciences since Harris (1952) inaugurated the systematic study of discourse-level relations in language above the traditional level of the sentence. Since the 1970's, many functional linguists and discourse analysts have made the study of text or discourse a central concern. These approaches have often been inspired by Halliday's definition of 'text' as any instantiation – spoken or written -- of the linguistic system in the form of language-in-context:

> When people speak or write, they produce text. The term 'text' refers to any instance of language, in any medium, that makes sense to someone who knows the language …
>
> (Halliday 2004 [1985]: 3)

'Text', which refers to the instantiation of a language system "in any medium", according to this view, is language that is separated from its material expression, e.g. as the co-orchestration and dialogical coordination of bodily dynamics in talk, or the material inscription of some organized arrangement of visible notational marks on a surface in the case of writing. In this view, talk and (written) text are both instantiated as 'text'. One consequence of this reduction is that the radically heterogeneous character of, respectively,

[1] I wish to acknowledge critically important discussions with Stephen Cowley, Bob Lumsden, Mark King, Neil Mercer, Kai Sorfjord, Sune Steffensen, Susan Stuart, and Feifei Zhou. All of these individuals have contributed in invaluable and inestimable ways to the development of the ideas in this paper. Marcel Kinsbourne also made insightful and helpful comments on an earlier draft. Paul Bouissac deserves special mention for his encouragement, patience, and support. I thank and extend my heartfelt appreciation to these individuals for their willingess to listen and to participate in critical exchanges that have helped me to clarify my ideas.
[2] Contact email: pauljthibault@gmail.com

talk and text is reduced to verbal abstracta. Text becomes the transcription of, above all, the lexicogrammatical choices and patterns in both spoken and written language. The term 'text' itself designates an abstract entity devoid of materiality. Moreover, the very different material dynamics of talk and text are conflated in the term 'text'. This conflation does not account for the material dimension of either as something that is intrinsic to language itself. The discussion below will focus exclusively on talk. However, it is important to clarify the issues discussed above in order to establish a new point of departure for thinking about the material dynamics of languaging behaviour – one which is not based on the more usual idea of 'text'.

The conflation of talk and text in transcription practices – both lay and professional – transforms the experience of the dynamics of talk. It is therefore 'non-neutral' for this reason (Ihde 1979: 21). Ihde draws attention to another facet of the phenomenology of technologically mediated observation and measurement such as the transcription practices under consideration here. Ihde refers to the effect of the "amplification-reduction structure" (Ihde 1979: 21) of these technologies. Whilst the technology of written transcription amplifies some features or aspects of the original event of talk, it reduces others. Ihde (1979: 21) points out that this two-sided effect is an invariant feature of all such uses of a technology to probe some phenomenon. The technological probe is not transparent, although this is so in various ways and to various degrees. In any case, the instrument that is used to probe or record the phenomenon changes our experience of the given phenomenon. In all such cases, the amplification tends to be foregrounded whereas the reduction tends to slip into the background to the extent that it may be forgotten or not even noticed (Ihde 1979: 21). In the case of the written transcription of talk-as-text, verbal abstracta and phenomena on relatively gross time-scales (e.g. discourse moves and turns) are amplified whereas the material, time-locked dynamics of events on smaller, faster timescales and their optic, haptic, articulatory, acoustic and other properties (Bouissac 2006) disappear pretty well altogether.

The resulting reification of verbal transcription-mediated-access to the altogether different phenomenon of talk and its dynamical properties has, of course, been changed significantly in more recent times with the use of more dynamic multimodal forms of transcription and simulation (Galantucci 2005; Steels 2006; Cassell and Tartaro 2007). The main purpose of these brief observations above is to draw attention to the power of writing and the associated doctrine of "scriptism" (Harris 1980: 1-18; see below, this Section) in shaping our understandings of talk in terms of the very different phenomenon of text in both scientific and lay practice. This has fed into a widespread belief in discourse-analytical approaches that talk can be transcribed and analysed on the basis of transcriptions and re-descriptions of verbal abstracta, i.e. what people, including analysts, report they have heard or seen, etc., *as if* these reports were the first-order data to be described and accounted for. In actual fact, these reports are second-order re-descriptions of the data (the material, always time-locked bodily dynamics of agents-in-interaction and their emergent inter-individual patterns) based on the description and transcription of verbal abstracta. Such re-descriptions fail to account for the material dynamics of talk and the small, fast time-scales on which these operate and in terms of which their specific properties and functional capacities need to be explained (see Cowley 2008; Steffensen, Thibault and Cowley 2010; Thibault 2008, 2011).

Overall, the amplification-reduction structure in operation here is implicit in almost all of our scientific, pedagogical, and folk-theoretical accounts of language and their associated practices. This is so to the extent that a folk-theory of abstract verbal forms and the so-called science of these are conflated. Language is analysed and taught for the most part in terms of a culturally promoted literacy-based phenomenology of formal abstracta such as words, sentences, texts, and so on that rarely gets beyond re-description according to the local folk-phenomenology of verbal forms and verbal reports about these forms. On the scientific view, these same formal arrangements are the artefacts on which presumed operations of abstract symbol processing are performed, whether mediated either by internal codes or rules in the brain or by external social conventions. Second-order abstracta also function as scripts that are utilized in educational settings to corral and regiment individual agency and creativity by orienting to standardized norms that fail to do justice to the fact that much of human intelligence, decision-making, and creativity are exercised through the materiality of first-order interaction dynamics between persons, not verbal abstracta (Steffensen, Thibault, and Cowley 2010). This has implications for teaching and learning that remain invisible to pedagogies that are based on standardized second-order constructs.

The text-based model of talk invites us to re-contextualize how bodies engage in real-time meaning-making activity as a form of text-to-be-read (Thibault 1994). Talk can better be better investigated and understood in terms of real-time dialogically coordinated bodily dynamics between persons, the forms of whole-body sense-making that this engenders, and the constraints on these (See analyses along these lines in Sections 3 and 31). As we shall see below, a culturally promoted belief in wordings embedded in associated language practices and interaction routines is one such constraint. Moreover, verbal patterns or wordings are disembodied and reified abstracta that inform our meta-semiotic ways of thinking about and analyzing language. These verbal patterns are described by linguists as corresponding to rule-bound regularities of grammatical or discourse levels of organization. Rules, whether descriptive or prescriptive, are, I contend, second-order cultural constructs and regularities. They are not in themselves the primary data, but second-order re-descriptions of the primary data. The resulting ‘texts’ are typically described and analyzed as verbal and other abstract patterns on different levels of hierarchical organization – e.g. phonological, lexical, grammatical, and discursive – arranged in a sequential order. These approaches – e.g. conversation analysis, discourse analysis - emphasize recurrent patterns. They posit that discourse consists of determinate higher-order units, e.g. discourse-level turns or moves. Discourse is thus seen in terms of program-like regularities of such formal units, rather than dynamical, time-bound properties and functional capacities of the co-orchestrated neural and bodily dynamics of persons-in-interaction.

Whole-body-sense-making-making is, therefore, reduced to predictable arrangements of determinate, recursively organized textual or discourse-level units that derive from a higher-order language code. Interaction reduces to text-in-context (Thibault 1994). Events get detached from time, timing, bodies, activity, environmental affordances and human relationships. Moreover, verbal patterns or wordings are disembodied and reified abstracta that come to inform our meta-semiotic ways of thinking about and analyzing language. These

verbal patterns are described by linguists as rule-bound regularities of grammatical or discourse levels of organization. The modeling of talk using analytical and descriptive abstracta that are based on written notation, including phonetic script, accounts in large measure for the modest progress made by twentieth century linguistics in overcoming its "written language bias" (Linell 2005). The bias identified by Linell is founded on the principle of "scriptism", as defined by Harris (1980: 1-18), i.e. the doctrine that writing is privileged over speaking in literate cultures such that writing constitutes the appropriate meta-linguistic model for talking about and understanding languaging behaviour in all of its manifestations.

In such models, the events in which talk is embedded are de-somatised and de-materialized as 'discourse'. Discourse is seen as sequential and recursive arrangements of abstract verbal forms. The model that informs this approach is that of using text 'in' context (Thibault 1994). Text-analysis *per se* is the basis of theory and description. This model confuses how we think about and use some kinds of written texts with how bodies engage in real-time meaning-making activity. Talk can better be better investigated and understood in terms of the material dynamics of real-time dialogically coordinated neural and bodily activity between persons.

2. The Separation of 'Language' from Dynamics: Trager's Distinction between Language and Paralanguage

The separation of 'language' from biologically grounded neural and bodily dynamics rests upon a number of prior assumptions. Two of these will be discussed here. First, there was the methodological decision, traceable to Saussure, to separate bodily processes from the definition of *la langue* as the object of study of linguistics (Saussure 1993 [1910-1911]: 189; see Section 5 below; see also Thibault 1997; Bouissac 2010). Saussure based his idea of *la langue* on the assumption that "it is the union of an idea with a vocal sign which suffices to constitute the whole of the language system" ["c'est l'union de l'idée avec un signe vocal qui suffit à constituer toute la langue"] (Saussure 1993 [1910-1911]: 189). Saussure's justification for this line of thinking is that phonation in the act of speaking is an individual fact that belongs to "the individual faculty" (1993 [1910-1911]: 189) and therefore that facts concerning the bodily articulation of language are not relevant to the study of *la langue*. In making this argument, Saussure failed to see that human vocal tract capacities belong to a population of agents as part of their evolved biological inheritance. Secondly, the theory of language based on the collective abstraction *la langue* that Saussure developed is a theory of formal abstracta (see Section 5). In constituting the object of linguistic science in this way, Saussure sets up, at least implicitly, an a priori distinction between the verbal and non-verbal aspects of first-order interactive behaviour between persons in talk. This distinction has been maintained in very many folk-theoretical and theoretical accounts of languaging behaviour to the present day. This brings me to the second assumption.

Accordingly, linguistics has proceeded on the further assumption that the verbal part of human behaviour can be isolated from the rest of communicative behaviour and that a science of language – viz. linguistics – based on the verbal part can be set up on the basis of this artificial and entirely arbitrary separation of the 'verbal' and 'nonverbal' components of talk.

For example, one of the earliest theorists of paralanguage, George L. Trager, argued that language was "accompanied by other communication systems, one of motion – kinesics, and one of extra-linguistic noises – vocalizations" (1958: 2). On this basis, communication, according to Trager, "was divided into language, vocalization, and kinesics" (Trager 1958: 2). Language, in Trager's framework, was defined as follows:

> Language will be described here only to the extent of saying that it is the cultural system which employs certain of the noises made by what are called the organs of speech, combines them into recurrent sequences, and arranges these sequences in systematic distributions in relation to each other and in reference to other cultural systems. That is, language has sound, shape, and sense.
>
> (Trager 1958: 3)

Language, in Trager's framework, is embedded in or situated in "the setting of an act of speech" (1958:3):

> When language is used it takes place in the setting of an act of speech. Speech ('talking') results from activities which create a background of voice set (1.2, below). This background involves the idiosyncratic, including the specific physiology of the speakers and the total physical setting; it is in the area of prelinguistics (Trager, 1949, 2-3). Against this background there take place three kinds of events employing the vocal apparatus: language (as described); variegated other noises, not having the structure of language-vocalizations; and modifications of all the language and other noises. These modifications are called voice qualities. The vocalizations and voice qualities together are being called paralanguage (a term suggested by A. A. Hill, who has been interested in the development of these studies). Paralanguage is part of the metalinguistic area of activity.
>
> (Trager 1958: 3-4)

Language is, therefore, made to correspond to the traditional three formal levels of linguistic coding that constitute the basic architecture of 'language', so defined, viz. phonology, syntax, semantics (Trager's 'sound', 'shape', and 'sense'). Definitions of paralanguage vary to include non-linguistic aspects of the speech signal and kinesics such as facial expressions, gestures and body postures. In all cases, the basic assumption is that the verbal and nonverbal aspects of speech behaviour can be separated off from each other as separate entities that are 'combined' in speech. The central contradiction of this view would have it that in speech behaviour abstract verbal forms, on the one hand, and other material, nonverbal audible and visible behavioural events, on the other, are 'combined', to echo Trager's term, though without specifying any causal mechanisms for this curious linkage of incompatibles (i.e. how does one combine formal abstracta with material bodily events?). Language in such a view – the classical formalist one – is identified with wordings, with verbal patterns, which are associated with phonological, syntactic, lexicogrammatical, or other form-based abstracta. In such a view, language is a code-like system of inputs and outputs across these levels that gets

separated from cognitive, affective, and bodily dynamics in real-time (Cowley 2008; Thibault 2008, 2011).

3. Developing Alternatives to Formal Abstracta as the Stuff of Language

How can we move beyond models of language based on codes and abstract forms (e.g. verbal patterns)? The theories of distributed cognition developed by Clark (1997) and Hutchins (1995a, 1995b) provided some early indications. Theories of distributed cognition questioned the view that human thinking is based on the internal manipulation of symbols. In the 1980's, connectionism, neuroscience, dynamic systems thinking, and theories of situated and embodied-embedded cognition also began to challenge the view that human cognition works in the same way that computers process information. Distributed cognition puts the emphasis on cognitive dynamics and on the embodied-embedded nature of cognitive dynamics. Where does language fit in? How can we rethink language in the light of this new thinking? Langacker's cognitive linguistic can serve as a useful starting point.

Cognitive linguists such as Langacker (1987: 57) assume that 'grammar' is an internal psychological representation of a speaker's knowledge of the "cognitive routines that are shaped, maintained, and modified by language use" (Langacker 1987: 57). Moreover, Langacker too, like Halliday, grounds his cognitive investigation of language in second-order abstracta. He argues, for instance, that "the functional and phenomenological characterization of mental experience is consequently more directly relevant to linguistic analysis than descriptions that refer to the firing of specific neurons. For linguistic purposes it is often sufficient merely to establish the existence of a higher-order cognitive entity, regardless of how it might arise from more basic processes … " (1987: 99). This is a species of functionalism that separates cognition from the physical means of its realization.

Accordingly, Langacker evidences no concern with how the conventionalized semantic structures of linguistic cognition are integrated to biomechanical and neural activity. Langacker makes a distinction between higher-order cognition, which he links to a phenomenology of abstract semantic structure. Semantic structure can be described without worrying about the physical stuff that realizes it. Thus, functional and phenomenological characterizations of mental experience can be undertaken without reference to the means of physical implementation of these functional and phenomenological properties. Langacker does not mention bodily processes in relation to cognition. With respect to the brain, he claims that cognitive events of sufficient regularity leave neurochemical traces that facilitate recurrence of cognitive events (1987: 100). On this view, neurochemical traces are the physical stuff that implement and realize cognitive events though this stuff is not necessary for the description of the cognitive event itself. Cognitive events are functional states of the brain. They can be realized by multiple neurochemical traces.

Language, including its 'grammar', according to the distributed language view, is related to interactive loops between brains, bodies, and the external environment rather than internal mental states and their representations or abstract verbal patterns as in current discourse-analytical and form-based approaches. Language is distributed between brain, body, and

environment. The emphasis is on concrete, real, living human individuals who are inter-connected with each other and with cultural artefacts and technologies rather than being mediated by abstract codes or systems. The latter are no more than reified generalities that have no explanatory value. Rather than saying that individuals are mediated by shared abstract systems, the sociality of our interactions is defined and is meaningful in relation to how it is integrated with what has gone before, what is going on now, and what is expected or anticipated to happen next. The distributed view of language emphasises how humans coordinate their actions with those of others in the service of common projects, rather than shared codes or abstract higher-order systems that are said to mediate the interaction. As Harris points out, signs, "provide an interface between different human activities" (2000: 69), rather than mediating between persons and an abstract system or code. The emphasis is on individuating dynamics: how do persons in interaction creatively adapt available resources in order to discover new possibilities of meaning-making, including the inventing of new means, new tools, and new semiotic resources for acting on and transforming the undiscovered potentialities of the materials and technologies to hand? The emphasis is thus on creation and invention rather than the mere use of pre-existing abstract codes, programmes and systems.

The latter view cannot account for human beings as persons: it treats them along mechanistic lines as information processors, mere end-users and instantiators of already defined form-meaning relations, and nodes in computational networks. The real challenge, then, is to investigate and to understand how new assemblages of persons, technologies, and learning systems help people to organize their individual and collective behaviours in novel ways – ways that allow them to organize their actions intelligently and ethically as agents who can evaluate and intervene in events and influence and guide them. On this view, persons are not only shaped and constrained by history; they are in history and are shaping it individually and collectively on, potentially, many different scales of place and time.

The concept of *assemblage* derives from Deleuze and Guattari (2004 [1980]). An assemblage consists of a set of historically individuated component parts and their capacities to affect other parts and to be affected by them. The component parts interact to form a new, emergent whole that is irreducible to the component parts. A key property of such newly emergent wholes is the heterogeneity of the component parts. This means that the parts retain a relative autonomy and identity with respect to the whole in which they function. Unlike structural-functional theories in biology, sociology, anthropology, linguistics, and so on, which assume that the component parts cannot exist independently of their intrinsic relations of interiority (DeLanda 2011: 3) that they have with each other, as assemblage is an emergent whole in which the component parts retain a relative autonomy. As DeLanda (2011: 4) explains, this means that the component parts can be detached or 'unplugged' from one assemblage and attached to another thereby entering into new relations and new interactions. An assemblage is defined by the interactivity between the component parts in terms of relations of exteriority (DeLanda 2011: 3). In an assemblage, the component parts are all historically individuated entities or component processes, not abstract types. Individual persons are individuals in this sense, but so are the events, the objects, the artefacts, etc. with which persons interact to form what I shall call social-affective-cognitive assemblages (Section 28). Assemblages can exist

on many different scales. A conversation between two or more persons is an assemblage. Interpersonal networks, social institutions and historical languages are all assemblages that exist on different spatio-temporal scales. They are all historically individuated entities that have a determinate temporal duration and spatial extension on some scale.

Languaging behaviour is grounded in biomechanics and the natural interactivity and expressivity of our bodies and brains. Discourse-analytical approaches privilege rule-bound sequences of abstract verbal patterns thereby separating such patterns from the individuating dynamics of the social assemblages that are constituted in dialogically coordinated interaction between persons and the matter-energy flows and their modulations that these interactions promote. The interacting components are soft-assembled in the social assemblage in the course of interaction. The component parts are not defined by rigid metric properties, but by their topological connectivity (DeLanda 2002: 56). Specific shapes of, for instance, bodily movements and gestures, including vocal gestures, matter less than their varying and fluctuating dynamical properties, e.g. the varying intensity of the dynamical properties of bodily movements (loudness, orientation, tempo, duration). Bodily dynamics can continuously vary and adapt such that, for instance, a syllable can be lengthened and adjusted in relation to the coupling of a given action to another action. Voice dynamics and other bodily movements can change to adjust to specific coupling relations between different component processes.

As Reed (1996a: 85) points out, there are various levels of bodily postures and nestings of these as well as controlled transformations from one posture to another (movement). A bodily movement is both a change of state from one posture to another at the same time that it is the maintenance of some postural orientation. Vocalizing is no exception. Body movements, seen in this light, have evolved through interaction with the world. Body movements are cognitively salient. They are oriented towards the intentional transformation of the future state of the organism. Vocal tract gestures, facial expressions, manual-brachial gestures, eye gaze, and other bodily resources are all adaptations of the body that sustain and enact intentional, meaning-guided action. The assuming of a particular postural stance, and the mobilization of the neuroanatomical resources required for its enactment, serve to orient the body and thus to indicate to observers the readiness of the agent for some course of action. Many of these postures are, of course, very rapid and subtle; they are micro-temporal or pico-scale bodily events (Section 4) that are rapidly mobilisable owing to the chaotic dynamics of microscopic body dynamics, thereby ensuring a high level of readiness of the system to mobilize and respond quickly to even slight perturbations (Freeman 2000: 87).

I will now consider some aspects of the cognitive significance of bodies-in-interaction in an analysis of a brief episode of infant-mother interaction.

Micro-temporal Phases in an Episode of Infant-Mother Interaction

Typically, a meaning-making trajectory integrates brain, body and world-side factors in short pulses or intervals that are connected to each other along their trajectory. This can be observed in the following episode involving a one-year-old boy, Luke, and his mother, Sheila.

Luke and Sheila are playing a fetching game involving plastic cubes.[3] Sheila encourages and prompts Luke to go and fetch one of the cubes on his own, without the assistance of his mother. In the episode, we see that verbal, postural, kinesic, prosodic, gaze, and facial movements all interact. Each of these controlled movements is a modification in the postural orientation of the body or some part of it. Each such movement is directed towards the realisation of a value in the situation, e.g. the mother wants her son to fetch the cube, to pay attention to her; Luke wants to please his mother, and so on. The main intervals of this brief slice of a trajectory may be summarised with reference to Table 1 as four micro-temporal phases, as follows:

1. The mother takes the green cube in her hand (00.07.28), adjusts her posture inclining backwards 900.08.52) and leans forwards towards Luke, holding the cube in front of him and engaging his interest and attention (00.10.00);

2. The mother shifts her overall posture and body orientation by leaning forwards towards the designated target of the proposed "fetch" with the cube in her hand (00.10.48): this is the onset of tokened or abbreviated "fetch" that the mother is promoting; Luke's gaze and body orientation shift to track his mother (00.11.00 – 00.11.68);

3. She then briefly re-establishes mutual eye gaze with Luke by turning her head to engage Luke's gaze (00.12.60); she then turns her head in the direction of her outstretched hand, holding the cube (00.13.56); maintaining this posture, her gaze then shifts back to Luke (00.14.16);

4. She then turns her gaze towards Luke while maintaining the forward leaning posture (00.14.60). This movement is followed by a rapid sequence of vocalizations, e.g., "go and fetch it", that are rhythmically synchronised with a sequence of rapid head turns towards Luke, on the one hand, and towards the target of the proposed fetch (the cube), on the other (00.14.60 – 00.16.40). The head turns serve (1) to engage Luke; and (2) to direct Luke to the designated target of the proposed "fetch"; Luke begins to move towards the cube (off screen) while his mother directs her gaze to him while holding the posture (00.16.40); Luke starts to crawl towards the cube, smiling, while his mother shifts her posture backwards to give Luke the space to initiate the fetch on his own (00.18.20).

[3] I thank Stephen Cowley for making the video recording of this episode available to me. The analysis is my own.

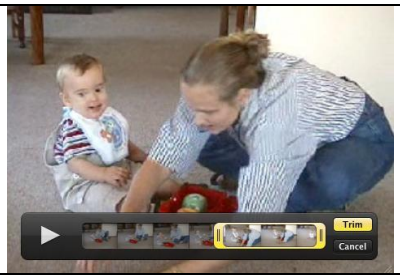
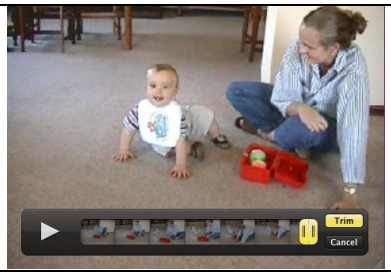
00.07.28	00.08.52	00.10.00
		
Phase 1 →		
00.10.48	00.11.00	00.11.68
		
Phase 2 →		
00.12.60	00.13.56	00.14.16
		
Phase 3 →		
00.14.60	00.16.40	00.18.20
		
Phase 4 →		

Table 1: Four phases in micro-temporal dynamics of tokened rehearsal (fetching game) between mother and infant (original video courtesy Stephen Cowley)

In this brief episode, we see that each of the four micro-temporal phases analysed above forms a pattern of behaviour that is distributed across different bodily postures and their transformations. Moreover, each phase, as well as the episode as a whole, shows the three properties of intentional behaviour identified by Freeman (2000: 116), viz. wholeness, unity, and intent. It makes little sense to split the behaviour into separate channels, codes or modes

of communication that are 'combined' (c.f. Trager 1958; see Thibault 2008 for discussion). Instead, they form a unity from the outset and must be studied as such (Arndt and Janney 1987: 107).

In the present example, we see how the concerted rehearsal that constitutes the fetching game between Luke and Sheila involves the two persons acting on each other, on the relevant objects (the plastic cubes), and on the situation through co-orchestrated and dialogically coordinated real-time bodily dynamics that allow them to generalise and abstract (rehearse) so as to create the meaning structures through which infants like Luke come to perceive and act in the world. Rather than holding an object, a hammer, say, in my hand and exploring it through the haptics and kinesics of tactile exploration, the increasingly compressed and abstracted forms of bodily dynamics enacted in concerted languaging *qua* rehearsal between infants and parents give rise to virtual perception-action cycles that enable infants such as Luke to grasp both in imagination and in action the culturally valued patterns of behaviour that are promoted by caregivers. Engaging in languaging as a form of concerted rehearsal is therefore a means of dynamically shaping and focusing trajectories of value-realizing actions in the human ecology (Hodges 2007, 2009).

Caregivers structure the daily life of infants into meaningful patterns. Meaningful patterns entail participation roles. Caregivers and infants together engage in structured forms of activity involving a differentiation of roles and specializations. Caregivers orchestrate the field of "promoted action" (Reed 1996a: 149) so as to help the child bridge the gap between the child's performance and a competent execution of the skill or action. Melser (2004, 2009) sees language as a mode of concerted rehearsal between individuals. Concerted activities such as languaging entail the acquisition of the requisite perceptual, (inter)actional and cognitive skills for taking part in activities of this kind. Rehearsal implies that the persons who participate in concerted forms of rehearsal such as languaging are readying or preparing themselves for a possible future real performance of the rehearsed action. Rehearsal may also modulate, prompt, guide, direct, orient, correct and organize one's participation in action.

A languaging event consists of a nested set of time-scales, or an array of time-scales that are manifested as different kinds of oscillatory behaviour on different temporal scales. Moreover, in languaging events of this kind different cycles of behaviour on different time-scales are not simply nested the one within the other, but may overlap. There is then the potentiality for radical scalar heterogeneity of time-scales. Biological oscillators may intersect with or overlap with social and cultural oscillators. The speech event unfolds pulse by pulse. Its metric structure is explainable in terms of the relations between oscillators on different time-scales. Rhythmic oscillators have the functional capacity to entrain to each other's temporal-rhythmic behaviour. Thus, two or more persons in conversation have the capacity to form an assemblage – e.g. an occasion of social interaction – by entraining to or synchronising with each other's speech rhythms (Section 17). They can also synchronize to rhythmical oscillations on large scales, both biological (e.g. the night-day cycle or the twenty-four hour rotational cycle of the Earth) and social (e.g. work cycles, classroom cycles, institutional cycles, the collective language dynamics of a population of speakers). Lefebvre (2004 [1992])

uses the term ‘dressage’ to explain how rhythmic patterns are established and integrated over biological, social, and historical timescales. Persons in socially organized assemblages are entrained to and are responsive to these rhythms.

Language is grounded in biomechanics, but is constrained by cultural processes, social conventions and historical traditions. The intrinsic interactivity and expressivity of the biomechanics set the parameters of the capacities of the interacting components of a conversation to couple their neural and bodily dynamics with each other in relations of co-orientation and co-affiliation that also organize the spatial extent of the interaction, its degree of intensity, and so on. Cultural conventions and historical traditions, including lexicogrammatical patterns and social norms, constitute a second parameter. This parameter stabilizes the identity of the assemblage as a whole. For example, ‘grammaticalized’ patterns act as second-order constraints that establish procedures and routines for interaction, the degree or level of formality of the interaction, normatively appropriate ways of differentiating the environment, and interactional obligations, responsibilities, and expectations. Some conversations are highly constrained in these ways, others much less so. Moreover, a conversation does not come into being out of nowhere. The component parts that form a conversation on a particular spatio-temporal scale always come into being when a whole has already come into being and uses its component parts and their capacities to constrain the interactions between them. Thus, persons have intrinsic biomechanical properties and capacities for interactivity and expressivity that can be exercised by these properties. These properties are bottom-up or micro-level contributions whereby individuals actively contribute to and modulate the formation of the overall assemblage. Conversations arise in communities, interpersonal networks, and institutions with their traditions and norms that pre-exist particular individuals. Historical traditions and cultural norms constrain and enable the possibilities for interaction of the component parts of the assemblage. They pre-exist individuals in the sense that they derive from cultural-historical timescales at the same time that they speak through and are immanent in the interactions between the components of the assemblage.

On this basis, the distributed language view has developed the distinction between *first-order languaging* and *second-order language* (Love 1990; Cowley 2007; Thibault 2008, 2011). Verbal patterns (words, wordings, grammaticalized semantic structures, etc.) are second-order cultural patterns that constrain the bodily dynamics of first-order interactivity between persons and between persons and aspects of their environments. A growing body of evidence shows that interactivity, not abstract symbol manipulation, content transmission, or information processing centred on the internal mental processes of the individual, is the key to human learning, cognition and intelligence. Text-based literacies have privileged pedagogies that abstract away from this basic fact because of their focus on a phenomenology of abstract forms and centralized social or cognitive codes. As we shall see below, it is interactivity that enables persons, through their coordinated languaging behaviour, to form the social-affective-cognitive assemblages whereby language operates on and catalyzes socially organized flows of matter and energy.

In the following Section, I define the distinction between *first-order languaging* and *second-order language*.

4. First-order Languaging and Second-order Language

My aim in this section is to offer some preliminary definitions of two terms that will be used to develop an alternative approach to the problems posed by the kind of approach that the early work of Trager exemplifies. First, the term *first-order languaging* refers to the co-constructed and dialogically coordinated body dynamics that occur in real-time interactive behaviour between persons. First-order languaging can therefore be distinguished from *second-order language*, i.e. verbal patterns such as lexicogrammar and discourse, which are stabilized patterns on longer, slower cultural time scales. Agents learn to exploit the dynamics of first-order languaging in ways that get linked to cultural norms and values that connect selves to patterns and traditions emanating from cultural-historical time-scales. The term 'languaging' was first used by Maturana (1970; see also Kravchenko 2003) to emphasise the biologically grounded and processual character of talk as a complex organization in time of biological and social processes oriented to the creation and maintenance of consensual domains due to the structural coupling relations enacted between languaging agents.

First-order Languaging

First-order languaging is embodied inter-individual activity that takes place in real time. It is grounded in body dynamics on fast time-scales measurable in milliseconds to fractions of seconds of fine-grained and co-orchestrated bodily activity between persons (Cowley 2008; Steffensen, Thibault and Cowley 2010; Thibault 2008, In Press). The researchers cited here have developed, in recent years, the notion of *pico-scale bodily dynamics* in order to theorize and to develop methodologies for the study of this scale of languaging behaviour. It is therefore more appropriate to focus on and develop explanations that are grounded in the fast time scales of these pico-scale bodily dynamics that are co-orchestrated and co-synchronized by persons in talk.

It has generally been assumed, for instance, that speech sounds imitate, encode, or express a cognitive-semantic operation or meaning that exists independently of, or is external to, the physical (bodily) activity on some other (more abstract) level of organization or coding. This pervasive view is frequently applied to both the articulatory and prosodic aspects of vocalizations. Instead, the cognitive-semantic operation or process derives **directly** from vocal tract or other intentionally modulated gestural activity. It is the patterns of neuromuscular activation involved in the articulation of a gesture – the neuromuscular program – that directly interact with and operate on the neuro-motor capacities of the addressee. In this way, cognitive-semantic operations are operated in languaging agents. For instance, in English and many other languages, rising tone is an expressive feature of utterances that articulates an anticipatory dynamic (Rosen 1985), i.e. something in the utterance or associated situation is incomplete or unresolved. The rising tone anticipates its completion either later in the same utterance by the same speaker, or by the addressee in his or her response. Rising tone does not imitate, encode, or express the meaning that is glossable

as "incomplete + anticipation of completion" or "unresolved + anticipation of resolution". It is not the arbitrary or conventional coding of such meanings.

Instead, it is an aspect of a dynamically, prosodically shaped vocal tract gestural activity in which the agent assumes an active, dynamical "postural-affective" (Werner and Kaplan 1984 [1963]: 24) state towards some environmental configuration or event that brings about a change in the agent's internal dynamics and organizes a response to it. Viewed in this light, first-order bodily dynamics are intrinsically expressive and cognitive, rather than the arbitrary coding of meanings that are external to them. The latter view is based on the idea that a static sequence of abstract phonemes, e.g. /kæt/, arbitrarily encodes the abstract meaning, "CAT", in the spoken word *cat*. This view arises when utterances are seen as disembedded from the "organismic matrix" (Werner and Kaplan 1984 [1963]: 25) which constitutes their motivation and in terms of which utterances are organized in response to changing environmental events to which the agent needs to adjust. As we shall see in Section 13 below, these observations of Werner and Kaplan tie in with the idea that the internal and external dynamics of the organism and its environment comprise a unified organism-environment system that is the foundation of languaging behaviour as a form of culturally constrained catalytical activity that is grounded in human biology.

The neuromuscular patterns of activation that articulate the rising tone directly enact a cognitive routine; they are not the expression *of* an external meaning that exists on some other level of linguistic coding. These patterns directly enact the cognitive routine in the sense that they bring about a corresponding change in one's interlocutor and therefore in some aspect of the relevant environment. The detection of the pattern by the interlocutor (addressee) may, in turn, result in further changes in the addresser's relationship with the environment as he or she adjusts to the change detected in the addressee. Body dynamics are therefore directly cognitive; they are not the 'expression of' something else in the form of abstract cognitive-semantic processes that pre-exist them. This observation is especially important in the context of first-order languaging because it puts the spotlight on the ways in which agents-in-interaction directly act on and modulate each others' cognitive operations and perceptions in and through their co-constructed micro-temporal bodily dynamics. These bio-physical dynamics are grounded in the natural expressivity of our bodies, i.e. the bio-physical patterns expressed by the body. Moreover, they are organized patterns that compress information and give agents the capacity to interact with other agents in ways that can individuate the self and the other. In second order language, the cultural processes of conventionalization are much more advanced to the extent that this primary biological motivation for the cognitive significance of speech sounds is bleached out though it is not entirely absent.

In actual fact, it is more accurate to say that the rising tone is a behavioural unit, or an aspect of one, in the sense defined by Reed (1996a: 30), i.e. a unit of action on some time-scale that "results in a change in the animal's relationship with its environment" (1996: 30). The 'grammaticalization' of behavioural units, i.e. constellations of phonetic gestures making up 'words' and larger units, is really a preservation or a persistence of pattern that is explainable in terms of the population-level dynamics of higher-order articulatory, acoustic and other

invariants in a community of speakers. These higher-order invariants can function as culturally constrained values owing to the selection pressures of a population of languaging agents. Populations of phonetic gestures habitually relate to populations of environmental affordances on the basis of a history of covariance relations that give rise to statistical effects in a population.

First-order language is time-locked inter-individual activity. It is a bio-cultural process that integrates many different time-scales of organization within and beyond the body and brain of the individual. It is based on co-constructed neural and bodily dynamics that are spread between people in complex processes of circular causation. These dynamics have the potential to connect persons across multiple time and space scales. Moreover, they have the power to coordinate experience around verbal patterns, bodily processes and aspects of particular situations. Furthermore, modulations of body dynamics can bias perception in ways that partition the world into meaningful categorizations. Body dynamics compress and synthesize semantic information that can modify the experience and perspectives of agents-in-interaction. This suggests that hominids that developed this capacity lived in a cultural ecology that allowed them to modify each other's perceptions, cognitions, feelings, actions, identities and values through co-orchestrated body dynamics in micro-time.

As we shall see below, this capacity is linked to languaging behaviour as a form of organism-environment catalysis grounded in experience. The co-synchronization in micro-time of the bodily behaviours of individuals in a particular population transforms perspectives, values and experiences. This capacity depends on the subtle and fine-grained coordination of body dynamics and features of the world that also exploit affect and cultural patterns. The term *languaging* might be a better term for helping us to escape from the limitations and reifying tendencies of the noun 'language'. We need new ways of thinking and new tools of thinking if we are to develop new ideas about 'language'.

Second-order Language

We can refer to verbal patterns – i.e. lexicogrammar -- as *second-order language* (Thibault 2011; see Love 1990 for the original use of this term). Second-order language is what most people have learned to recognize and talk about as 'language' on the basis of a personal and cultural history of metalinguistic practices grounded in literacy and the phenomenology of language form which literacy practices promote. Language is abstracted from bodily behaviour and accordingly treated as a separate and distinctive modality of communication with its own characteristics, sometimes called "design features" (Hockett 1960). When language is theorized in this way, as the example of Trager showed in Section 2, it is separated from the total continuum of communication that occurs in conversation (Harris 1990: 22) and turned into an arbitrary and abstract symbolic code.

Traditionally, language has been viewed as the processing of abstract symbolic forms. Reified synchronic forms are the 'encoders' or repositories *of* meanings, seen as separate from body dynamics. Body dynamics are separated from cognition and first-person experience. Synchronic abstractions replace change and dynamics. This perspective fails to show how

second-order cultural patterns (wordings) get integrated, in real time, to first-order languaging dynamics, not encoded by them. Moreover, first-order languaging dynamics modulate wordings just as the latter constrain the dynamics. First-order language dynamics makes use of audible and visible bodily behaviours in real-time. Our brains tune into and exploit these dynamics in ways that connect experience to normative patterns and values. Body dynamics, not formal abstracta, have the functional capacity to transform perception and action. Co-constructed body dynamics (first-order languaging) have causal and cognitive powers. Interactants are influenced by and respond to these dynamics in micro-time. They exploit pitch, rhythm, cadence, tempo, duration, loudness and other aspects of vocalizations as well as, for example, facial and gestural activities.

5. Typological Thinking vs. Population Thinking: Saussure's Error

Languaging behaviour cannot be reduced to the verbal patterns that linguists transcribe and theorize as 'language' by abstracting verbal patterns from real-time bodily dynamics. On this view, the 'object of study' is a reified abstraction, a hypostatization that has no naturalistic grounding in human biology. This view treats verbal patterns and their purported meanings as abstracta that are separated from the neural, affective, and bodily dynamics of real-time languaging behaviour. Many of the phenomena that are placed under the rubric of paralanguage, kinesics, and so on, are aspects of the very same time-locked dynamics that constitute languaging activity in living, interacting, intentionally charged human agents with bodies and brains. This conception is very different from one that views abstract language forms as the inert and lifeless outcomes of an input/output machine on analogy with the digital computer. Language is grounded in the neurobiology of our brains, our physiology, in perception, in action, and in body morphology. The reduction of language to code-like input/output systems and their regularities denies this grounding. Consequently, language is treated as something that cannot be naturalistically grounded in human biology and first-person experience. It is seen as disembodied and arbitrary cognitive or social processes based on computational models of abstract symbol processing. Cowley (2008) points out that we hear, feel and understand what we say in ways that ground language and interaction in the phenomenology of first-person experience (see also Reed 1996b). On the other hand, if we separate language from biology and first-order experience, we fail to see how languaging is integrated with our feelings and our bodies in productive ways.

In the recent history of linguistics, this tendency can be traced back to the methodological decision made by Saussure when he set up the then nascent science of general linguistics in the early decades of the twentieth century. Saussure's positing of a collective abstraction, viz. *la langue*, as the object of linguistic sciences, has helped to stymie a proper appreciation of the radically heterogeneous character of language (c.f. Saussure's exclusion of *le langage*). In insisting on a homogenous abstraction as the object of linguistic science, Saussure privileged essentialist and typological modes of thinking traceable back to Aristotle that have remained stubbornly ingrained in linguistics to the present day. Thus, on analogy with the exception-less laws of classical (Newtonian) mechanics, *la langue* was seen to be the repository of everything which the linguist held to be "essential and universal" ["essentiel et universel"]

while at the same time putting aside other factors deemed to be "the particular to the accidental" ["le particulier de l'accidental"] (Saussure 1993 [1910-1911]: 192). On this view, *la langue* is a "collection of abstractions" ["un ensemble d'abstractions"] (Saussure 1993 [1910-1911]: 192) from which everything "essential and universal" about language(s) can be logically deduced as constant formal regularities.

In shifting the focus from specific historical languages (*les langues*) to the theoretical abstraction of *la langue*, Saussure viewed *la langue* in terms of abstract forms and general categories rather than as historical individuals or singularities on some spatio-temporal scale. Consequently, the material dynamics and historical processes which have lead to the morphogenesis of these individuals are set aside. Causal explanation based on material dynamics was replaced by the description of formal regularities, most notably the union of the two mental abstracta – acoustic image and concept – in the formation of the linguistic sign (Saussure 1993 [1910-1911]: 190). Moreover, *la langue* is viewed as being logically necessary for the production of this union (Saussure 1993 [1910-1911]: 190) even though the specific causal mechanisms responsible for this union elude such an account. In adopting this view, Saussure effectively eschewed productive causes in favour of a theory (and its object) founded on formal regularities and abstracta. It was on this basis that Saussure formulated a small number of general principles or axioms, notably the *linearity principle* and the *principle of arbitrariness*, on the basis of which a large number of consequences are deduced. Saussure's rejection of a naturalistically grounded mode of enquiry therefore paved the way for a theory of language based on essences, rather than material dynamics and productive causal relations.

Saussure's theoretical object, *la langue*, is a collective synchronic abstraction. Saussure's essentialist and typological view can be contrasted with the population thinking that was first developed by Darwin. As the brief observations above show, Saussure's way of constructing the theoretical object, *la langue*, makes thinking about variation and change difficult or problematic. Variation is thus seen by Saussure (1993 [1910-1911]: 192) as "the particular and the accidental" and therefore is a departure from the typological criteria established by the abstract and homogenous theoretical object, *la langue*.

Population thinking, on the other hand, views formal abstracta and types as no more than theoretical and heuristic tools rather than hypostatized ontological levels of language. Populations constitute the basis for groupings of individuals. Causal relations between individuals and individual variants are a central part of the explanation of a population and its dynamics. Populations are not abstract totalities, but concrete historical individuals on some spatio-temporal scale. They have extension in space and time. A language (e.g. Catalan, Italian, Putonghua, Urdu) is, then, a concrete historical individual on its spatio-temporal scale just as an individual person is on a very different scale.

On this view, a historical language is an individual on its particular spatio-temporal scale. It is not an abstract system that is instantiated as specific instances that conform to varying degrees to the typological criteria established by the formal regularities and schematic properties that are seen to comprise or define the essence of the language system. Instead, the relationship

between a language and the smaller scale individuals that comprise it is one of parts to whole. Thus, a language qua historical individual on its far greater spatio-temporal scale is made up of the dynamical interactions between the smaller scale individual component processes (e.g. languaging events) that are the component parts that comprise the larger-scale entity. The critical difference for our present purposes between the two views – typological and population thinking -- resides in the fact that a population is not defined in terms of structural arrangements of formal abstracta, but in terms of its material dynamics and the causal processes that change the system over time. A system based on formal abstracta is unable to achieve this kind of causal explanation. Moreover, the population view takes it as entirely unproblematic that a given population may be characterized by varying degrees of heterogeneity or homogeneity without feeling the need to postulate an abstract totality, viz. Saussure's *la langue*, as the object of study.

In population thinking, the relations between individuals on different scales are one of parts to wholes, rather than a type-token or a system-instance relation between different ontological levels, e.g. the system level and the instance level. It is the causal interactions among the individuals – seen as historical singularities – on any particular spatio-temporal scale (not ontological level) that leads to the emergence of the next larger-scale individual. A population on any given scale is an assemblage of co-articulated material processes that has a given spatial extension and that exists over a given time-span. It is created and maintained by its material dynamics and the productive differences these give rise to, rather than the purely abstract and negative conception of difference on which Saussure based his conception of linguistic *valeur*. The fact that a language consists of many sublanguages and sub-varieties of various kinds is itself a contingent consequence of population processes in the socio-cognitive-interactional dynamics of, say, the members of a particular interpersonal network. For present purposes, the important point is that population thinking provides a better theoretical explanation of language at whatever scale in terms of time-locked processes and material dynamics on different scales rather than in terms of synchronic states, formal abstracta, and outputs of formal systems.

'Grammaticalization' is sometimes seen as a matter of compositionality, i.e. the composition of abstract grammatical forms from smaller component parts such as words. In actual fact it is, in the first instance, the preservation and persistence at the population level of very high-order physical invariants in a repertoire of phonetic gestures. These invariants afford access to a world of virtual cultural entities once one has learned the cultural tricks for accessing them. The phenomenon of grammaticalization refers to the automatization of regular, repeated sensori-motor routines in a population in order to reduce the cognitive and sensori-motor costs of linguistic interaction (Bybee 1998; Bybee and Thompson 2000; Argyropoulos 2008). These invariants are a form of higher-order perceptual information that can catalyze in self and others apperceptions of virtual objects and events that need not be physically present in situations or accessible to on-line perception. Once such virtual entities became established in a population as culturally salient and valuable entities that matter to agents, so too did meta-linguistic models and procedures for re-categorizing in socially distributed ways the phonetic gestures as populations of lexicogrammatical forms that are analysable as belonging to

different word classes and their possibilities of recursive recombination. The two developments would have emerged in parallel, viz. in a nutshell, reflexivity and recursivity, the coupling of which induces its own catalyzing effects. On this view, the re-categorization of a repertoire of phonetic gestures as lexicogrammatical categories in a given population of agents is a form of cultural scaffolding. Specifically, it is a cultural scaffolding of the growing diversity of interaction routines and associated situation-conventions that the phonetic gestures afford agents as more and more pressures for standardizing the coordination of agents builds up.

The extraction of prototypes on the basis of the covariance relations between environmental events and the expanding repertoire of recursive phonetic gestures give rise to statistical properties at the population level. These properties arise from the perceptual sampling and exploration of phonetic gestures in a given population. They enable individual agents on particular occasions to determine or interpret the intended meaning of any given gesture through repeated, habitual participation in and observation of the interaction routines and behaviours in which these covariance relations arise. This development would in all probability give rise to selection pressures for agents to generalize from their experience of the situations in which the gestures habitually occur. Through informal and formal learning processes, agents pool these generalizations as part of collective narrative wisdom, again at the population level, such that the accumulated wisdom is potentially accessible by all members of the population. In turn, the pressure to generalize would bring about further selection pressures to segment the continuously varying character of the gestures into component parts that are associated with or can be associated with differential aspects of situations.

Grammaticalization arises in a population as these associations themselves are subjected to more and more layers of cultural constraints as the process of segmentation spreads, owing to its augmented capacity for generalization and the enhanced cognitive and semiotic reach that this capacity affords. The spread of these processes catalyzes the rise of digital semantics and, hence, the expanding pluriverse of virtual cultural entities and events that enable humans to keep track of self and others over an expanding diversity of space and time scales and niches in the extended human ecology. The output of these combined processes catalyzes particular effects. In turn, these effects become the input to new processes that give rise to cycles of self-organizing iterative linguistic structures. These structures exhibit more and more properties of global coherence and identity that persist on the long historical-cultural timescales of a historical language. The units of selection – prototype extraction, generalization, segmentation, association of phonetic gestures with covarying environmental affordances -- are positive feedback processes giving rise to an autocatalytic web of socio-cognitive interaction strategies and behaviours that are perpetuated on cultural timescales as second-order lexicogrammatical and discourse-level and other constraints on these same behaviours.

6. The Scalar Heterogeneity of Language

It is commonplace to say that language consists of different levels of organization. Generally speaking, these levels are seen as inter-related aspects of the formal architecture of a language that can be localized in code-like mechanisms. Such mechanisms are postulated as existing inside the brain of the individual or in externalized artefacts such as 'text' (Section 1). From the perspective of population thinking, language is a radically heterogeneous phenomenon that is causally spread across a diversity of different spatio-temporal scales (Lemke 2000a, 2000b; Thibault 2004a, 2004b; Wheeler 2005), rather than being isolable on any one of them or homogenized as inter-related levels of abstract forms. A focus on any particular scale may be an entirely legitimate line of enquiry for certain purposes, but it cannot be the whole of language. Language and the explanation of language cannot be confined to any one spatio-temporal scale.

According to population thinking, smaller scale entities are the parts whose interactions comprise the next higher-scalar entity. The classic example in biology goes somewhat (and approximately) like this: genes are parts of chromosomes; chromosomes are parts of cells; cells are parts of specialized organs; organs are parts of organisms; organisms are parts of groups; groups are parts of subpopulations; subpopulations parts of species, and so on. On each of these scales, the entities that comprise that scale have the capacity to reproduce. Moreover, there is variation and this variation can affect reproduction and rate of reproduction of the entities that exist on any given scale. For these and other reasons, the different scales proposed in this hierarchy show Darwinian characteristics. This does not mean, incidentally, that all population dynamics are Darwinian ones. Many aspects of social and cultural dynamics entail reproduction though not necessarily through Darwinian processes involving parent-offspring lineage in the biological sense.

As for language, we can postulate that language is spread over a number of scales, populated by their respective entities. The scales across which language is spread can be approximately and by no means exhaustively outlined as follows:

i. Neurons are the component parts of populations of neurons (Edelman 1987, 1989). The central nervous system is a system of oscillators and the modulatory relationships between oscillators. Neurons and neuron circuits are oscillatory and are modulated by influences from other neurons and neuron circuits in distributed networks of brain activity. The patterns of interactions of neurons with other neurons form stable patterns of oscillatory networks of neuronal activation leading to the emergence of populations of neural networks. The content of these networks is a prototype that is created by sampling the objective statistical properties of a population - properties that can be transformed into a subjective felt experience of a linguistic category;
ii. Pico-scale bodily dynamics measurable in milliseconds to fractions of seconds are stable yet varying patterns of bodily behaviour that form populations of dynamical bodily events in the observable languaging behaviour of individual persons and groups of persons;

iii. Populations of phonetic gestures are formed from the entrainment of lower-scalar pico-scale bodily dynamics to higher-scalar patterns of vocal tract articulatory and auditory gestural activity. These stable patterns are based on rich phonetic memory that is grounded in first-person experience of phonetic events (Browman and Goldstein 1992, 1995; Fowler 1986, 2010; Port 2007, 2008), not abstract phonological schema;

iv. Inter-individual interaction routines based on the co-synchronization and entraining of an individual person's repertoire of vocal tract behaviours to group level dynamics and strategies requiring the social coordination of persons gives rise to inter-individual patterns that are not reducible to the component individual persons on the scale below;

v. Language practices of particular social groups associated with particular situation-types and the conventions for coordinating persons in situations;

vi. Lexicogrammatical patterns are the statistical results of cultural processes of institutionalized teaching, standardization, and transmission across generations (Halliday 1991) that emerge through populations of language practices. In this way, a given interpersonal network of interacting agents preserves (and values) higher-order physical invariants of populations of phonetic gestures and associated interaction routines. Phonetic gestures are, accordingly, re-categorized as lexicogrammatical patterns, i.e. normative patterns that stabilize the identity and the socio-cognitive dynamics of an entire interpersonal network or community of interacting agents.

Unlike the relation between a higher-order system and its instantiations, the relationship of smaller scale entities to larger scale ones is one of parts to whole. For example, populations of phonetic gestures are the parts that comprise populations of interaction routines that coordinate persons. Moreover, whereas the system-instance instantiation relation is based on degree of conformity of the instance to the more schematic criteria established by the higher-order system category, the relation of parts to whole in a population is causal. Wholes emerge through the interactions among component parts as a consequence of the specific causal mechanisms at work on a given scale and the often complex interactions over time across scales. For example, interaction routines emerge through the ways in which phonetic gestures entrain to and synchronize with higher-order inter-individual patterns that, in time, characterise particular social groups.

Saussure distinguished the social and individual dimensions of language in terms of *la langue* and *la parole*. These two terms were derived from *le langage*, which Saussure put out the door of linguistic theory on account of its unmanageable heteroclite character. Saussure's reasons were as follows:

> Le langage est un terrain complexe, multiforme, héteroclite dans ses différents aspects. Un consequence, c'est qu'on n'arrive pas à la classer pris dans son tout avec d'autres faits humains. Il est à cheval sur des domains divers (domaine physique, psychique, ou encore: domaine individual, social.) (On ne sait comment lui conférer l'unité.)
>
> (Saussure 1993 [1910-1911]: 276)

On the distributed view, language just is a radically heterogeneous phenomenon that is spread across a diversity of timescales, as outlined above. Saussure's *le langage* can serve as a new point of departure such that real-time dialogically coordinated languaging events between persons are not the outputs or instantiations of a homogeneous and abstract *la langue*, but the emergent outcome of the interactions between component processes and their capacities deriving from a diversity of timescales. A languaging event is an assemblage in which populations of interacting component processes and their capacities on the scale of that event are constrained and enabled by populations of entities and their capacities on other scales that extend from the neural to the cultural-historical, as shown above, and are integrated to that event. The component processes on any given scale are individuals -- populations of individuals, to be more exact, i.e. populations of neurons, of bodily dynamics, of words, of persons, interpersonal networks, and so on. As I said before, populations of component processes on smaller scales (e.g. patterns of neural activation, intentions and body dynamics) are sub-personal component processes of the individual person; they have the capacity to modulate the behaviour of that person just as the coordinated behaviour of the individual person modulates the larger-scale interaction-system between persons that is a conversation. On larger scales, populations of cultural-historical patterns that have arisen in a community in part as the unintended statistical consequences of the very many interactions between persons entrain the dynamics of the lower scales to their own dynamical patterns such that the latter constrain and enable lower scalar component processes.

Grammaticalized patterns therefore are normatively constrained procedures that embody in the dynamics of a population the successful and valued patterns and routines of the past (Section 28). The capacity of grammaticalized patterns for increased abstraction means that they can function as cultural scripts that compress and synthesise cultural-historical information. They are performance and dispositional scripts that have the functional capacity to prompt us as to how to act, orient, attend, perceive, feel, and think, and so on, in specific circumstances. When activated, they get coupled to bodily dynamics, which they shape and entrain in culturally specific ways that lead persons to act in appropriate (and inappropriate) ways. Grammatical scripts can therefore serve to prompt and guide behaviours, though one needs to have access to the culturally compressed information in order to ensure that grammatical patterns are appropriately coupled to bodily behaviours. This is clear in most language pedagogy, which teaches disembodied grammar scripts as verbal forms on the basis of a misleading code view of linguistic 'communication' without attending adequately to their implications for bodily performance in interactional encounters with others and the social-affective-cognitive assemblages that these give rise to.

Another big difference with typological thinking is that in populations of whatever scale the entities on any given scale are historical individuals or singularities that operate at a particular spatio-temporal scale and which interact with entities on other scales. Thus, an historical language, like a biological species, is a historical singularity that exists on a particular spatio-temporal scale just as a single neuron is also an individual existing on a very different spatio-temporal scale. On the other hand, the relationship between a higher-order language system and specific instances is based on the Aristotelian assumption that this relationship is a

hierarchical one. In a relationship of this kind, the entities on different levels of generality are postulated to exist on different ontological levels and therefore to correspond to different ontological categories.

Population thinking emphasises the unique and the singular rather than the typical and the general. Individual neuronal networks are unique just as individual phonetic gestures and individual languages are unique on their respective scales. The collective interactions among the entities on a particular scale, as Mayr (1976) pointed out, can only be described in statistical terms. Lexicogrammatical patterns are, in Mayr's terms, statistical abstractions deriving from the heterogeneity (not homogeneity!) of the interactions among the individuals that comprise the population. However, the collective character of these interactions means that some patterns are selected, taught, and transmitted such that they propagate throughout a population from centres of teaching, power and influence and are reproduced or replicated as normative patterns to which individual behaviours are entrained and to which they are expected to conform for the good of the collective. This does not mean that the statistically unlikely state of homogeneity prevails, unlike Saussure's idealized and homogenized conception of a synchronic *langue*.

Lexicogrammar is a mode of normative empatternment that evolves within the socio-cultural dynamics of collectivities of persons and their institutions. However, the spatio-temporal extension of a particular language qua historical singularity (not idealized system, norm or type) means that heterogeneity and variation are the likely normal state of affairs. In the sense that lexicogrammatical forms are the statistical consequences of collective and socially distributed interactions between persons, they are virtual entities that individual persons learn to perceive and orient to as actualized in the own and others' linguistic behaviour as specific wordings. This achievement takes place through participation in meta-linguistic practices of learning to attend to aspects of the dynamics of vocal behaviour qua higher-order articulatory, acoustic and other invariants that are entrained to and shaped by cultural attractors. In the first instance, these practices begin, most fundamentally, in the parent-infant dyads that are created in the first year of the infant's life (Halliday 1975). We learn to talk about talk in ways that shape and (self-)regulate talk around norms that the socio-cultural milieu requires us, either informally (e.g. a sense of solidarity or belonging to a group and having an identity in that group) or formally (e.g. pedagogical enforcement of language standards in schooling), to conform to for the collective good. Moreover, conformity brings its own emotional rewards at the same time that nonconformity is punished, e.g., group solidarity brings a sense of belonging and therefore a sense of group identity; non-conformity can lead to forms of social stigmatization of non-standard varieties and their speakers.

The vertical coupling of scales is central here. For example, an individual person's brain consists of populations of neural networks and the interactions between these that occur on a smaller spatiotemporal scale with respect to the individual. Neural dynamics are embodied in a larger-scale entity, viz. the central nervous system and body of the individual organism. Second-order cultural patterns such as lexicogrammatical differentiators are control parameters that index emergent, self-perpetuating virtual relations in a population of

interacting agents. Second-order patterns of this kind persist on the long, slow historical-cultural timescales of a population of languaging agents. Our ability to adopt the 'language stance' (Cowley 2011) on the dynamics of languaging behaviour and to perceive lexicogrammatical patterns (wordings) in real-time languaging events means that these patterns are also seen and heard as intentional semantic contents on the short timescales of real-time interaction. Intentional semantic contents are part of the endlessly evolving hierarchies of control structures (Juarrero 1999; Van Orden and Holden 2004; Thibault 2004b. They emerge and are perpetuated in time by means of circular causality as systems of symbolic associations (Deacon 1996) on cultural timescales.

Living systems originate in the circular, positive feedback processes of chemical autocatalysis. In autocatalysis, the output of a chemical process becomes in its turn its input and catalyzes the same process. Autocatalysis thus perpetuates the reaction in continuous cycles of chemical reaction. The chemical reaction itself appears, on its time-scale, as a coherent, self-organized, iterative structure. Because lexicogrammar patterns are perpetuated in time as webs of symbolic associations by autocatalytic processes, they have the functional capacity to constrain processes on smaller, shorter timescales of real-time interaction between individual persons. This helps to explain lexicogrammatical patterns as both normative constraints emanating from cultural timescales, on the one hand, and intentional semantic contents that are attributed to individual persons in interaction as aspects of what we perceive in the dynamics of first-order languaging behaviour, on the other. The developmental and individuating trajectories of persons vertically couple more and more scales or organization and process.

The vertical coupling of different timescales on account of biofeedback processes makes this possible. Vertical coupling means, for example, that neural events on their time-scale are not seen as random fluctuations, as seen, say, from the viewpoint of the cultural time scale. Instead, they are coupled to events on longer, slower cultural time-scales (amongst others). Bio-feedback processes in complex bio-cultural systems such as in languaging behaviour are vertically coupled on many different timescales, e.g. those stipulated above, taking in the neural to the cultural. As Van Orden and Holden (2002: 94) point out, this leads to the emergence of "fractal patterns of long range correlation", e.g. between the neural and cultural time-scales, or between intentional semantic contents of utterances of individual persons and the lexicogrammatical constraints emanating from the cultural scale.

Linguistic utterances are operators on implicit underlying networks of representations rather than encodings in propositional form of representations of states of affairs (Harris 1991; Bickhard 2004; Thibault 2005a, 2005b, 2011). In individual brains, these representations are stable patterns of neural activation leading to neuronal networks the content of which is created by the individual agent's sampling of the objective statistical properties of a population of, e.g., observed covariances between parts of utterances (e.g. words) and aspects of situations. In a particular situation, both standing constraints and contingencies will conspire to activate and transform this stored prototype into an intersubjective experience when two brains come to resonate with each other owing to the co-synchronization of each

person's neural and bodily dynamics in their dialogically coordinated languaging behaviour. Because representations are also conventionally and normatively associated with particular situation-types and their conventions, which utterances operate on and transform, the intention that modulates the utterance and flows through its trajectory, can be inferred or interpreted due to the fact that utterances take place in relation to situations which are normatively constrained. If the utterance, "Can I help you?", uttered in a shop by a salesperson, has acquired the semantic-pragmatic status of a Sales Bid, it is because it takes place in a situation and in relation to its implicit underlying representations. This means that agents are able to correctly interpret the customary intention that guides the speaker's uttering of this vocalization as a bid to initiate a sale with a potential customer. The intention becomes customary and therefore easy to interpret owing to the intensity of the significance that persons attribute to the situation and the frequency and density of the agent-situation interactions in relation to which the intention is interpreted in a given population of interacting agents.

7. Abstract Symbolic Forms and/or Material Bodily Dynamics

Languaging behaviour is forever under construction. Every utterance that occurs is a historical singularity. Many current approaches to language processing assume the opposite: they assume that the same conditions yield approximately the same output, or that the same choices are instantiated from the same abstract system, consisting of a priori schema and component processes. Input/output processing models of language are based on the idea (in various versions) that language consists of a set of inter-related levels of abstract form – e.g., phonology/graphology, morphosyntax/lexicogrammar, semantics -- mediated by coding relations. Such approaches assume that language has its basis in smaller component atoms that combine through rules to create language output.

On this view, high-level abstracta such as phonology and morphosyntax mediate and filter out irrelevant small-scale detail on any given level of the coding cycle so that only the abstract formal categories of the system are considered relevant to its internal architecture and functionality. Small-scale bodily dynamics and fine-grained details of the situation do not enter into the description of linguistic inputs and outputs assumed by the views described above. They are not therefore seen as relevant to or as affecting languaging behaviour. How do such approaches explain context-dependent languaging behaviour? One way is to assume that stable linguistic components combine in situationally specific ways according to stable social or generic conventions and expectations that skew the outputs of the system. The outputs of the system in this approach are described as 'text' or 'discourse' in many recent accounts. Text and discourse are seen as the manifestation of formal regularities and patterns above the level of the sentence, which previously was the main concern of classical linguistics. There are two critical problems with this approach.

First, the preoccupation of linguistics with formal abstracta is preserved though it is now shifted or extended to higher levels of linguistic organization 'above' or 'beyond' the level of the sentence. Secondly, the components of linguistic behaviour are never found to exist in context-independent ways that allow us to define any context-independent criterion for the

establishment of stable forms that are reliably the same from one situation to the next. The attempt to understand languaging behaviour on the basis of stable, abstract forms and the regularities of their combinations has resulted in languaging activity being removed from and studied independently of its biological implementation as inter-individual bodily dynamics in a population. Albeit with varying emphases, this is no less so of many current cognitive and functionalist approaches to language and discourse. These theories are based on a 'levels' account of language, seen as comprising a number of inter-related levels of abstract form and the coding relations between the different levels.

For example, Halliday (2004 [1985]: 24-26) proposes that the content stratum of language is internally stratified into two layers: an 'interfacing' stratum (semantics) and an 'organizing' stratum (lexicogrammar). Thus, the semantic stratum interfaces "with what goes on outside language: with the happenings and conditions of the world, and with the social processes we engage in." (2004 [1985]: 24). Halliday argues that this involves two steps, as follows: First, "experience and interpersonal relationships" in the ecosocial environment are transformed into meaning. This is the semantic stratum. The semantic stratum interfaces with the world outside language and transforms it into semantic content. The semantic categories of language construe the ecosocial environment in which language is embedded. Meaning, in a further step, is transformed into wordings (lexicogrammar) on the next level down. On this view, lexicogrammar is the stratum on which meanings are organized as wordings. Cognizers and perceivers are not explicitly mentioned in this account; the emphasis is on the internal stratification of the content plane of language into semantics (content substance) and lexicogrammar (content form). In any case, the account is in conformity with formal abstract symbol processing models of 'meaning'.

Such models assume the existence of a world or environment 'out there' that is transformed into meanings by the categories that are internal to the language system. Halliday does not discuss the biological basis of this assumption. Rather, there is an implicit division of labour between the work of the linguist and the work of the neuroscientist. The task of the former in Halliday's account is to specify the functional architecture of the linguistic system and the role it plays in the overall process of symbolically processing or coding the world as meaning according to the semantic categories of language. In spite of the absence of any explicit discussion of the cognizers and perceivers who participate in these processes, the implicit assumption shared with the symbolic processing view is that information in the form of "experience and interpersonal relationships" is transformed into meanings. This entails some kind of representational process that underlies these processes of transformation. In sidestepping the role of cognizers/perceivers in order to focus on the architecture of the language system, we are left with a split between the "ecosocial environment" and the "linguistic system." Moreover, the absence in this account of cognizers/perceivers qua active agents provides no way of seeing that languaging behaviour is inseparable from and is an extension into the environment of the bodily processes of agents. In the alternative account to be explored and developed below, languaging is a mode of behaviour that catalyzes or mediates the organism-environment relation as a unified system. In this account, lexicogrammar acts as constraints on behavioural events whereby agents actively catalyze

their environments. Rather than the symbolic processing or construal of experience into meaning, as in Halliday's account, unified organism-environment catalysis just is experience.

On the other hand, careful analysis of real-time languaging behaviour shows that the components of linguistic behaviour exhibit a wide degree of natural and intrinsic variability. It is therefore difficult to sustain the view that the dynamical properties of real-time language events are reducible to instantiations of higher-order systemic or schematic properties. The view of language as consisting of levels of abstract form that express or realize semantic functions of various kinds is a logical reconstruction based on formal type-token or system-instance instantiation relations internal to a synchronic language system. Instead, I shall argue that language is not adequately definable as a well-defined synchronic 'object of study', but consists of populations of dynamical diachronic processes that are spread across diverse spatio-temporal scales ranging from the neural to the cultural. In languaging activity, these scales are integrated with each other in varying ways and to varying degrees in real-time languaging behaviour between persons.

8. Functions of Language: Etiological and Normative Approaches

According to functional theories of language, functions are expressed or realized by formal properties of linguistic utterances. Certain formal properties are seen as having evolved to perform some semantic or other function. Semantic functions are thus realized by properties of lexicogrammatical form on the content stratum. These functions are abstracted from the material dynamics of utterance activity. In the dynamics of first-order languaging behaviour, it is more problematic to separate function from the material dynamics in this way. Can first-order dynamics be subsumed by and reduced to a set of abstract semantic functions? The complex material dynamics of first-order languaging cannot be reduced to or explained by in terms of a set of abstract functions instantiated in abstract form. In other words, there is a problem of instantiating, in a levels account of language, abstract semantic-pragmatic functions with their own formal properties and modes of realization within the material dynamics of vocal tract gestural events.

A vocal tract gesture is a physical act of an individual that is guided and modulated by both intentionality and meaning. It is an exploratory act, or a 'choice' from a seeking system, in the sense that in acting on some aspect of its world by means of the gesture, the individual seeks responses that will provide further information about the world. Consider an agent who is intent on getting a cup of coffee in Starbucks -- he or she produces the utterance, *a tall cappuccino*, on being called to the counter by the salesperson. The choice of a particular action trajectory reduces the uncertainty of 'what next?' Both the speaker's desire for a cup of coffee and the salesperson's readiness to provide the desired service could be said to elicit or call forth the utterance, which then reduces the uncertainty of what will happen next. The utterance qua physical gesture thus serves as a constraint on the entropy production of the event (Salthe 2011: 423). It serves the function of getting the speaker what he or she desires – the tall cappuccino. The utterance does not *have* this function; this function is not in any way encoded in the formal organization of the utterance. Bickhard (2008) has discussed at length the difference between etiological approaches to function (i.e. something *has* a function) and

his idea of emergent normative function (i.e. something *serves* a function). The utterance does not have the function of requesting a cup of coffee; its component parts and their relationships cannot be explanatorily reduced to the function of requesting coffee.

On this view, the parts of the whole utterance and their inter-relations have functions in the following sense: The parts that constitute the whole have dependency relations with each other that presuppose a system of relations in which the parts have the functions that they do. Both the parts and the whole have the functions they do because of the functional relations that derive from the structures and forms of organization of the system in which they occur. On the other hand, normative function in Bickhard's sense can be modelled in terms of how something *serves* the function of contributing to the self-maintenance of a far-from-equilibrium system. In its own modest way, the utterance *a tall cappuccino* serves the organized functioning of the self-maintenance of the agent qua far-from-equilibrium system. By the same token, the agent's recursive self-maintenance requires some means of differentiating environments and the interaction potentials of these environments. Differentiations must be connected to possibilities for interaction. The utterance *a tall cappuccino* has this differentiating function. But first the speaker must assess the multiple possibilities of interaction with the relevant environment and make a selection. The desire for coffee, the opportunity this presents to chat with friends, the social milieu of Starbucks, etc. are all culturally saturated factors that potentially inform the selection of a particular interaction. Thus, the speaker determines that these factors provide indications as to the appropriateness of the environment for the selection of an interaction of the requesting coffee kind. The indications functionally presuppose that the conditions for the success of the interaction are valid in this environment.

As Bickhard (2008) shows, the interaction can only succeed if the environment is one that is appropriate to the interaction. The utterance qua interaction with the environment is a differentiator. It seeks to differentiate some aspect of the situation that is selected as the current locus of cognitive processing. Differentiations can be judged to be true, false, correct, incorrect, appropriate, inappropriate, etc. If the environment supports the differentiation in the anticipated way, then the presuppositions of the indications received are verified and the interaction will succeed. On the other hand, if the environment falsifies the presupposition, then the interactive differentiation in the utterance is falsified for the agent precisely because the agent can detect that it is a differentiation that is not supported by the environment in question. Vocal tract gestural activity is a sub-personal component of the agent that has this function.

Human agents are living systems with a specific form of embodiment. Our embodiment confers the required degree of stability on the agent such that vocal tract and other forms of gesture can be articulated with the requisite degree of reliability and stability from occasion to occasion. The differentiations that can be made by these gestures must be stable and guaranteed by the embodiment of the organism. A stable embodiment ensures that the agent can engage in the relevant environmental differentiations and interactions. The selection of a given utterance functionally presupposes that the current environment is an environment that

is appropriate for the agent-environment interaction indicated by the utterance. The example thus presupposes that the environment is one in which the uttering of the utterance will lead to a transaction resulting in a tall cappuccino. If, on the other hand, I said the same thing to the sales person in the Post Office, the current environment may not be one that is appropriate for the indicated interaction. The selection of the utterance presupposes there is an environment in which coffee can be obtained. If the presupposition is wrong, then the interaction is dysfunctional with respect to the given environment (Bickhard 2008).

It is now possible, I think, to clarify what it means to say that an utterance is a form of seeking activity. A particular environment typically has many possibilities for interaction. Starbucks is no exception. Of the multiple possibilities, one has to be selected and then acted upon. As I said before, the particular action trajectory that is selected serves as a constraint upon the entropy production of the event, guiding and honing it towards its conclusion. The multiple possibilities of the environment constitute a search space -- a global informational entropy – that must be reduced. It is precisely the choice of a particular action that narrows the space to a more determinate trajectory through that space. This implies a Shannon-type reduction of uncertainty. The action is pulled into its future by the speaker's desire for coffee in the service of contributing to the self-maintenance of the agent's far-from-equilibrium condition. Difference plays a productive role in the driving of this process. Differences, e.g., the desire of the person wanting a tall cappuccino and the readiness of the salesperson in Starbucks, mesh productively with each other so as to yield new connections in the assemblage as a whole. The productive differences between the two individuals – the desire of the customer and the readiness of the salesperson – constitute an intensive difference (Deleuze 1994) when the two individuals are meshed in a new assemblage constituted by the interaction between them. This assemblage is, as Deleuze would say, in the first instance a continuum of intensity that is progressively differentiated as capacities are exercised and tendencies released (DeLanda 2011: 91).

Some normative functional capacities of utterances are:

i. Operate on and activate implicit presupposed systems of representations through apperception;
ii. Transform the social status of participants;
iii. Coordinate actions, perceptions, understandings, etc. between persons;
iv. Locate and track individuals (human and nonhuman) over space and time;
v. Modulate the bodily feeling states of persons;
vi. Enact social-affective-cognitive assemblages of persons and other classes of individuals by coupling languaging activity to the material world;
vii. Catalyze flows of socially organized matter and energy in social-affective-cognitive assemblages;
viii. Direct attention of persons and create joint attention frames between persons;
ix. Create virtual experience of cultural entities and events that are constrained by an attunement to linguistic patterns.

Persons qua living systems are recursively self-maintenant systems. Languaging behaviour is one of the sub-personal behaviours that persons have so as to maintain themselves over time through their transactions with their environment. Davia (2006: 260) argues that if we see that

in living systems there is no distinction between function and metabolism, then the notion of function becomes redundant. If we take this step, we can focus on languaging behaviour as an aspect of the biological process – a culturally constrained aspect – whereby persons and social groupings of persons metabolize their self-maintenant relationships with their environment. The material dynamics of languaging behaviour, not abstract form, is in itself a key means whereby humans catalyze their experience of their worlds. On this view, the properties of the material dynamics constitute metabolic processes that serve to maintain the person's relationship with the world.

9. The Intrinsically Expressive and Cognitive Character of First-order Languaging Dynamics

How does orderly, patterned inter-individual languaging behaviour that is shaped by culture arise? How do globally coherent representations of environmental phenomena arise? How can we explain language as intentionally and affectively modulated behaviour between persons? What is the place of living, feeling, moving human agents in such an explanation? If language is reduced to input/output processing models, we will never develop satisfactory answers to these questions. The view from dynamical systems theory helps to provide some new perspectives on these questions. Languaging behaviour is self-organizing in the context of the larger system of human agents embedded in their environments (Thibault 2004b: 48-49). Rather than positing internalized codes or rules, abstract social conventions, and input/output processing models as the basis for explanation, languaging is a self-maintaining process that exploits gradients of uncertainty both in the environment and in the internal milieu of individual organisms (Hollis, Kloos and Van Orden 2009: 212). In languaging behaviour, a number of potential states are possible at any given moment until one particular outcome is selected by circumstances. The set of potential outcomes is determined by the constraints that are in operation. Constraints may exist in the agent or in the environment, but they do not determine actual behaviour. Accordingly, actions are graded as more or less likely. Hollis, Kloos, and Van Orden (2009) make the following relevant observations:

> On this basis, we equate the states of gradient potential sets with affordances and effectivities in cycles of action and perception (Davia, 2005; Gibson, 1977; Swenson & Turvey, 1991). *Affordances* are descriptions of the environment directly relevant for action, with reference to an organism and its effectivities; *effectivities* are descriptions of the organism directly relevant for action, with reference to an environment and what it affords (Turvey & Shaw, 1995). To this we would only add that *gradient potential states* are also descriptions of the history of an organism, directly relevant for action, with reference to its immediate future.
>
> (Hollis, Kloos, and Van Orden 2009: 216)

Utterances are patterns of behaviour that have the potential to change the organism's relationship to its environment. Such patterns persist in space and time and therefore have the ability to exert consistent selection pressures on a given population of interacting individuals (Reed 1996a: 30-31). On this view, meaning is the result of the individual's efforts to come

into a relationship with the affordances of the environment when the individual interacts with those affordances. A key concept here is *interactivity*. Interactivity is in the first instance the active sensori-motor exploration of the affordances of one's world. As Gibson (1983 [1966], 1986 [1979]) showed, perception is an active process of seeking stimulus information in the environment. This information has the capacity to change the organism's relationship to its environment. Perceptual exploration is a mode of searching for information that constrains the information that is detected. In interacting with the environment, the organism enacts changes in the organism-environment system that yields differentiations that may be useful to the organism and its further interactions with the environment. Differentiations are internal interaction outcomes that guide further activity and provide a basis on which context-dependent apperception con proceed (Bickhard 1998: 198). Perceptual exploration is a specific kind of interaction with the world that yields stimulus information that is picked up by anatomically specialized perceptual systems (Gibson 1983 [1966]).

Utterances too are interactive in the sense that they operate on some aspect of social reality and transform it (Bickhard 1998: 205). However, many interactions are not confined to specialized sub-systems such as the classical perceptual systems. Perceptual systems are just a sub-class of a wider range of possibilities for exploring the environment in order to differentiate it. Exploratory activity is a means of modulating the environment in order to differentiate it. Languaging activity is an extension of this basic fact, not an exception to it. For example, vocal tract action is a means of acting on and transforming the social world in ways not necessarily tied to the immediate environment. Vocal tract gestural activity is a way of modulating the environment of other persons in order to obtain information about them as well as to allow for the tracking of apperceptions of social realities across diverse time and place scales. Lexicogrammatical patterns are second-order cultural constraints on languaging activity that provide conventionalized solutions to the problem of coordinating the diverse understandings that agents have both of the situation they find themselves in and of each other's understandings of the situation.

The search space of all possible ways to interact with some phenomenon is a virtual continuum of unactualised potentialities until one is selected by a particular search path. This virtual continuum of unactualised potentialities is nevertheless weighted such that some pathways are more likely than others. The continuum of virtual possibilities is value-weighted so that the continuum realizes what Hollis et al (2009: 216-217) call a "gradient of uncertainty": potential states are graded according to their likelihood of being realized.

> Gradients of uncertainty are like energy gradients in physical systems in several ways (however, see Keijzer, 2003; Turvey & Shaw, 1995). Each gradient is a potential set for action that includes the likelihood of respective actions without fully prescribing the particular action that will occur.
>
> (Hollis et al 2009: 216-217)

Exploratory activity changes the content of uncertainty gradients (Hollis et al 2009: 217). A particular body movement or shift in orientation, no matter how subtle or fine-grained,

changes the uncertainty gradient. Each time an act of perceptual exploration occurs, new potentialities for action and perception are created that change the potential set. Hollis et al point out:

> Actions perpetually update potential sets, which ensure a locally unstable system - a system close to a critical state. *Critical states* are states in which oppositional "forces" (constraints that favor one or another available outcome) are precisely balanced against each other. Critical states are thus a kind of boundary between qualitatively different behaviors.
>
> (Hollis et al 2009: 217)

These constraints are the embodiment of a population's history of exploring the search space and the action trajectories it adopts to explore that space. Successful action trajectories are selected and consolidated on the basis of habit. Habit is the principle process whereby a system acquires definite boundaries and contours that are maintained through time. The habitual repetition of an action thus integrates present and past and thereby gives rise to the anticipation of potential futures that can reliably be expected to occur when the given action is performed. Habitual patterns arise through a history of success in dealing with potential future interactions. In time, these habitual patterns become more valued such that they are selected as conventional routines on account of their success in dealing with the search for the information that the organism requires for the self-maintenance of the far-from-equilibrium conditions essential for its integrity and survival. These nonholonomic constraints – the emergent conventional patterns assembled through a history of organism-environment interactions – are control parameters that constrain and direct the flow of dynamical pico-scale bodily events in languaging behaviour.

10. The Auditory System is calibrated for the Perception of Source Events via Higher-order Acoustic Values

Agents perform the activities of articulating and/or perceptually exploring environmental events, e.g. vocal tract gestures. In doing so, they deploy bodily skills that allow them to integrate certain activities – e.g. optical or auditory scanning – with the given event. The (rapid) movement of our eyes along a given graphic arrangement of notational items on, say, a page gives rise to a changing array of optical stimulation in time. We perceive and acquire sensitivity to the patterns we detect on the basis of our active engagement with environment events. Perception-action invariants therefore reside in the relationships between an organism and its environment. For example, we learn that the visual-spatial stimulus information that we pick up is constrained as a function of the movement of the eyes, the head, and the torso through space, in the process of optically scanning the given optical array. The information that we pick up covaries with the movement of our eyes, our head, and the orientation of our bodies as we scan the sequence. Similarly, perception of phonetic gestures qua environmental events is dependent upon our active orientation to and exploration of the acoustic array of stimulus information that is caused by someone's vocal tract activity during the act of speaking. The time-locked perceptual exploration of the acoustic properties of speech (and

other acoustic) events is also motivated by the behaviour of the source that caused the acoustic event (Warren and Verbrugge 1981: 224).

Warren and Verbrugge (1981: 225) refer to experimental results that "support the general claim that sound in isolation permits accurate identification of classes of sound-producing events when the temporal structure of the sound is specific to the mechanical activity of the source … " (Warren and Verbrugge 1981: 225). They further point out: "If higher-order information is found to be specific to events, while values of lower-order variables per se are not, then it may be more fruitful to view the auditory system as being designed for the perception of source events (via higher-order acoustic functions), rather than for the detection of quasi-stable sound elements." (Warren and Verbrugge 1981: 225). In other words, the identification of the source event and its behaviour is the primary task of the auditory system. Higher-order acoustic functions are values that we perceive in source events such as a person's vocal tract activity. The higher-order perceptual information specific to that event can be used by the listener during the course of the event to assess, for instance, the speaker's feeling-state as well as the feeling which the speaker's vocal (and other) activity induces in the listener (Sections 11, 24). Or the higher-order information may specify and provide off-line access to experience of virtual cultural entities (Sections 23-25).

Bodily dynamics are intrinsically expressive and cognitive. Persons respond to the real-time dynamics of languaging behaviour as both expressive and cognitively salient in their worlds. Unlike code and symbol processing views, the theorist can focus on how agents hear and feel voices and respond to dynamics. Individuals align with and orient to affect, feeling, and aspects of events through expressive dynamics such as pitch, cadence, rhythm, tempo, duration, loudness, etc. Moreover, semantic priming shows how individuals dialogically coordinate and align their vocalizations in ways that give rise to the creation of shared semantic categories and values. Coordinated first-order languaging behaviour biases and shapes perceptions. It is by means of the emergence of social strategies in the form of inter-individual interaction routines that persons are able to partition their worlds on the basis of digital semantic categories (Ross 2007). A fundamental question, therefore, is: How do the bio-physical properties of languaging behaviour enable persons to partition their worlds using digital semantics in order to modify what persons perceive, orient to, and value?

In the following Section, I shall examine some possible answers to this question.

11. Body Dynamics, Felt Experience, and Values: The Feeling of What Happens

Time-locked and continuously changing bodily dynamics in interaction enable individuals to co-synchronize and to coordinate their actions and perspectives with those of others. Dynamics, not hypostatized symbolic forms, are what catalyze linguistic cognition. Cognition arises out of coordinated languaging behaviour as diverse factors are semantically synthesized in linguistic catalysis and integrated to core consciousness (Damasio 2010: 205-207). Change and dynamics are the key factors here. Language dynamics can induce in the self an emotional reaction that "alters the master interoceptive maps, a modification of the proto-self

ensues thus altering the primordial feelings." (Damasio 2010: 205). Damasio continues, as follows:

> Likewise, the sensory portal components of the protoself change when an object engages a perceptual system. As a consequence, the regions involved in making images of the body are inevitably changed at protoself sites—brain stem, insular cortex, and somatosensory cortices. These varied events generate microsequences of images that are introduced into the mind process, by which I mean that they are introduced into the image workspace of the early sensory cortices and of select regions of the brain stem, those in which feeling states are generated and modified. The microsequences of images succeed each other like beats in a pulse, irregularly but dependably, for as long as events continue to happen and the wakefulness level is maintained above threshold.
>
> (Damasio 2010: 205-206)

Continuously varying language dynamics alter the "sensory portal components of the protoself" (Damasio 2010: 205) such that the micro-sequences of images that are fed into the mind process have a first-person feel (see also Cowley 2006, 2007, 2008: 331, 333), which Damasio has designated as "the feeling of what happens" (Damasio 1999). Damasio (2010: 2007) stresses the likely central role played by timing as continuously varying dynamics instigate changes in the protoself, as follows:

> Timing is likely to play a role here too, when the causative object begins to be processed and changes in the protoself begin to occur. These steps take place in close temporal proximity, in the form of a narrative sequence imposed by real-time occurrences. The first level of connection between modified protoself and object would emerge naturally out of the time sequence with which the respective images are generated and incorporated into the cortege of the mind. In brief, the protoself needs to be open for business—awake enough to produce the primordial feeling of existence born out of its dialogue with the body. Then the processing of the object has to modify the varied aspects of the protoself, and these events have to be connected to each other.
>
> (Damasio 2010: 207)

In first-order interaction, one agent's dynamics affect or inhibit how the other perceives, acts and orients. The other modifies his or her responses on the basis of experience-based sensitivity to aspects of events as they unfold in a narrative-like sequence that is imposed by the time-locked character of the unfolding dynamics. This establishes a first-level connection between the protoself and first-order languaging dynamics. Felt experience is narrative-like. Schögler and Trevarthen (2007: 291) have shown how the changing dynamics of inter-individual movement patterns, consisting of coupled shifts in posture, orientation and vocalizations, engage participants and observers with temporally and interpersonally coherent and intentionally modulated mimetic displays that elicit 'narrative' interpretation (see Thibault 2011 for further analysis and discussion). As agents modify the environment of each other's actions through their inter-individual dynamics, they hear, see and feel what their

languaging behaviour means as its dynamics continuously modify aspects of the protoself in core consciousness.

Prosody, for example, thus continually perturbs core consciousness such that the resulting modifications of the proto-self's primordial feeling states "now become differentiated feelings of knowing relative to the engaging objects." (Damasio 2010: 209). Cognition dawns. In becoming "differentiated feelings of knowing" with respect to the dynamics and their relations to aspects of the world (the situation, the environment) with which agents co-engage in their interaction, the "objects of knowing" are made salient and are assigned values at the same time that they are placed in the felt self-perspective of the agent in core consciousness. Language dynamics therefore bias action and perception in self and other in ways that are value-weighted. This gives rise to semantic priming, based on how agents coordinate aspects of the dynamics with aspects of situations. In time, this is how we perceive wordings in dynamical languaging behaviour when we learn to attend to syllabic patterns and coordinate them with objects and events – both actual and virtual -- with which they covary in a given population of languaging agents.

12. The Sensori-motor Basis of Sign-making

The pre-linguistic vocalizations and related behaviours of infants are, of course, frequently and primarily motivated by the need to fulfil basic metabolic and physiological needs and requirements of the organism. For instance, babies will cry when they are distressed, feel hungry, or are in pain (Stark et al 1974; Demos 1988: 33). Many aspects of infant vocalizations are not exclusively motivated by metabolic processes, but by changes in the cognitive and communicational dynamic of the dyad linking caregiver and infant as a distributed cognitive-interactional system. The dyad itself is thus operated on and transformed in ways that may also lead to the satisfaction of the infant's metabolic requirements. However, the operations and resulting transformations occur at the higher level of the dyad in ways that are not fundamentally metabolic, but cognitive. Infant vocalizations are, with respect to metabolic processes, second-order processes. These may certainly play a role in the regulation of the infant's metabolism (getting food, warmth, etc.).

By the same token, the sensori-motor patterns generated by the infant's vocal and related bodily activity play a key role in extending the infant's agency beyond its body and of coordinating the infant with selected aspects of its world. Vocal activity is therefore about more than the self-regulation of first-order metabolic processes of the organism. It is also a second-order form of sensori-motor coordination with the world beyond the organism. The prosodic patterns of the infant's vocal activity are behavioural patterns that enable the infant to tune into and to interact with forms of organization in its environment, notably the behaviour (vocal and otherwise) of caregivers (Cowley et al 2004). Early infant semiosis is dependent upon the embodiment of the agent. The infant's embodiment affords a repertoire of sensori-motor behaviours such as vocalizing, exchange of gaze, smiling, facial expressions, etc. that are constitutive of the infant's interactions with his or her world.

The dialogically coordinated character of the infant's engagements with others from the earliest moments of life and their constitutive basis in the infant's sensori-motor coordinations with the environment (the dyad) suggest that first-order languaging has its developmental basis in earlier forms of exploratory activity in pre-linguistic caregiver-infant dyads (Halliday 1975; Cowley et al 2004; Thibault 2005c). It is a form of exploratory activity aimed at seeking and obtaining information about the dyad by acting on it and getting responses from the other. Utterance-activity embodies an anticipatory dynamics in the form of stances of attention and expectation formed by intentional dynamics in the limbic system (Freeman 2000). The entire sensory cortex is therefore primed to be in an active state of anticipation. Sensori-motor behaviour is exploratory because it seeks and obtains responses from the environment of the agent – responses that confirm or deny these hypotheses formed in the intentional dynamics of the limbic system. These hypotheses are tested by "state transitions giving amplitude modulation (AM) patterns that converge into the limbic system, in mammals into the entohinal cortex." (Freeman 2000: 97). The formation of a new hypothesis primes the agent for a range of possible responses, each with its means of being effected as sensori-motor action in the world.

This is, I argue, the foundation of all dialogically coordinated processes of cognition and communication. The infant's sensori-motor coordinations with its world place tight constraints on the kinds of explanations we can come up with. Given (1) that human organisms are complex, dynamic, self-organizing systems in far-from-thermodynamic-equilibrium conditions and (2) that they actively maintain themselves in this statistically improbable state by continuously producing and synthesizing the processes that maintain them along their time-locked trajectory, we can say that the infant's sign-making is not a progression towards an evermore sophisticated means for encoding and decoding the world. Instead, in exchanging matter, energy, and information with their environments through their active engagements with it, infants learn to participate in and to semantically synthesize sign-making behaviours that come under more and more cultural constraints in the course of their development as persons (Cowley et al 2004). They do so by actively exploring their worlds and obtaining information about it in the form of the responses to their own and others' activity. Signs do not encode inner and outer states of affairs qua representations of an independent world; instead, human agents (infants, adults), manufacture signs for themselves and for others through the catalyzing actions of their languaging behaviour. They integrate and semantically compress information about their worlds from multiple sources and sensory modalities (Cowley 2008). This takes the form of sensori-motor coordinations and couplings that have the potential to bring about adaptive behaviours or orientations to behaviour that change the agent's relationship to its world (Reed 1996a).

In the following Section, I shall discuss more fully the catalytic nature of linguistic processes in relation to recent development in the biological sciences.

13. The Catalytic Nature of Languaging Behaviour

Davia (2006) has proposed a new theory that grounds biology and neuroscience in a general theory of the process of catalysis. In the following Sections, I shall further explore the

implications of this general theory for our understanding of language. Davia's hypothesis is that at every scale, "living processes are processes of catalysis, and that all biological processes mediate transitions in their environments, employing the same mechanisms as enzymes." (2006: 255). Enzymes are the prototypical catalysts: autocatalytic reactions occur when an enzyme catalyzes a chemical reaction without itself being changed (Davia 2006: 261). Following the theory of autopoesis developed by biologists Maturana and Varela (1980, 1987), Davia has questioned the view that cognition represents information about an independent world 'out there' that is external to the organism. The theory of autopoesis proposes instead that organisms create their worlds through their interactions with it.

Davia's theory of catalysis goes further and argues that the organism and its environment are unified by the mediating or catalyzing role played by solitons or travelling ways in relation to its environment. Carpenter and Davia (2006: 1081) draw attention to extensive neurobiological research that shows "the wave-like or resonance-like nature of perceptual experience itself." (Carpenter and Davia 2006: 1081). Carpenter and Davia further remark:

> These observations may reflect a single, unifying principle, namely, the wave-like processes are the way by which living organisms mediate (catalyze) their environment, and they relate directly to the organism's experience (Davia 2006).
>
> (Carpenter and Davia 2006: 1081-1082).

Organisms catalyse or mediate their environment by means of wave-like processes, including the languaging behaviour of human agents. The environment does not exist independently of the organism, but is a result of the organism's catalyzing activity. Organism and environment comprise a unified organism-environment system on account of this catalytical activity.

According to Carpenter and Davia ((2006), the nervous system is an excitable medium on account of the metabolism of glucose and other nutrients:

> We argue that the brain is best understood as an excitable medium as a consequence of the metabolism of glucose and other essential nutrients. This energy gradient is dissipated by neural activity that is constrained by the organism's structure, including that arising from its history and ongoing activities. Thus, the non-linear neural waves can be seen as a self-maintaining and self-sustaining dynamic, a solution to the boundary conditions implicit in the structure of the organism and the relation between the organism and its 'environment.'
>
> (Carpenter and Davia 2006: 1082)

Carpenter and Davia (2006: 1083) generalize catalysis from the prototypical instance of enzyme catalysis. A soliton-like wave is "a localized, nonlinear solution to the boundary conditions that constitute its environment." (Carpenter and Davia 2006: 1083). On this view, catalysis on all scales (enzyme, cell, organ, and organism) involves the overcoming of structural constraints to dissipate energy (Carpenter and Davia 2006: 1083). Boundary conditions constitute a problem space of what Deleuze would call "virtual multiplicities"

(Deleuze 1987: 23-25; 1994: 212) in the sense that "the virtual possesses the reality of a task to be performed or a problem to be solved" (Deleuze 1994: 212). In this sense, catalysis is a solution to a problem space of virtual possibilities. In individuating a specific organism-environment relation grounded in experience, a solution to the problem posed by the current balance of constraints is individuated.

Carpenter and Davia argue that a living organism is a unitary process of catalysis at all fractal levels of, for example, the human brain and body, seen as media of catalysis. A soliton-like wave is a disturbance of a medium that propagates through that medium, which is able to replenish its energy. A medium of this kind is what Carpenter and Davia define as an "excitable medium." On all levels, soliton-like waves facilitate these transitions. On this view, catalysis is generalized and applicable to many scales of organismic organization from enzymes, cells, organs, and organisms (Carpenter and Davia 2006: 1083): "A living organism can be understood as a unitary process of catalysis, mediating its environment." (Carpenter and Davia 2006: 1083). Extrapolating, languaging activity consists of soliton-like waves of a unitary process of linguistic catalysis comprised of waves of neural and sensori-motor activity. In the first instance, phonetic gestures are time-locked behaviours that exhibit patterned, wave-like properties. Phonetic gestures disturb the medium of air in structured ways and are propagated through the air to impact on the excitable media of an observer's brain and body. Furthermore, this applies across all scales of the organism-environment system as massively parallel and inter-connected processes from cell, to organ, to organism, to organism-environment system. The brain does not solely determine or control this process, which depends on structural constraints arising from a history of the organism's interactions with its environment, including interactions with other persons, perception-action invariances, normative constraints and values emanating from cultural timescales, contingencies of the situation, and affective and other subjective and inter-subjective gradients.

The perceptual pick up of first-order languaging activity and its higher-order invariances constitutes what Carpenter and Davia (2006: 1083) call a "unified dynamic." This dynamic, Carpenter and Davia continue, "is a solution to the boundary conditions arising from our eye movements, our head movements, and so forth, interacting with impinging energy (see also Noë 2004). Our experience manifests the perception-action invariance as a unified event that is more thermodynamically stable than non-unified patterns." (Carpenter and Davia 2006: 1083). A linguistic event is only implicit in a discrete set of statistical invariances in a phonetic gesture and consequent acoustic array in the environment. The unity of such an event is implicit and depends upon the organism's ability to confer structure on it through its own catalyzing activity as the perceiver orients to, attends to, and perceptually explores the event. The organism's active exploration of the sensori-motor contingencies of the event therefore is a catalytic process that is dependent on the way in which the organism actively experiences the event as a complex, multidimensional neural standing wave that is associated with the perception of the event. The ability to discern the implicit linguistic structure of such an event depends on a history of the organism's culturally constrained experience of such events.

According to Davia, the nervous system is an excitable medium, i.e. a landscape replete with energy that can be consumed and replenished (see above). The brain is just such a medium in which glucose is consumed and replenished. A soliton or travelling wave is a time-locked invariant structure that emerges and exists as a coherent, ordered entity on some scale within an excitable medium. There is, then, no such thing as an environment that is independent of the organism. Carpenter and Davia point out that motor activity is "what an observer perceives of the organism's catalytical process." (2006: 1083). Carpenter and Davia provide the following example:

> Thus, a millipede's solitonic motion as it crosses a sand dune (to morph Herb Simon's classic example of the ant) is an observer's perspective of the millipede catalyzing an aspect of its environment. These findings, if generalized, may point toward a non-reductionist reconciliation of the observation of waves in physiology and the observation of waves in behavior; that hypothesis, however, requires further development.
>
> (Carpenter and Davia 2006: 1083)

Languaging events clearly consist of behavioural waves that can be observed from diverse points of observation in a given dialogical array (see Hodges 2007, 2009 for this term). Seen from this perspective, it is the speaker's catalyzing an aspect of his/her environment. Gibson's (1986 [1979]) ecological theory of perception identified invariances that arise from the organism's interactions with its environment. Invariances are not 'out there' in the world, but are the relations of the organism's perception-action. O'Regan and Noë (2001) developed their theory of sensori-motor contingencies in order to explain how the organism makes these invariance relations explicit through its experience. Invariances are detected through the organism's active exploration of and interaction with the events and objects that constitute its environment. First-order languaging activity is bodily activity that dynamically changes the environment. It is a mode of performatory activity that makes information available to other agents in the environment (Reed 1996a: 155) and exploits the potentiality to affect other individuals (and be affected by them). It is also and simultaneously a mode of exploratory activity that seeks information from the environment by engaging with others and eliciting responses from them – responses that provide information about the agent and the agent's world to others in the dialogical array.

For example, how the listener/observer hears and feels the dynamical properties of the speaker's utterance-activity dynamically changes the listener/observer's bodily feeling states, her body dynamics, her focus of attention, and locus of cognitive processing in ways that are detected by and affect the first speaker. These changes provide information to the speaker concerning the listener, the listener's affective states, her intentional orientation, and so on. Agents' perceptions depend upon the action they perform. Perceptual experience is created or enacted as a result of the organism's active engagement. The invariants that are detected crucially depend on and acquire content for us as a result of what Noë calls "sensorimotor knowledge" (2004: 9). Noë elaborates on this notion as follows:

> Genuine perceptual experience depends not only on the character and quality of stimulation, but on our exercise of sensorimotor knowledge. The disruption of this knowledge does not leave us with experiences we are unable to put to use. It leaves us without experience. For mere sensory stimulation to constitute perceptual experience – that is, for it to have genuine world-presenting content –the perceiver must possess and make use of *sensorimotor knowledge*.
>
> (Noë 2004: 10)

We access and scan articulatory events through movements of, for example, our eyes, our ears, and our head as we attend to and perceptually explore such events. Moreover, much of this movement is self-actuated in response to the changing dynamics of the event. It is through self-movement that we actively probe and gauge the event. In this way, we learn what the relevant patterns of sensorimotor dependency are (Noë 2004: 13). The dynamical properties of such events do not necessarily yield a single or unitary visual, auditory, tactile, or other experience. In languaging too, their properties arise through our active exploration, which provides information concerning things like syllabic patterns, pitch, tempo, rhythm, degree of loudness, tone, and so on. In this way, we obtain information concerning the organization of the event on the basis of the ways in which perceptual stimulus information covaries with our exploratory activity (auditory scanning, positioning of the ears, head turning and adjusting to track the event, and so on).

14. Linguistic Catalysis and Semantic Synthesis: A Biologically Grounded Alternative to Language as Constructed Code

In the tradition of ecological psychology founded by Gibson (1977, 1983 [1966], 1986 [1979]), Verbrugge (1977, 1985) had already argued that linguistic utterances are catalysts that have the functional capacity to either activate or inhibit flows of affect, cognition, action, and so on, in both self and others (see Thibault 2011). Enzymes are the prototypical biological catalysts (Section 13). The catalytic process greatly increases the rate at which molecular reactants form a thermodynamically more stable product. The products of this process are thermodynamically more stable than the reactants.

> In enzyme catalysis, the reaction ultimately occurs because the product(s) is/are more thermodynamically stable than the individual reactants. The catalytic process facilitates the transition from the reactant(s) to the product(s) by overcoming the structural constraints of the reactants' structure and dynamics. Research suggests that catalysis takes advantage of the invariance (symmetries) of the biological structure (the protein-substrate complex) to deliver energy where it is needed to change the molecular structure. The process appears to be 'vibrationally-assisted,' a wave-based facilitation that involves a type of localized, non-linear wave, called a *soliton*.
>
> (Carpenter and Davia 2006: 1082)

Something analogous occurs on the very different organism-environment scale in utterance construal. Here I draw on Cowley (2008), who points out that changes in how we report the dynamics of talk are not reports of objectively heard patterns, but reports on how we feel the dynamics in core consciousness (Cowley 2008: 329; Sections 11, 14, 24, 30). Dynamical languaging events are, in part, coordinated waves of vocal tract and other bodily activity that are propagated into the environment when they perturb and restructure environmental media such as air and light. They are the perceived dimensions of the agent's languaging behaviour as the agent catalyzes an aspect of his or her environment. This dynamical wave-like activity triggers events that tend to settle on a stable pattern in a process of *semantic synthesis*, not encoding/decoding. This process is akin to Barbieri's (2003) organic process model of protein synthesis (Cowley 2008: 330-331). However, we do not only hear lexicogrammatical patterns; we also hear voice dynamics (e.g. prosody). Construal of a linguistic utterance as a stable percept integrates diverse factors during the catalyzing process. These factors include: lexicogrammatical constraints (e.g. co-occurrence restrictions) and lexicogrammatical patterns heard as 'wordings'; voice dynamics; recent events held in working memory; prior experience; and speech timing. To quote Cowley:

> … utterances set off events that stabilize on a pattern. In settling on a sense, participants exploit dual recognition. While hearing wordings, orientation (or reorientation), they also use prosody. As Rączaszek et al (1999) emphasise, symbols constrain dynamical events (c.f. Pattee, 2000). Using dual recognition, dynamics are integrated with a heard pattern. Organic coding thus offers a model of how we use connotations. Neural 'openness' uses recent experience and timing to integrate recognition that manufactures cognitive-communicative fusion. While akin to Cowley's (2006) felt response, in the work of Rączaszek et al (1999), the parallels with organic coding apply to neural functioning. The percepts that are identified become manufactured syntheses that arise as artefacts (utterances) prompt brains to fill out incomplete information.
>
> (Cowley 2008: 331)

The semantic synthesis of a self-organizing semiotic artefact (a linguistic utterance) depends on dynamics and timing. Following Pattee (2000), the fusion of symbols (e.g. wordings) with dynamics means that symbols constrain the dynamics. Symbols, according to Pattee, are rate-independent as distinct from rate-dependent phenomena such as the material dynamics. When symbols are integrated to dynamics, the resulting linguistic catalysis works to either amplify or to inhibit flows of cognition, affect, feeling, coordination, and action in self and others (Verbrugge 1985). Phonetic gestures qua behavioural waves and their physical properties do not stand in a correspondence-type relation to the semantic synthesis mentioned above. Instead, our experience of the semantic synthesis of diverse factors corresponds to specific patterns of neural activation (Maturana and Varela 1987: 22), not with the behavioural wave and its specific, measurable properties per se. The resulting meaning is unique to the individual just as the specific patterns of neural activation are. The physical form of the wave does not have meaning in itself; it elicits the construction of meaning in observers (Freeman 2000: 14-15). Maturana and Varela (1987: 22) show that the phenomenal experience of seeing

the colour green "can be triggered by a number of different light perturbations" such that "we can correlate our naming of colors with states of neuronal activity but not with wavelengths." (Maturana and Varela 1987: 22). The "states of neuronal activity" – neural standing waves – associated with the experience of seeing a particular colour are triggered (not caused) by "a number of different light perturbations." (Maturana and Varela 1987: 22). The same observation is valid for perceptual experience generally.

For example, the experience of hearing a particular wording in a given phonetic gesture can be triggered by different properties of the sound wave in different persons. A particular wording does not correspond to fixed physical properties of the sound wave, as analysed in different ways by phonetic notation, spectrographic analysis, PRAAT-type acoustic analysis, and so on. The soliton-like wave of languaging behaviour catalyzes a self-organizing, time-locked process of semantic synthesis of diverse, previously loosely structured factors. These factors include lexicogrammar qua virtual patterns from cultural timescales, felt dynamics, first-person experience, expectations, contents of working memory, and physical events (e.g. vocal tract gestural events). These factors are synthesized as semantic artefacts when brains are prompted to make explicit otherwise implicit and incomplete information (Cowley 2008: 331; Carpenter and Davia 2006: 1084). If the invariances are not 'out there', objectively speaking, in the world, but are in the relations of the agent's perceptions and actions – in the agent's sensori-motor contingencies – then it is, as Carpenter and Davia point out, the agent's experience that makes them explicit through its exploratory activity (Carpenter and Davia 2006: 1083; see also Maturana and Varela 1987: 22). The particular patterns of neural activation that are triggered by a particular phonetic gesture and therefore the particular semantic synthesis that is experienced by that person depend upon the intrinsic properties and structure of that person – the person's internal complexity – at any given time as well as that person's history of experiencing such events. Socially coordinated languaging behaviour is a way of concerting the unique, experience-based meanings of individual persons through attunement to gestural structures that afford processes of successive approximation of each other's meanings so that person's can assimilate to the extent that is necessary their own neural structures to the meanings of others (Freeman 2000: 15).

Brain and body are the media of catalysis in linguistic cognition. As the discussion above shows, linguistic utterances are not ready-made, unified events, e.g. form-meaning pairings or codings. The processes of semantic synthesis described by Cowley arise by virtue of the agent's experience, by neural standing waves. We do not in the first instance perceive a continuous, unified dynamic in the form of an already constituted linguistic event. Our perception of a dynamical languaging event, e.g. vocal tract gestural activity, is an exploratory activity that takes place in time. Soliton-like waves of neural and bodily activity unify the boundary conditions arising from the organism's interactions with its world. The localized, robust qualities of phonetic gestures qua soliton-like behavioural waves afford their active, reciprocal perceptual exploration by observers on the time-scales of dialogically coordinated utterance-activity between persons. Moreover, their properties are non-arbitrarily related to and embedded in the invariances and boundary conditions of the environments in which they occur and in part constitute and respond to (Section 4). Unlike the idea that

language encodes or represents phenomena in an external world in terms of the categories internal to the language system, linguistic catalysis is one instantiation of a more general phenomenon whereby biological processes mediate or catalyze experience by channelling energy via structure (see above). Enzyme catalysis does it when:

> The protein chains of the enzyme may support soliton waves that alter the conformation of the enzyme-substrate complex, affecting the width of the energy barrier. The conformational change lessens the distance between specific parts of the enzyme and thereby lessens the distance between the molecular reagents that are bound to it. This shortening increases the possibility of 'quantum tunnelling' and increases the reaction rate.
>
> (Carpenter and Davia 2006: 1084)

A catalytic reaction occurs because the products of the reaction are more thermodynamically stable than the individual reactants. The catalytic process helps to bring about and indeed accelerates the transition by overcoming the structural constraints of the reactants' structure and dynamics (Carpenter and Davia 2006: 1082). Similarly, complex, globally ramifying patterns of oscillatory behaviours across many scales of neural and bodily dynamics in languaging behaviour constitute macroscopic coherent states of matter and energy that unify many fractal scales of the organism's organization as whole-body sense-making (Sections 30-32). This suggests that oscillatory patterns of neuronal activation are embodied aspects of the organism's catalytic processes rather than representations of an external world. Linguistic catalysis is facilitated by the wave-based properties of phonetic gestures. These properties have the capacity to exploit the invariances and symmetries of biological structure "to deliver energy where it is needed" (Carpenter and Davia 2006: 1082) to change that structure.

Freeman (2000) has described the transition to a new global state of coherent oscillatory patterns that ramify across the entire cortex in these terms. Meaning, Freeman argues, is the transition from one thermodynamically stable state to another; it is triggered by the processes of assimilation when the self adapts its brain and bodily dynamics to the waves of languaging behaviour of self and others during the process of perceptual exploration referred to above (Freeman 2000: 121). Freeman explains meaning in these terms as follows:

> Meaning emerges in sequences of global AM patterns of oscillatory neural activity coordinating the neurophil of an entire cerebral hemisphere. The amplitude of the oscillations is high in local patches and low in others, as viewed through brain imaging, in which, as with all patterns, both the highs and the lows are necessary. Each pattern is a construction of the brain, with onset and termination by global state transitions. The contributions to the pattern are both local and large scale. Local details are provided by synapses that have been modified by previous learning and now shape local bursts as they emerge in patches of the forebrain, including the primary sensory cortices, the limbic system, and the brainstem nuclei. The interactions of these patches with each other and with the brainstem create a global state, which organizes and constrains local

> activity in a process of circular causality. At any given time, this state –which is constantly in flux as individuals grow and learn—is the meaning in a person.
>
> (Freeman 2000: 143)

Global states of complex oscillatory patterns as described by Freeman unify or semantically synthesise the discontinuities between energy and the diverse factors mentioned above in relation to the processes of semantic synthesis. As previously noted, a phonetic event, corresponding to a complex, multidimensional soliton-like behavioural wave, does not constitute in itself a continuous, unified dynamic. The perception of the wave affords its active exploration by the observer. This means that time-locked neural waves associated with the sensorimotor contingencies involved in our exploration of that event constitute a unified dynamic, e.g. a semantic synthesis. This dynamic is an emergent, time-locked solution to the boundary conditions arising from our bodily exploration of such events – e.g. eye movements, ear movements, head movements, etc. – interacting with the energy of the soliton-like wave that impinges on the excitable media of the brain and the body as whole-body sense-making at all scales of organismic organization (Section 32). The physical properties of the behavioural wave amount to the complex intersection of virtual futures that is immanent in the complex of tendencies and forces that accompanies the event and which anticipates a process of resolution in the form of, e.g., actualized experience. The catalytic process actualizes in experience a solution to this virtual problem-space created by the multiple physical properties of the event. The phonetic event itself occurs in a complex, symbiotic relationship to aspects of its environment - the situation, the persons participating in it, their past experience (of such events), the feeling states of each participant, cultural constraints, and so on. The phonetic event is itself an aspect of a dialogically coordinated relationship between the persons involved. The ability to hear and make semantically salient certain kinds of patterns, e.g. "wordings," in such events, or to perceive and feel a "friendly" voice as distinct from a "neutral" or "unfriendly" one are actualized patterns catalysed by observers along specific vectors of interest, attention, relevance, motivation, feeling states, and so on.

The phonetic event qua soliton-like behavioural wave physically expresses a virtual multiplicity of patterns in the auditory array of perceptual information that is specified about the environmental event (the vocal tract gesture) that caused a disturbance in the array. The adopting of a 'language stance' (Cowley 2011) by an observer in the dialogical array actualizes a specific vector – let us call it the 'wording' or lexicogrammar vector' – as an actualized singularity by virtue of the catalyzing process of the observer. Semantic synthesis and the kinds of thinking, acting, feeling, and so on that it affords agents in linguistic catalysis are based upon these processes of the actualization of singular patterns from the virtual multiplicity that the behavioural wave constitutes. It is not the encoding of a pre-existing experience as a semantic content into linguistic form. Meaning and value do not reside in any of the components that are synthesized, but emerge in and through the processes of semantic synthesis that is linguistic catalysis. First-order languaging dynamics and the affordances of their physical arrangements (patterns) therefore afford virtual multiplicities that we selectively

actualize as particular, individuated singularities according to the balance of constraints – cultural, biological, situational, subjective, and so on – at play in particular situations.

As Cowley's discussion shows, the parallel with protein synthesis is useful because we can see how the higher-order percept – the perception-action invariants – is manifested by our experience as a unified event (a semantic synthesis) that is, as Carpenter and Davia 2006: 1083) put it, "thermodynamically more stable than non unified patterns." These non-unified patterns include the various factors mentioned by Cowley, e.g. wordings qua virtual patterns emanating from cultural timescales, the bodily feeling of what we hear, expectations, norms, physical events, and contents of working memory. The self-organizing semantic synthesis of linguistic artefacts entails the parallel operation of many different processes, rather than a linear or sequential ordering of encoding/decoding terminating in the reception of an encoded message. Instead, semantic synthesis involves parallel, relatively unstructured and loosely coupled processes that offer enhanced possibilities for innovation by altering, say, the coupling or timing relations between the diverse factors involved in the synthesis that I mentioned above with reference to the work of Cowley. Participation in cultural routines gives rise to globally coherent neural activity, which Carpenter and Davia characterize as a soliton-like wave "that maintains its organization, mediating the impinging transitions and giving rise to the organism's experience." (Carpenter and Davia 2006: 1083). It is the catalytic, mediating work of this globally coherent activity that gives rise to a unified semantic synthesis of, for example, the various factors mentioned above.

Unlike enzyme catalysis, linguistic catalysis is a bio-cultural process involving self and other(s) with complex internal organization. In linguistic catalysis, the addressee of the catalysis is, in many respects, causally responsible for the cognitions, feelings, actions, and so on, which may be catalyzed. Persons have internal dispositions, inclinations, and capacities which enable them to adopt and transform in the service of their own projects the observed behaviours, routines, ideas, and so on, of others. If enough people adopt the observed behaviour in this way, the catalytic process has a chance of spreading more widely though the population.

Soliton-like waves do not reduce to activity of the nervous system; they are emergent phenomena whose existence depends on the current balance of constraints. Constraints on languaging activity derive from biology, culture, a history of interactions grounded in first-person experience, norms, current circumstances, the situation, the current states of the agent, and so on. Constraints hone and fine-tune the possibilities of cognition and action before it occurs by narrowing down or restricting the many degrees of freedom so that orderly activity can occur (Hollis, Kloos, and Van Orden 2009: 214). Constraints do not directly cause behaviour. Instead, they alter the probability of particular behaviours occurring by narrowing down the potential set to a smaller one consisting of fewer possibilities. On the other hand, the development of new skills and capacities means that prior constraints are relaxed and the degrees of freedom are increased such that the cognitive and contextual reach of the organism is extended. The potential for the soliton-like waves that characterize languaging behaviour is therefore dependent upon the agent-environment system, rather than being located in either

the agent or the environment per se. Capacities, dispositions and constraints may be located in either of these, but none of these as such causes behaviour. Capacities and constraints narrow the probabilities of behaviours occurring, but actual behaviours are triggered by the contingencies of particular situations. Contingencies make particular behaviours happen. Capacities, dispositions, and constraints specify the probability landscape in which these can occur, but do not determine them.

The notion of linguistic catalysis suggests that language is not something we 'use' by instantiating tokens from a system of types or a code that we deploy to decipher an independent world. In animated first-order languaging activity, participants do not have a sense that they are separate from their languaging or that the objects and events that are coordinated with their languaging are separated from the self. Instead, they have a felt sense of being caught up in an event and its flow. The idea of language as an input/output machine consisting of different levels of coding localizes language to some specific performatory and receptive functions of the body or reifies it as 'text'. However, the catalytical view suggests that life processes at all levels are unified by self-similar fractal patterns on multiple scales of the entire body-brain system: languaging is whole-body sense-making in this sense. On this view, linguistic catalysis triggers large-scale complex oscillatory patterns in the nervous system that constitute large-scale unified states that potentially ramify across all scalar levels of the organism's organization and affect it (Section 32).

Subtle modulations in first-order languaging dynamics can alter the gradient of uncertainty in interaction (Hollis, Kloos, and Van Orden 2009: 216; see Section 9 above). So too can changes in intensity of affect, the bodily orientation of interactants, and aspects of situations, amongst other factors. First-order languaging dynamics and the reciprocities they create promote habits and routines in caregiver-infant dyads that lead to a process which Stern (2004) called "moving on." The processes of moving on lead to the creation of a repertoire of now-moments. As dialogically engaged persons such as infant and caregiver move along together, they enact and create a series of now-moments that have their origins in the narrative processes generated in core consciousness by the proto-self (Section 11). These brief narrative units lasting no more than a few seconds give rise to intersubjective moments of affect-charged "meeting." According to Hart, meetings of this kind "increase the ability of the nervous system to intensify and coregulate with someone else's activity, and they occur only when the infant's and the caregiver's nervous systems have been able to engage in mutual adjustment and self-regulation, the condition that Stern refers to as moving along." (Hart 2011 [2006]: 27). The now-moments reciprocally bind the caregiver's and the infant's neural and bodily dynamics in states of intersubjective entanglement (Section 32), resulting in mutual recognition and shared experience. Moments of meeting intensify the flow of affect through the entangled dynamics of the two agents. Meeting "promotes the nervous system's capacity for self-regulation and attention control." (Hart 2011 [2006]: 27). Moments of meeting, as defined by Stern and Hart, are, to quote Hart, "a dyadic expansion of the consciousness." (Hart 2011 [2006]: 27).

The intersubjective processes of moving on and the resulting now-moments described by Stern and Hart are the precursors of conversational dialogue. Dialogue can thus be seen as successive loci or rhythmic pulses of intersubjectively coordinated and deictically grounded experience in the form of short pulses or temporal intervals that are connected to each other along their trajectory. Lieberman (1967) showed the link between speaking and expiration in this connection. Expiratory airflow from the lungs is one of the parameters in terms of which dialogue is organised. The so-called "breath-group" is one of the factors that chunk speaking into temporal intervals that correspond on higher levels to different syntactic units – clause, phrase, and word – though in ways that tends to be language-specific rather than universal. Pulses of experience are catalysed in and through cycles of dialogically organized catalytic activities that loop between participating agents in dialogue and their environment(s). Dialogue is a succession of now-moments experienced in this intersubjective space-time that are deictically grounded as metricised rhythmic pulses or intervals of intersubjectively coordinated and deictically grounded experience that interactants co-orient to. Seen in this light, second-order lexicogrammatical resources of clause grammar and the resources for combining clauses into larger complexes (Halliday 2004 [1985]: chap. 7) are forms of cultural scaffolding that enable agents to solve problems of coordination as they attempt to move along together in dialogue.

15. Phonetic Events as Travelling Waves

Brain and body are the media of linguistic catalysis. Phonetic events resulting from vocal tract activity are not in themselves unified phenomenon. They are unified by the agent's experience. The neural waves associated with the perception-action invariances of an event of this kind constitute a unified dynamic. This dynamic is a solution to the boundary conditions arising from our ear and eye movements, head movements, and so on, interacting with the energy from the phonetic event as it impacts the organism. It is experience that registers the perception-action invariance as a unified event that is more thermodynamically stable than are non-unified patterns (Section 14). It is the skilful exploration of audible and visible patterns of languaging behaviour that give rise to coherent neural activity in the form of a soliton-like wave that maintains its organization as it mediates the transitions in the energy impinging on it and give rise to the agent's experience. How can we connect these observations to phonetic events? In the remainder of this Section, I make some preliminary observations directed to the answering of this question with reference to the gestural phonology of Browman and Goldstein (1992, 1995) and Fowler (1980, 2010).

Browman and Goldstein (1992, 1995) showed that phonological entities are temporally overlapping or co-articulated gestural activities of the vocal tract that they have called "phonetic gestures." Phonetic gestures are bio-physical actions, not abstract mental categories (Fowler 2010). Phonetic gestures are biomechanical activities of the vocal tract that create and release constrictions in the vocal tract in different ways and to different degrees. For example, the consonant /p/ is articulated by two gestures: (1) lip closure or constriction of the lips; and (2) a devoicing gesture constricting the vocal folds in the larynx. A phonetic gesture is a distinct unit of action that takes place in the speaker's vocal tract. For example, in the lip-

closure gesture in /p/ the two lips typically come together. This involves the movement of three independent articulators that all play their part in effecting this particular configuration of the lips, i.e. lip closure. These movements are: (1) the upper lip is displaced downward with respect to the lower teeth; (2) the lower lip is displaced upward with respect to the lower teeth; and (3) the raising of the entire mandible. The lip closure gesture is thus the result of a synergy of interacting factors that give rise to a functional unit called a coordinative structure (Turvey 1977).

During the time-bound performance of the lip closure gesture, the three movements mentioned here form a coordinative structure in the sense that the diverse articulatory movements actively synchronize with and modulate each other. The three movements are independent variables that can function in other structures, but in the lip closure gesture they form a cooperative coalition or synergy of the variables that lasts for the duration of the performance of the gesture (Fowler 1980). Lip closure is thus a unit of action that can be described microscopically as involving a very large number of interacting co-articulated muscular movements at the same time that it can be described macroscopically as a unit of action involving a much smaller number of variables required for the performance of a given task (e.g. lip closure) (Browman and Goldstein 1992, 1995).

A defining characteristic of the articulatory phonology of Browman and Goldstein (1989, 1992, 1995) is the view that the goals of speech production are defined by local constrictions effected by specific organs of speech in contrast to the traditional view in linguistics that the goals of speech actions are defined in terms of the properties of the sounds that result from the actions of vocal tract gestures. Browman and Goldstein (1989) argue that organs constitute an intrinsic partitioning of possible vocal tract gestures into categorically distinct types. For example, lips and tongue body are "intrinsically different structures" rather than "points on some continuum" (Browman and Goldstein 2002: 9). Phonologically, the actions of distinct organs are basic. For example, most if not all languages effect contrasts based on vocal tract organs, e.g. Labial (lips), Coronal (tongue tip/blade) and Dorsal (tongue body) stops, and Nasality (velum). Browman and Goldstein (1989) also cite developmental evidence based on the experimental work of Meltzoff and Moore (1977; see also Kuhl and Meltzoff 1982, 1984) on infant facial mimicry of adult facial gestures. When experimenters performed facial gestures involving specific facial organs (e.g. lips, tongue, eyes), infants produced movements that approximated the adult movement in organ-specific ways. For instance, the infant may imitate a tongue protrusion task, in which the experimenter protrudes the tongue sideways, with a straight out movement of the tongue. The early ability of infants to individuate and match organs on the basis of optic and kinaesthetic information is of course underpinned by mirror neurons. Studdert-Kennedy and Goldstein (2003) examined a corpus of phonetic transcriptions of infant speech - their own and that of Vihman (1996) – to show that children show a lot of consistency in the vocal tract organ that they use to articulate the initial consonant of a word. To quote Browman and Goldstein (2002: 10):

> For example, a given word might always begin with a lip gesture, but vary as to constriction degree (stop, fricative, or approximant) or in voicing. Thus, children appear

> to acquiring (sic) a relation between actions of distinct organs and lexical units very early in the process of developing language. This relation is apparently not mediated by a phonological unit specified for additional features. At this stage, it is the organ identity itself that infants incorporate in early lexical items.
>
> (Browman and Goldstein 2002: 10)

On this view, a gesture is a constriction action of one vocal tract organ, as we saw above in relation to the lip closure gesture. In that case, three articulators – upper lip, lower lip, jaw -- form a functional synergy to effect lip closure. Of course, a number of vocal tract organs, not just the lips, are involved in the formation of constrictions during speech. Vocal tract organs include lips, tongue tip, tongue body, velic aperture, and glottal aperture. Vocal tract variables pertain to each organ. Table 2 summarises and adapts discussion in Browman and Goldstein (2002: 12).

Vocal Tract Organ	**Vocal Tract Variable**	**Articulators Involved**
lips	lip protrusion	upper and lower lips, jaw
	lip aperture	upper and lower lips, jaw
tongue tip	tongue tip constriction location (location of constriction along vocal tract)	tongue tip, tonguebody, jaw
	tongue tip constriction degree (size of variable serving as rest position for constriction action)	tongue tip, tongue body, jaw
tongue body	tongue body constriction location (location of constriction along vocal tract)	tongue body, jaw
	tongue body constriction degree (size of variable serving as rest position for constriction action)	tongue body, jaw
velum	velic aperture	velum
glottis	glottal aperture	glottis

Table 2: Vocal tract variables and articulators for the five vocal tract organs (adapted from Browman and Goldstein 2002: 12)

Lip closure and devoicing are two discrete gestural units that have the potential to enter into productive relations with each other, e.g. in /p/, or with other combinatoric phonological

units. Thus, /p/ is an organized assemblage formed from the constriction relations between the two discrete gesture units. Talk is decomposable into a small language-specific repertoire of such action units that have the additional properties of discreteness and context-invariance. These gesture units are not the same as the segmental units and features of conventional phonological representation, though they are of a similar degree of graininess (Browman and Goldstein 1995). They are in any case the basic units of phonological structure. This is so in two senses. First, they exhibit properties of the combinatoric units characteristic of traditional phonology. Secondly, they are measurable bio-physical processes qua units of action. In the first sense, they are seen as time-invariant, discrete, and low-dimensional – all key properties of combinatoric systems. In the second sense, they are seen as time-varying, continuous, and high-dimensional – all key properties of real-time bio-physical processes.

In the traditional view, the two sets of units were viewed as incommensurate. Properties pertaining to the first sense were not found in the physical speech signal. The solution was to postulate that formal phonological units belonged to an abstract mental code that existed in the brain/mind of the individual. On this view, the physical properties of real-time co-articulation have no systematic relation to the abstract formal units of this internal mental code. It was the task of the speaker/listener to infer the abstract units of the code from the messy details of real-time co-articulated speech events (see Fowler 1980, 2010). Moreover, Browman and Goldstein (1995) showed that the two senses – the formal, combinatoric and the bio-physical – can mutually constrain each other in various ways. The solution, according to Browman and Goldstein (1992, 1995), lies in shifting away from abstract formal units as the components of combinatoric systems to distinct action units, i.e. phonetic gestures. These action units have the properties of discreteness, context-invariance, and are small in number. They are the most basic units of phonological structure. They therefore fulfil the requirements of both formal combinatoric and bio-physical processes, simultaneously.

The creation and release of constrictions in vocal tract gestural activity articulates a unified and dynamical wave-like pattern of energy and organization in the form of the combinations of phonetic gestures that articulate words and larger units. The constriction gestures that form and then decay and which we associate with specific phonological segments, traditionally defined, turn out to be concurrent, rather than sequential. Browman and Goldstein (1992, 1995) have called this overlap "co-articulation." As we saw above, the consonant /p/ involves the co-articulation of two overlapping gestures, i.e. lip closure and devoicing. These two gestures are concurrent, rather than sequential. In utterance-activity, temporally overlapping phonetic gestures are performed by distinct vocal tract organs in the production of phonological units such as consonants and vowels, as shown by Browman and Goldstein. In the case of /p/, this means that the lips can yield a closure action at the same time that Voicing occurs. Consonants and vowels are, then, complex, multidimensional events that express physical information about the behaviours of more than one organ simultaneously.

In Gibson's terms (1986 [1979]), this is information about a source-event (an organ-specific gesture or constellation of gestures) in the environment rather than information about the acoustic signal (Fowler 1980, 2010). The organ-specific character of phonetic gestures means

that the gestures produced by these organs do not blend during speech production (Browman and Goldstein 2002: 18). The discreteness of the source organs involved in the articulation of consonants and vowels multiplies the differentiations that are effected by a particular phonological unit. The increasing capacity for complex, multidimensional differentiation by means of phonetic gestures in a given population of interacting agents augments the available pathways for energy dissipation through linguistic catalysis. The fine-grained sensorimotor differentiations made by phonetic gestures mean that languaging agents are able to constantly fine-tune and adjust their catalysis to respond to a vast range of different environmental events (Thibault 2004a: 184-187). This accounts for the vast range of environmental differentiations that a person is able to articulate and experience in and though linguistic catalysis. On this view, phonetic gestures are physical (bodily) activities that are linguistically significant qua physical (not mental) events. They are the actions that cause the acoustic speech signal to be propagated in its environment from a source-event (a speaker's vocal tract activity) relative to some dialogical array and its observer positions. In keeping with a central tenet of Gibson's ecological theory of perception, Fowler (2010) points out that "the relation between acoustic structure and corresponding perceptual object is considerably more direct than it can be if language forms are mental categories." (2010: 291). Fowler further argues that language forms are adapted to two critically important parity constraints.

First, they are physical events that are perceived, rather than mental ones that are inferred. Secondly, language forms are preserved throughout a communicative event such that speakers and listeners do not lose them and therefore do not have to recover them in the course of that event (Fowler 2010: 292). They are held as contents in working memory on timescales that persist throughout a communicative event. Different phonetic gestures physically and causally structure the air differently such that the patterning in the air over time provides information about the gestures themselves and their source qua environmental event. On this view, Fowler argues that language forms are public events that are "adapted to the achievement of parity." (Fowler 2010: 292). Seen in this light, phonetic gestures appear to be good candidates for travelling waves or solitons. Soliton-like waves are localized and robust in the sense that they do not easily dissipate (Carpenter and Davia 2006). Solitons need not be solitary waves, as in the case observed by J. Russell Scott in the mid-nineteenth century of the solitary wave of water that maintained its structure for two miles when a boat suddenly stopped in a canal. Solitons can be complex, multidimensional patterns.

Phonetic gestures are time-locked environmental events that persist and are preserved during a languaging event as coherent, ordered events exhibiting higher-order perceptual invariants within the environmental space and energy of an excitable medium, e.g., the air in the external environment, the auditory cortex in auditory perception, and the brain. The formation and persistence of phonetic gestures qua soliton-like waves is related both to the invariances and symmetries of its environment and to context-varying information. Simple invariant acoustic and other properties have proved elusive owing to the fact that the acoustics are not produced by a given, localized segment, but by the time-locked trajectory of the vocal tract gesture as a whole during its formation. Phonetic gestures exhibit another general characteristic of soliton-like waves insofar as they mediate transitions by channelling energy

into structure. They do so in time. The simultaneous production by the vocal tract of vowel and consonant restrictions and the transitions between these in a complex, time-locked temporal trajectory of vocal tract activity is just such a channelling of energy into structure. If vocal tract gestures truly were sequential rather than temporally overlapping or co-articulated, they would not be capable of expressing the large amount of high-dimensional information – the rich physical detail – that is in fact characteristic of vocal tract gestural activity.

Catalysis in general involves the overcoming of structural constraints to dissipate energy, as in the paradigmatic case of enzyme catalysis. Body and brain are excitable media with travelling waves of excitation and inhibition. Carpenter and Davia (2006) extend the domain of catalysis to apply to the macroscopic level of biological systems embedded in their environments. According to Carpenter and Davia, all biological processes mediate transitions by channelling energy via structure. As pointed out before, these researchers argue that this occurs at all scales from enzymes, cells, organs, and organisms. At all of these scales, a living entity just is a unitary catalytic process of mediating its environment. Phonetic gestures are one instance of the complex, multidimensional behavioural waves whereby an organism unifies the boundary conditions resulting from the organism's interactions with its environment. Brain and body are the media of linguistic catalysis.

Co-articulated gestures, as shown in the work of Browman and Goldstein (1989, 1992, 1995), overlap in real-time, are continuous, and show considerable context-sensitive variation from speaker to speaker and occasion to occasion just as prosodies do. Decades of research and experimentation in phonetics (e.g. Shankweiler et al 1977; Browman and Goldstein 1992, 1995; Port 2007, 2008) have shown that topological-continuous variation in the dynamics of articulation is, in actual fact, the norm as distinct from the long-standing assumption that the articulatory and auditory space of speech events is segmented into discrete, invariant formal segments, e.g. phonemes. This observation applies to parameters such as voice-onset time, vowel quality, consonant place of articulation, the duration of vowels and consonants, and the temporally overlapping, context-sensitive variation that is characteristic of co-articulation (Browman and Goldstein 1992, 1995; Fowler 2010; Port 2007, 2008). As pointed out in Section 13, observers of such events actively explore and probe these and other parameters in order to discover the relevant patterns of sensorimotor dependency (Noë 2004: 13). The parameters mentioned above as well as vocal tract organ-specific constrictions are some of the invariants that are detected in phonetic gestures. They are aspects of very fine-grained sensorimotor discriminations that effect transitions in co-articulated phonetic events. As noted in Section 13, these invariants acquire content through the application of the perceiver's "sensorimotor knowledge" Noë (2004: 9) as perceivers mediate these transitions during linguistic catalysis.

A languaging agent who is sensitized to the patterns of invariance in the given phonetic gesture will manifest a travelling wave across his or her auditory cortex as he or she attunes to the acoustic signal and its dynamical, time-locked properties. The wave is a complex pattern of AM activity (Freeman 2000): neurons firing at particular amplitudes constitute the wave structure. The wave is context-sensitive such that the environmental events with which the

wave covaries in experience give rise to statistical learning in a population of interacting agents. This means that populations of travelling wave patterns consisting of phonetic gestures or gesture-types are associated with particular classes of environmental events in apperception (Verbrugge 1985, Thibault 2011). Phonetic gestures therefore exhibit another feature of travelling waves: they propagate a fixed structure and its associated energy within a dialogical array on account of the unity of structure and energy that is characteristic of travelling waves. They are, as Fowler shows, typically preserved rather than lost during a languaging event. They therefore do not dissipate easily. Also, phonetic gestures are stored as rich and detailed (high-dimensional) auditory and articulatory memory for linguistic patterns (Port 2008: 4). The capacity of phonetic gestures to be preserved over a given languaging event lends support to the idea that it is the rich, physical detail of such events which agents remember and recall on future occasions on the basis of their direct experience of such events, rather than sparse phonological schema (Port 2007, 2008).

Languaging behaviour consists of cycles of dialogically coordinated catalyzing activity that loop through agents and their environments. Human agents make explicit organized linguistic structure in the implicit structure of phonetic gestures by their catalyzing activity (Section 13). This is an active and skilful process whereby agents adapt their bodies to the skilful sensorimotor exploration of the phonetic event, in the process making body and brain in some respects conform to and become similar to the event (Freeman 2000: 120). As Freeman also points out, the forms that one discovers in such events when one shapes body and brain to the event in this way do not pass from event to brain. Instead, the application of sensorimotor knowledge and the discovery of forms – high-order physical invariants – in such events catalyzes in the imagination flows of cognition, affect, action, and understanding such that we are able "to create the internal structures with which we can act and understand." (Freeman 2000: 120).

In the following Sections 16 to 19, I shall consider in more details how this happens. These Sections (and others) will appear in Part 2 of this article. Part 2 will be published separately in the following issue of PJOS III.3 (2011).

16. Intentions as Control Parameters and Grammatical Order Parameters

Intentions flow through and modulate the dynamics of articulatory trajectories. Thus, the intention to articulate the lexeme /pat/ rather than /bat/ is, in part, an intention to articulate a lexeme with an initial voiceless consonant /p/ in contrast to the initial voiced consonant /b/. Gafos (2006: 67) points out: "The intention to communicate a lexeme with a final voiced consonant, in particular, is defined as a part of a dynamics that attracts the order parameter toward the intended voicing. In turn, intentions are constrained by the grammar dynamics, namely, by how forms 'should be produced' in specific contexts." Gafos argues that "intentions are communicative goals." (2006: 67). I would see this as only partially true. Intentional dynamics, following Juarrero (1999), flow through the entire action trajectory from its onset to its conclusion and modulate it. They are not simply target goals though I am in agreement with Gafos, who also says that, "Intentional dynamics adds an attractor at the required value of voicing {-x0, x0}, where x0 '=' [-Voiced], -x0 '=' [+Voiced]." Intentions

are non-grammatical control parameters that can vary in intensity. This control parameter interfaces with and interacts with a grammatical order parameter (Gafos 2006: 68). Gafos observes: "Order parameters describe the macroscopic form of phonology and grammar principles refer to such parameters … " (2006: 68).

Lexicogrammatical patterns and formats are socially and culturally distributed re-descriptions of fragments and combinations of fragments of many utterances experienced by speakers as rich, high-dimensional phonetic gestures in the contexts in which they covary with aspects of human experience. Port (2010: 316) shows that fragments of these concrete memories "are assigned to categories in many ways by speakers of the language related to what we call semantics (e.g., singular/plural, etc.), phonology (e.g., +/-voice, /b, d, g/), and syntax (e.g., grammatical categories)." (Port 2010: 316). Grammar order parameters arise in a population of speakers owing to the increasing conventionalization and standardization of aspects of rich, high-dimensional phonetic gestures as phonological, lexicogrammatical, and semantic categories. Such categories represent formalizations of different aspects of first-order languaging dynamics as formal categories in a second-order rationalization or re-description of language, seen as consisting of different levels of abstract form in a language code. Aspects of dynamics are, accordingly, frozen or hardened as normative constraints on the dynamics through selection pressures arising from informal and formal teaching and learning, the enforcement of cultural norms, the modeling of speech in terms of writing, and so on.

Thus, phonetic gestures get re-categorized as socially distributed and increasingly norm replicating phonological forms that stabilize pronunciation routines. In turn, these phonological routines get associated with semantic categories in ways that are stabilized at the population level as regular, habitual patterns of association between phonological forms and semantic categories. The cultural standardization of these patterns of association thus gives rise to 'grammar' as the re-categorization of these routines and their habitual semantic associations such that the two are mutually constraining and linked to particular situation-types and their conventions. By the same token, the emergence of 'grammar' goes hand-in-hand with the augmented cognitive and meta-linguistic capacity to reflect on and to selectively analyse (aspects of) lower-scalar dynamical processes in terms of culturally sedimented formal categories that provide the basis for folk-theoretical ways of scaffolding and coordinating interactions between persons around appropriate norms. The formal categories thus provide a distorting lens through which lower-scalar dynamics are perceived, analysed and evaluated. 'Grammar' therefore gives rise to an increasingly self-maintaining and quasi-autonomous cultural dynamic of meta-organizations composed, in turn, of real-time lower-scalar dynamical processes on which 'grammar' induces us to reflect in culturally regimented and constrained ways. Lexicogrammatical patterns are population-level (e.g., cultural) constraints on the bio-physical dynamics of phonetic gestures. The gestures compress in their dynamical properties accumulated historical-cultural information and higher-order invariant properties (Section 10) that have the functional capacity to evoke culturally specific forms of virtual experience and virtual semantic entities. These 'digital' semantic differentiators bias action, feeling, cognition and perception in value-weighted ways that give rise to higher-order behavioural control (Sections 4, 11, 18, 25).

Seen in this light, 'grammar' specifies culturally constrained and standardized order parameters that also interface with and are modulated by the intentions of languaging agents. Intentions flow through real-time neural and bodily dynamics of first-order languaging behaviour and are not reducible to the order parameter set by culturally regimented grammatical form. By the same token, the intentional dynamics of agents are attracted to and flow into the attractor spaces that are set by lexicogrammatical order parameters. Intentions are control parameters because they flow from an intentional source – the neural dynamics of an agent – through body dynamics and out into the environment along the temporally unfolding action trajectory, e.g. a vocal tract gesture. In this sense, they are an aspect of the mechanisms whereby agents control and modulate their own and others' behaviours relative to that environment. At the same time, 'grammar' is an order parameter in the sense that it is a mechanism-independent quasi-causal constraint or attractor qua future cause that shapes, guides and sculpts the development of the action trajectory as it is pulled into its future. This occurs in ways that are constrained by cultural norms that serve to orient, scaffold, and coordinate the intentions of agents when they interact with each other. Intentions, then, are not so much goal-directed, but anticipatory.

Gafos points out that "intentional strength is a scalar variable" (2006: 70) that varies continuously along an interval or cline. The intention to express this or that contrast, e.g. the contrast between [Voiceless] and [Voiced] in the /pat/ vs. [bat] pair is motivated by contextual factors. However, intentions are not causes of actions. They are aspects of the processes themselves (the actions) that may change a system from one state to another. They are immanent in the material dynamics of action trajectories such as phonetic gestures. Intentions refer to the actual world at the same time that they interface with the virtual world of 'grammar' by virtue of the multiple attractors that compete for the dynamics of the unfolding action-trajectory, defined as a distribution of singularities. A particular phonological value, e.g. Voiced or Voiceless, is a singular point or a singularity -- a stable point in phase space towards which vocal tract trajectories tend to converge. A trajectory is defined as a stable behaviour that changes as the behaviour changes state, i.e. it passes through a series of points in phase space as it converges towards its attractor.

Attractors are singular points in this sense. A state space can have several attractors. Its singular points are thus surrounded by a basin of attraction, a region that affects the behaviour of trajectories. If behaviour starts in a particular basin of attraction it will end up at the attractor. Attractors thus define regions of stability; they stabilize an attractor around a set a values that restrict its degrees of freedom. Intentions, on the other hand, are control parameters that function to take a stable system (an order parameter) to one of its critical thresholds. They are scalar variables that vary in intensity such that a critical threshold is reached and a bifurcation occurs, i.e. one stable distribution of attractors (an order parameter) is changed into a different, contrasting one, as shown above in the discussion of the lexemes /pat/ and /bat/. Intensive variables such as intentionality are control parameters that can transform one distribution of attractors into another one once a critical threshold of intensity is crossed. Whereas order parameters define regions of stability, control parameters such as intentionality define the potentiality for phase transitions to occur that dynamically transform

one stable distribution of attractors into another, different one.

Control parameters and order parameters together define an immanent space of possibilities. Intensive properties such as intentions qua control parameters are not therefore extrinsic to this space of possibilities, but are immanent in them. A phase space so defined corresponds, in the Deleuzian ontology, to a *multiplicity* (Deleuze 2004: 230-235). Multiplicities define the spontaneous capacity of a material process to generate pattern without external intervention (DeLanda 2002: 26). A multiplicity is defined by the distribution of singularities in a material system that define its tendencies, and the critical transitions that take the system from one phase state to another. Intentions can thus be seen as a topological-continuous space that progressively differentiates and specifies itself into the partitionings of vocal tract variables into stable, discontinuous regions that afford the reciprocal coupling and mutual attunement of speakers to each other's vocal tract dynamics.

Gafos sums up his discussion as follows:

> There is a parameterization in terms of an order parameter and a control parameter, in (11a, b) respectively. Order parameters describe the macroscopic form of phonology and grammar principles refer to such parameters (see Gafos 2002 on gestural coordination relations). In our example, the control parameter is intentional strength. As shown in (11c), there is also an 'interface', the hypothesized model relating these two parameters, dx/dt = G(x) + intent * (xREQ– x), where G(x) = – k + x – x^3. Crucially, however, this 'interface' does not translate symbols to continuous signals. Rather, it states a dynamic linkage, in the form of a testable relation, between a grammatical (order) parameter and an extra-grammatical (control) parameter. The linkage is dynamic because the two parameters it relates are interdependent and changing quantities, …
>
> (Gafos 2006: 69)

Intentional strength is a scalar phenomenon. Once the intentional strength interfaces with the grammar, the latter moves away from equilibrium. If the intentional strength is made intense enough, the grammar system can undergo a phase transition to a new stable state. Intentional strength is defined by continuous intensive properties that can differentiate into discontinuous regions corresponding to particulate phono-grammatical categories. These processes of differentiation occur because of the intrinsic partitioning of possible vocal tract gestures into distinct categorial regions that afford attunement between the members of a population of speakers and therefore serve to differentiate lexicogrammatical units. Constriction parameters of the vocal tract are assigned quantitative values (Stevens 1989).

17. Constriction Parameters, Mutual Attunement and the Dialogical Coupling of Languaging Agents as Bio-dynamical Social Engines

Stevens showed that the relation between constriction parameters and their acoustic properties exhibit nonlinear properties. A restricted range of these values will afford attunement between speakers, whilst others will not. The topological continua of vocal tract variables can

therefore be partitioned into discrete regions that afford attunement. Stevens shows that the relation between constriction parameter values and acoustic properties divides into stable and unstable regions. In the stable regions, small changes in the constriction parameter result in small changes to the acoustics. In the unstable regions, small changes to constriction parameters result in large changes to acoustic properties. According to Goldstein and Fowler (2003: 35-39), the former afford mutual attunement between speakers because their properties afford imitation by the members of a population of speakers whereas the latter do not. The interaction between intentions and vocal tract variables can be many and varied. Intentions are immanent in the dynamics of these material processes.

Strictly speaking, imitation in the early months of the infant's life refers to the infant's capacity to relate to and to mirror the external behaviour of others (Hart 2011 [2006]: 23-24). However, mutual attunement is more than imitation so defined. Imitation alone could not get the dialogical coupling of persons through their reciprocal languaging off the ground, so to speak. Mutual attunement is based on the capacity to affect each other's **internal** dynamics and to modulate them. Mutual attunement is affective attunement and, as Hart points out, it "is essential for our ability to feel other people and to feel that we are felt, which facilitates the development of attachment capacity and enables us to relate to significant others throughout our life span." (Hart 2011 [2006]: 23). The capacity to affect each other's internal dynamics is therefore an essential foundation for the capacity of listeners to covertly simulate through the activation of neural structures that mirror the vocal tract gestures of the speaker, as proposed in the motor theory of speech perception (Liberman and Mattingly 1985; see also Kinsbourne 2005).

Sander (1977) showed that the infant's ability to self-regulate develops in the earliest stages of infancy. Caregiver and infant engage in interaction routines that promote affect-based forms of co-regulation (Trevarthen 1998; Bråten and Trevarthen 2007; Bråten 2007). The recticular activation system in the brain stem creates a temporal structure for brain activity. Neurotransmitters in this system control sensorimotor exploration, attention, and motivation. Motivation arises from the interaction between the recticular activation system, the diencephalon, and limbic system (Hart 2011 [2006]: 48). Bråten (2007), Trevarthen (1998), Stern (1977), Cowley et al (2004) and others have further shown that synchronized interaction between infant and caregiver gives rise to joint motivation and attention. The mutual regulation of arousal between infant and caregiver also results in physiobiological and psychobiological attunement between them (Stern 1984). As Hart (2011 [2006]: 49) points out, caregiver and infant modulate each other's energy states through their mutual attunement and the interpersonal routines in which this is embedded.

New patterns of neural activity form in the brain as a result of the interactions between persons and their environments. Regular, repeated experiences and intense affect establish new patterns of neural activity that reactivate more readily (Hart 2011 [2006]: 49). When mutual attunement occurs, there is 'deep coordination' (Hart 2011 [2006]: 49) of the autonomic and limbic centres in the brain, which are important for the fine-tuning of the emotional regions of the brain amongst other things such as "heart rate, breathing, digestion"

(Hart 2011 [2006]: 49). Synchronized attunement, which is based on the limbic regions, is referred to as "limbic resonance" by Hart (2011: [2006]: 49). Limbic resonance, Hart (2011 [2006]: 49-50) explains, "requires that an internal state be expressed externally." (Hart 2011 [2006]: 49). The many fine-grained motor distinctions of the face, the hands, and the vocal tract are able to articulate many very fine-grained motor discriminations and combinations of these, which partly come under the control of the cranial nerves in the brain stem. The cranial nerves connect to circuitry involving the diencephalon, the limbic system, the insula, and the amygdala. They also "later connect with an area deep inside the frontal lobes: the orbitofrontal cortex, which enables us to feel and understand facial expressions." (Hart 2011 [2006}: 50). In this way, the human brain and that of other primates can link the perception of gaze and facial expression to emotion, motivation, and meaning.

Simondon (2010: 385) points out that motivation are not based solely on individuals, but on the potential energy that the individual person has in relation to the social group. Individuals, Simondon argues, are partially autonomous modulators of social realities. Individuals can activate ["déclencher"] flows of energy qua intentions and motivations through action trajectories that operate on and adjust the relational dynamics of selves in their social milieu. For this reason, individual persons may be described as *bio-dynamical social engines*. The material dynamics of vocal tract activity is what enables the flow of energy through the system, i.e. along its time-locked trajectory. Forms qua topological constraints on the dynamics modulate the flow of energy. Stable regions of vocal tract activity yield organ-specific gestures that have the potential to specify differential contrasts in phonological units such as /p/ in contrast to /b/ (Section 15). Contrasts can be either typological-categorial or topological-continuous. Perceived contrasts in vocal tract and other bodily behaviours enable languaging agents to mediate transitions in their environments as they catalyze their worlds (Section 13). Contrasts -- topological and typological -- mediate transitions in the flow, distribution and modulation of the energy that flows through the action trajectory and into the environment in which it is articulated. Matter and form are thus unified in linguistic catalysis as structured energy fields that make information available to agents willing and able to respond to it.

As shown in the work of Fowler (1986, 2010), speech sounds provide observers with information about the more fundamental environmental event – the vocal tract gestures of the speaker –, which the listener attunes to so that the coupling of speakers and listeners can take place. The capacity for covert simulation in one's own neural structures of the other's vocal tract gestures clearly is a case of the speaker affecting the listener's internal dynamics through the coupling of their internal (neural) and external (bodily) dynamics. The entrainment of a population of languaging agents to similar patterns of vocal tract dynamics thus provides a solution to the problem of assimilating one's own internal dynamics to the behavioural and neural dynamics of others. It is in this way that persons can assimilate their own behavioural and neural dynamics to the experiences, feeling, intentions, points of view, and the meanings of others in socially coordinated ways that have no need for abstract codes. Languaging behaviour is an external resource whose regular, habitual, norm replicating patterns in the form of vocal tract gestures in a population give rise to population-level entrainment effects.

Kinsbourne (2005) further points out that this means that entrainment is not only in terms of bodily interactional synchrony, but also involves the entrainment of the neural structures required for the covert simulation of the other's vocal tract gestures in the forebrain of the listener. This means that the neural structures involved are entrained to the intrinsic potential for complex, fine-grained differentiations that vocal tract gestural activity affords.

Stevens (1989) showed that the stable regions of the articulatory-acoustic map are the basis for phonological units. According to Browman and Goldstein, "the partitioning (sic) such continua into discrete regions emerges from the constraint of mutual attunement under conditions where the relation between constrictions and acoustics exhibits such non-linear maps." (2003: 14). The attunement constraint means that, in time, agents will converge on and eventually settle on the same constriction values. The attunement constraint is therefore responsible for the partitioning of the constriction degree continuum into discrete regions that are interactionally salient for agents due to the population-level entrainment effects mentioned above.

As we noted in Section 15, the phonological units of a language are formed from combinations of discrete organ-specific phonetic gestures. The combination of the organ-specific gestures of lip closure and devoicing is a value-realizing action that has the value /p/. The lip closure gesture in /p/ is also present in the initial consonant of words like *mat, pick,* and *bat* in combination with other organ-specific gestures that combine with lip closure to form consonants such as /m/ and /b/. Thus, the combination of the gesture primitives [lip closure] and [devoicing] produces /p/; in contrast, the combination of [lip closure] and [voicing] yields /b/. The phonological units /p/ and /b/ can occur in initial position of the lexemes *pat* and *bat*.

The vocal tract actions of different individuals in the same community must be attuned to each other – around six months infants begin to align their own vocal tract activities to the range of constriction parameters that afford attunement among speakers. The constriction parameters that afford attunement in a population of speakers draw attention to the fundamentally ecological, value-realizing character of languaging (Hodges 2007, 2009). A primary value that is realized by languaging is the building of dialogical connections with others. The mutual attunement afforded by the given range of constriction parameters is therefore a fundamental value that is realized when speakers achieve mutual attunement to each other's vocal tract actions. Persons physically modify their vocal tract gestures in order to attune to others. Moreover, the phenomenon of 'gestural drift' (Sancier and Fowler 1997) shows that speakers can modify their vocal tract actions throughout life in response to the changing demands of the different speech communities in which they participate. Attunement is a primary value because it underpins all other forms of value-realizing that can be enacted in languaging (Hodges 2007, 2009).

The constriction parameters that afford attunement constitute a most basic way in which languaging contributes to the recursive self-maintenance of persons through their languaging behaviour. Speakers contribute to their own self-maintenance by attuning to the constriction parameters of the given population and tuning out of those that are not pertinent in that

population. Attunement can be considered to be an appropriate interaction outcome because it serves the function of contributing to the maintenance of the far-from-equilibrium stability of the person as a member of that speech community. The detection of the salient constriction parameters in the vocal tract actions of others implicitly predicates that the relevant environment is functionally appropriate to a range of interaction possibilities and that these possibilities can hold or not hold in that environment (Allen and Bickhard 2011: 109).

A given constriction parameter is an indicator of a set of interaction possibilities. If one of these possibilities is articulated, i.e. by a particular constriction parameter, then the interaction will succeed, e.g. mutual attunement will be achieved, if the environment is appropriate to the indicated possibility. The observables of the environment are the constriction parameters that are detectable in the vocal tract actions of others. These parameters implicitly predicate that the current environment is one that is appropriate to the possibility of attunement with other speakers. This entails the learning of a linkage between the appropriate parameters and the implicit predicate to be learned that is internal to the person. That is, the agent generates an internal model of its agent-environment transactions that operates on a slower timescale with respect to the agent's *in vivo* interactions with its environment.

Here, the environment consists of the other speakers with whom one can be mutually attuned through the detection in other's vocal tract actions of the appropriate parameters just as others can detect the appropriate parameters in one's own vocal tract actions. What governs the behaviour of languaging agents, on this view, are the specific features of two environmental qualities: (1) the co-presence of other speakers; and (2) attunement to the same constriction parameters. Rather than notions like shared meanings, we see that our discussion is focused on receptors (auditory, visual, haptic, etc. systems) and actuators (vocal tract activity, facial expressions, hand-arm gestures). Receptors and actuators are tightly coupled with each other and with what von Uexkull (1928) called the organism's *Innenwelt* in the creation of the action circuit whereby the *Innenwelt* is determined through organism-environment interactions. The *Innenwelt* that is so created just is the world for that organism. Moreover, the interactions that give rise to the *Innenwelt* are value-realizing. Value-realizing behaviours ensure that the encountered constriction parameters are not just a motley collection of discrete entities, but they are integrated to a meaningful whole that realizes a value for the organism.

Constriction parameters are elementary differentiators that have the capacity to indicate possible interactions with the relevant environment. In this sense, they constitute anticipations of possible interaction outcomes. The constriction parameters that are salient in a given population serve this function because they differentiate the environment in relation to the agent's value-realizing behaviours. In the first instance, this means that they have the functional capacity to achieve mutual attunement between speakers. Importantly, this shows more clearly why the constriction parameters have value for agents in their worlds – not merely because they participate in an abstract play of differential oppositions between phonemes, but because they are positive and productive differentiators (Deleuze 1994 [1968]: 254-256) whereby agents affect others and are affected by them. Their value resides in the fact that languaging agents are highly attuned to the behaviour of others. Moreover, specific

constriction parameters can directly affect interaction outcomes by amplifying or inhibiting the responses of others. This has important implications for a catalytic theory of language.

Hodges (2009: 638) reports on recent work in ecological psychology using non-linear dynamical analyses that can help to better understand these issues in relation to fractal patterns:

> Recent work in ecological psychology (e.g., Van Orden, Moreno, and Holden 2003; Van Orden, Holden, and Turvey 2005; Van Orden, Kello, and Holden, in press) using non-linear dynamical analyses, have revealed pervasive, long-range patterns in linguistic performances (e.g., word pronunciation, lexical decisions, semantic categorization). These data are startling and have important implications for our understanding of intentional activities. If, for example, someone repeatedly says a word thousands of times, and measures such as the latency and duration are taken, aperiodic waves of variation in amplitude occur across time. Instead of random variation around a "true score" of latency or duration, the waves of variation at different scales of times (e.g., blocks of 10 trials, 100 trials, 1000 trials, 10,000 trials), are similar at every scale. The similarity is so precise that if the amplitudes are graphed relative to frequencies on a log/log scale, the relationship is linear. Such a pattern is *fractal* (sometimes referred to as 1/f scaling).
>
> (Hodges 2009: 638)

Hodges further points out that no one scale or unit of measurement can capture or describe the fractal phenomenon. These observations readily apply to the constriction parameters discussed above. Hodges identifies three ways of looking at such patterns: the patterns are global, collective, and paradoxical (Hodges 2009: 638). Fractal patterns are global because the larger the sample and the more precise the measurement device, the more variability is revealed. Fractal patterns are collective because they cannot be reduced to any single causal principle or module. Instead, they are the emergent results of the interactions between many different components and many different processes and are irreducible to any single principle or component process. The fractal pattern is immanent in the interactions between the component processes over the space-time of the interaction between the components. It ceases to exist once the components cease to interact. Fractal patterns are paradoxical because a single movement, e.g. an organ-specific constriction parameter in vocal tract action, correlates with an increasing diversity of patterns across scales. Hodges observes: "Actions that occur on the scale of 0.2 sec show dependencies on other activities that are on the order of 11,000 sec (Van Orden, Holden, and Turvey 2005). Interdependence implies that each part reflects something of the whole in its behavior." (Hodges 2009: 639). For example, the correlations between organ-specific constriction parameters, combinations of organ-specific gestures, and morphemes, etc. suggests the interdependence, context-sensitivity and global inter-dependence of the different scales discussed by Hodges (2009: 639).

18. Mutual Attunement, Phonological Structure and Value

Mutual attunement of vocal tract gestures arises and is maintained in a population of languaging agents not only because of the reciprocal coupling of the organ-specific gestures of agents via the medium of air; it also arises in response to the need to solve problems of coordination between agents in emergent social situations. To participate in social situations means that one is, implicitly, a social person (Bickhard 2006). To promote oneself as a social person is to make usually implicit and inherently normative claims as to one's capacity to participate in such situations. Moreover, it means that one's own actions, including utterance-activity, occur in and are promoted within a heteroglossic field of others' desires, motives, intentions, and plans. This discovery goes hand in hand with the discovery that the desires, motives, etc. of others may contrast with one's own. One learns that the social world is populated with other persons' points of view (Reed (1996a: 167). Reed points out: "There is thus strong pressure on children to learn to select utterances that produce desirable effects." (1996: 167).

Learning about others' points of view is, Reed argues, "a fundamental cognitive milestone" (1996a: 167) in the life of the child. This developing awareness of the points of view of others entails an inherent reflexivity in the coordination problems that arise between agents in social situations and in their solutions. The social coordination of diverse points of view in the dialogical array of observer positions and perspectives in some situation calls for solutions that can provide some sort of resolution of, say, your understanding of my understanding of your understanding … in the given situation. The heteroglossic diversity of points of view that underpins this inherent reflexivity are then "key elements in structuring the child's linguistic environment in such a way as to provide information concerning the significance of varying speech structures and that set the child on the road to making discoveries about the generative patterns inherent in the language around him or her." (Reed 1996a: 167).

The mutual attunement of vocal tract gestures is, as Goldstein and Fowler (2003: 35-39) show, a process of attending to and tracking the information about the other's vocal tract gestures that is made available through the restructuring of the medium of air that occurs in the act of speaking. In this sense, phonetic gestures are public, observable events, not private, mental ones; they take place between persons in social situations. They are events which persons actively and intentionally select and organize for the purpose of presenting information to others. In so doing, they make others aware of relevant environmental information even when the other person is not physically present in the situation and has no direct, online access to the situation.

Attunement begins in the earliest stages of infant-caregiver interaction so that infants align their vocal tract behaviour to that of the others in their environment (Halliday 1975; Vihman 1996; Cowley et al 2004; Kuhl 2007). Reciprocal attunement of vocal tract gestures takes place through the acoustic medium which functions to couple the vocal tract behaviours of agents to one another though visual and haptic information can also play a role in these coupling dynamics (Meltzoff and Moore 1977; Kuhl and Meltzoff 1982, 1984). Browman and Goldstein (2002) make the following pertinent observations:

> Because the relation between constriction parameters and their resulting acoustic properties exhibits some nonlinear properties (e.g., Stevens, 1989), certain ranges of values of the constriction parameters will afford attunement among talkers (i.e., those values can be effectively imitated or mimicked), while other ranges of values will not. In this way, the continua corresponding to tract variable dimensions can be intrinsically partitioned into discrete modes or ranges that afford attunement. We will now describe a simulation of attunement that illustrates this idea.
>
> The idealized relation between a constriction dimension and (some dimension of) the sound that it produces can be seen in Figure 1, adapted from Stevens (1989). The relation between the constriction parameter values and the acoustic result shows different kinds of regions—stable and unstable. In stable regions (labelled I and III), small changes in the constriction parameter result in only very small changes to the acoustics, while in the unstable region (labelled II), small changes to the constriction parameters result in large acoustic changes.
>
> (Browman and Goldstein 2002: 13)

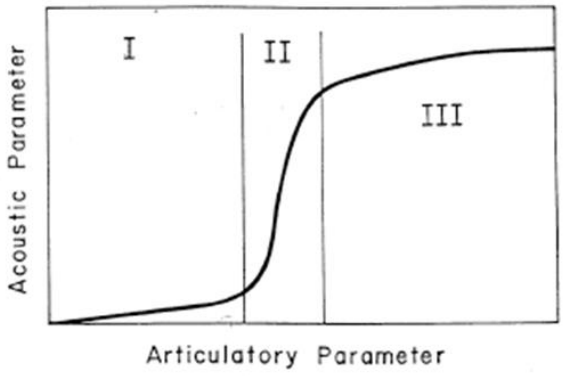


Figure 1: Non-linearity in the articulatory-acoustic map (after Stevens, 1989)

Browman and Goldstein discuss experimental results that demonstrate that the partitioning of constriction degree into discrete regions (e.g. closure, critical, narrow, mid, wide), as shown in their Figure 1 above, is "an intrinsic consequence of the attunement constraint, and the non-linear map relating articulation to acoustics." (2002: 15). Discrete categories of constriction parameters arise in the processes of reciprocal attunement that occur in a population of speakers. Browman and Goldstein model gestural structure in vocal tract activity in terms of two parameters: (1) the common distinct constriction degrees (closure critical narrow mid wide); and (2) common distinct constriction locations (labial dental alveolar palatal velar

uvular pharyngeal). The descriptors used here are no more than labels to approximate the "ranges of values of these parameters that are typically used contrastively." (2002: 16). Contrastive values are set up in a particular language when its speakers converge on these values through their reciprocal attunement to a stable constriction parameter (see above). These parameters are value-weighted such that their intrinsic tendency towards partitioning into discrete modes affords reciprocal attunement whereas unstable regions that do not lend themselves to mimicking do not. Contrastive values of gestural units in a given language arise through self-organizing processes of reciprocal attunement of speakers to one another's vocal tract gestures constrained by the shared biomechanics of vocal articulation and the coupling constraints of the acoustic medium (and to a lesser extent other media such as the optic and the haptic).

Mutual attunement is possible in the first instance because of the adaptation of phonetic gestures to the achievement of interpersonal parity (Fowler 2010: 292; Section 15). The achievement of interpersonal parity through the developmental processes of attunement referred to above means that organ-specific phonetic gestures are bonded or combined to form, in utterance-activity, gestural scores that have the capacity to selectively organize and present information in ways that foster awareness between persons (Reed 1996a: 171). Discrete organ-specific vocal tract gestures combine to form phonological units in which the discreteness of each of the gestures that combine to form a given unit is retained (Goldstein and Fowler (2003: 2). The gestures that combine in this way are potentially contrastive in the given phonological unit. For example, *bin*, *vin*, and *win* all begin with gestures of the lips, but the gestures contrast in the degree (state value) of Lip Aperture (or degree of lip constriction). *Bin* has the highest degree of lip constriction; *vin* has less lip constriction; and *win* the least (Goldstein and Fowler 2003: 35). The variable degrees of lip constriction exhibited along the continuum of Lip Aperture in the three examples are points along a scale of topological-continuous variation (of degree of lip constriction). The question arises as to how these continua are partitioned into contrasting states or values (of the vocal tract).

According to Goldstein and Fowler (2003: 35), it is the public, observable status of vocal tract gestures that provides the basis of the explanation. Vocal tract gestures, rather than being the imperfect realization of the abstract categories of an internal phonological code, are observable public events that are shaped by and attracted to the phonological norms of a particular speech community. Phonological forms are socially distributed second-order constructs with respect to the first-order phonetic gestures which speakers produce in their vocal tract activity. Phonological units and phonological structure in languages – the phonological forms of a specific language – are abstract structures with respect to the articulatory and acoustic properties of actual phonetic gestures as they are articulated and perceived in context-dependent vocal tract activity. Phonological forms are relatively more abstract and context-independent structures that are decomposable into combinations of discrete, organ-specific vocal tract gestures that are potentially contrastive, as noted above. They are low-dimensional attractor spaces with respect to the high-dimensional articulatory and perceptual detail of actual phonetic gestures in acts of speaking and listening. As low-dimensional attractor spaces, phonological forms are quasi-causal, mechanism-independent

constraints on real-time vocal tract activity. They are emergent normative constraints on the vocal tract activities of a population of interacting agents. As pointed out above, Goldstein and Fowler (2003: 35-39) show that mutual attunement is the causal mechanism that enables the emergence of phonological forms at the population level:

> Mutual attunement must be accomplished primarily through the acoustic medium. Because the relation between constriction parameters and their acoustic properties is nonlinear (Stevens 1989), certain regions of a vocal tract continuum will afford attunement, while others will not. Thus, the categories we observe could represent just those values (or regions) of the tract variable parameters that afford attunement. They are an example of self-organization through the public interaction of multiple speakers.
>
> (Goldstein and Fowler 2003: 36)

Building on the findings of Stevens (1989) referred to in the above quotation, Goldstein and Fowler (2003: 36) conducted preliminary simulations of agents attempting to attune their actions:

> In a preliminary simulation designed to investigate the partitioning of a tract variable constriction continuum into discrete regions, agents interacted randomly under the following three conditions: (a) Agents attempt to attune their actions to one another. (b) Agents recover the constriction parameters used by their partners from the acoustic signal, and that recovery is assumed to be noisy. (c) The relation between constriction and acoustics is nonlinear. The simulation investigated an idealized constriction degree (CD) continuum and how it is partitioned into three categories (corresponding to stops, fricatives, and glides). Figure 4 (in referenced document) shows the function used to map constriction degree to a hypothetical acoustical property, which could represent something like the overall amplitude of acoustic energy that emerges from the vocal tract during the constriction. The crucial point is that the form of the nonlinear function follows that hypothesized by Stevens (1989) for constriction degree and several other articulatory-acoustic mappings. Regions of relative stability (associated with stops, fricatives, and glides) are separated by regions of rapid change. The constriction degree continuum was divided into 80 equal intervals.
>
> (Goldstein and Fowler 2003: 36)

In recovering the constriction parameters from a noisy environment in which the relation between constriction and acoustics is nonlinear, agents in effect overcome structural constraints that enable them to channel energy via structure such that attunement between agents in a given population arises and, on that basis, dialogically coordinated means of mediating the environment in languaging behaviour. The developmental findings cited by Goldstein and Fowler (2003: 38) lend support to this:

> While children's early words are consistent in the oral constriction organ employed, and match the adult models in this regard, they are quite variable in within-organ properties,

> such as constriction degree (or constriction location). The simulations suggest that within-organ categories emerge only from attunement, which presumably takes some time. This conclusion is further bolstered by the recent perceptual findings with infants 10-12 months of age (Best and McRoberts in press), showing reduced discrimination for within-organ contrasts, even when the contrasts can be found in the language the child is about to acquire. At this age, infants have only begun to attune their vocal behavior to the language environment (de Boyssson-Bardies et al. 1992), and therefore partitioning of within organ categories is expected to be incomplete.
>
> (Goldstein and Fowler 2003: 38-39)

Children learn in time to control bodily dynamics, including vocalizing, in ways that control perceptual input. They do so through reciprocal attunement with others and through the honing of "within-organ properties." Initially, the infant has little control over this input, sending random commands to the muscles. In time, he or she learns to correlate certain random commands and within-organ properties with specific responses in the world, especially responses from other persons. This correlation is established on the basis of the consistencies that are established between the motor command and the perceptual input. In time, the infant can elicit desired responses in others by calling up the appropriate motor command (e.g. proto-imperatives). He or she can control vocal and other bodily behaviours in concert with others (and later solo) that establish a consensual domain of consistent motor-sensory relationships. These motor-sensory relationships bias perception in value-weighted ways that lead to higher-order behavioural control. Thus, control of vocal tract and other gestural activity mean that the gesture can be used to get others to fulfil one's needs and wants. The use of the gesture for higher-order control of this kind just is, from the child's point of view, the meaning of the gesture.

Mutual attunement via the acoustic medium to each other's vocal tract gestures is afforded by those more stable regions of vocal tract activity that, in time, give rise to the emergence of the phonological values of a particular population of speakers. In their articulatory and perceptual activities, speakers and listeners actively seek after and orient to (phonological) values in each other's vocal tract gestures that have the potential to alter their cognitive and affective relationships to their environment. The contrastive values that may be articulated and/or detected in vocal tract gestural activity are information that is specific to the affordances of that activity. The detection of such information has the functional capacity to change the agent's relationship to some aspect of its environment (Reed 1996a: 97). Moreover, the pick up of this information through the listener's exploratory activity of the highly structured energy field – the soliton-like wave – that acts on an excitable medium (air, body, brain) covaries with the agent's awareness of a change in the agent's relationship to the environment (see Reed 1996a: 98).

In my view, meaning, as distinct from value, is definable as just this reflexive awareness of the interpreted change that the detection of a given value brings about in the agent's (self-)awareness of its relationship to the relevant environment. The processes of mutual attunement described by Goldstein and Fowler (2003: 35-39) show how a given population of

languaging agents has evolved both biologically and culturally to be motivated to orient to and to relate to the specific affordances of vocal tract activity in particular culturally shaped and socially distributed ways that have specific values for that population (see Cowley 2008: 339-340; Section 15). Phonological values are, then, ecological facts, as defined by Reed (1996a: 100-103). They do not exist as a private mental code of stored phonological categories, but as normative, socially distributed constraints on vocal tract behaviour – behaviour that is public and observable. Infants are induced and rewarded by caregivers to seek out in their own and others' vocal tract activity the culturally salient phonological patterns that serve to promote culturally valued meanings and activities that may bring about modifications of agents' relationships to their environment.

19. The Second-order Reification of Vocal Tract Gestures as Speech Sound Segments

Mutual attunement as discussed above provides a plausible foundation, developmentally speaking, for the emergence of conventional resources that enable persons to solve the many coordination problems encountered by them in the different spheres of social life. Coordinated vocal tract gestural activity between persons in languaging behaviour requires a specific organization of this behaviour qua action systems that will tend to persist and recur at the population level as a consequence of the values they have for agents in the situations in which they serve to coordinate social interaction between persons. Phonological units and structures are, on this view, the result of the bonding of organ-specific gestures to form thermodynamically more stable action systems, or what I shall call *phonosemantic operators*, in utterance-activity between persons. Phonosemantic operators consist of vocal tract gestural scores in utterance-activity.

The term *phonosemantic operator* serves to remind us that (1) vocal tract gestural scores are not the arbitrary vehicles for the encoding of a separate, more abstract semantic or cognitive content by a sender and its subsequent decoding by a receiver; they are intrinsically meaningful (Section 4); (2) they function as operators on situations, their conventions and the understandings that persons have of the given situation and its associated conventions; (3) they are normatively constrained and culturally shaped and standardized resources that have the functional capacity to evoke situation conventions that enable them to solve coordination problems between persons in particular situations; (4) their conventionalization entails the typification of utterance-activity as differentiated kinds that are iterable across different situations and on different occasions such that they become abstracted from particular situations and standardized as cultural types through processes of 'grammaticalization'.

Iterability arises when writing systems serve to model the topological-continuous character of first-order dynamics in terms of discrete elements such as phonemes that are repeatable across speakers and occasions. One consequence of iterability in literate cultures is the standardization of speech sounds by means of institutionalized repetition in order to achieve error-free transmissibility as distinct from imitation of the topological-continuous aspects of speech (Harris 1991: 388). Harris explains this as follows:

> Transmissibility without error-compounding requires not only that the elements of the

> material to be transmitted be discrete, but also that they be pre-set in sender and receiver. When the speaker and the hearer are referring to a set of elements known to both, the hearer need receive only enough of a signal to distinguish a particular element – phoneme or word – in contrast to all other elements that could occur there. When the hearer then transmits (repeats) the utterance, he pronounces his own rendition of the pre-set (i.e. known) elements which he has distinguished. This means that both must learn to recognize a set of grammatical elements, primarily particular phonemes (or phonetic distinctions) and secondarily vocabulary (morphemes, words), in respect to which they speak and perceive utterances. It is this public institutionalization that makes the transmission of an utterance a repetition, whereas an attempt to redo or transmit something whose elements are continuous or not pre-set is an imitation.
>
> (Harris 1991: 388)

Pre-set elements are established in a population of speakers as norms of, for example, pronunciation and phonological spelling based on the presumption of the objective existence in the speech signal of segmental speech units, e.g. phonemes, which are taken as normative phonological units. Harris, in keeping with the overall formalist-structuralist premises of twentieth century linguistics, nonetheless preserves the idea that these segments are real features of speech sounds rather than the artefacts of a particular way of doing meta-language. He also assumes, incorrectly, in my view, that the continuous or non-discrete dimensions of vocal tract gestures are not also subject to processes of conventionalization. As Port (e.g. 2007: 351) shows, this view is a consequence of using the discrete, serially ordered graphic elements of alphabetic notation to transcribe and 'represent' the very different dynamics of talk as if the dynamical characteristics of talk can be attended to in terms of segments such as phonemes, syllables, and so on. The presumption by many linguists of the objective existence of segmental speech units belies the scientific evidence concerning the rapid perceptual learning that characterizes the infant's early encounters with the speech patterns of the ambient language (Port 2007: 353).

Infants learn to perceive and to model the very rapid patterns of co-articulated vocal tract gestures in ways that bear no resemblance to the much slower time-scales implied by the unnatural segmentation of co-articulated speech gestures into discrete grapheme-like segments, i.e. phonemes. Through statistical learning, infants achieve within the first year the ability to discriminate the salient articulatory and auditory patterns of the ambient language through constant habituation to the rapid rate of articulation of normal speech (Kuhl 2007). The mutual attunement of languaging agents in early language development to those regions of the vocal tract that afford partitioning into discrete regions and their repetition by speakers sets up an early perceptual focus on organ-specific contrasts that are amenable to their subsequent reification as segments on analogy with the graphic units of alphabetic notation.

Writing is a form of external cognitive and cultural scaffolding that affords the development of the skills of institutionalized reification of, attending to, repetition of, and transmission of sequences of discrete elements such as phonemes and syllables. These elements compress socially distributed and standardized information that enables languaging agents to recognize

determinate sequences of discrete elements as grammatical units such as words, phrases, sentences, and so on. Such units are repeatable and are, as Harris puts it, "uniquely appropriate for error-free transmission of utterances." (1991: 388-389).

Organ-specific vocal tract gestures are potentially contrastive. In the first instance, it seems likely that conventionalized proto-utterances or phonosemantic operators, consisting of holistic vocal tract gestural scores of co-articulated contrastive features, were associated with different aspects of experience through processes of statistical learning in a population. Organ-specific phonetic gestural contrasts were the basis of the discrimination or partitioning of the environment by increasingly standardized gestural scores into largely implicit representational topologies that formed the basis for the collective pooling of experience. Organ-specific contrasts proved amenable to reification as the basis on which co-articulated gestural scores were re-described on analogy with units of written notation as frozen segments of speech sounds. These segments form the basis of second-order phonological categories. The latter serve the meta-linguistic purpose of distinguishing between different classes of gestural score and subjecting these to increasing pressures of conventionalization. In this way, the reified aspects of standardized gestural scores and their distinctions are decomposed and then preserved as segmental distinctions in sound (e.g. pronunciation) that can be preserved under diverse conditions of transmission across persons and situations. Once discrete segmental distinctions emerge in this way, it follows that such distinctions, serially ordered in determinate sequences as norm replicating linguistic objects, would soon give rise to further meta-linguistic criteria for the segmentation of gestural scores into proto-word like objects that were defined according to the ways in which sound segments were combined. Moreover, the principle of decomposition at work here also means that these early word-like objects could be modified by the affixing of further standardized sequences of sound segments for the purpose of differentiating categories of experience such as time, plural, animate, gender, and so on.

Phonological forms, as we saw above, are socially distributed cultural types that constrain and shape vocal tract gestural activity. They provide standardized solutions to the problems of coordination that arise in emergent social situations. The conventionalization and standardization of phonosemantic operators leads to their re-categorisation as lexicogrammatical patterns (Sections 16, 25). The mutual attunement of speakers' vocal tract gestures afforded by the acoustic medium is an inherently and implicitly reflexive and self-organizing process in which persons seek to coordinate their diverse, not always congruent or harmonious, points of view. The achievement of interpersonal parity that is consequent upon mutual attunement in language development means that persons become increasingly sensitized to the context-dependent nature of different utterance-types qua phonosemantic operators and therefore to the ways in which utterances can have different effects and consequences, depending on the context. This goes hand-in-hand with the increasing understanding that different utterance-types have the power to evoke, operate on, and transform an increasing diversity and range of social situations.

In the first instance, utterances operate on implicit situation conventions and their reflexivities

that are social, but not linguistic. This is what infant protolanguage (Halliday 1975) (cf. Reed's 1996a: 167 "indicational language") achieves, i.e. the capacity to act upon and to transform participants' awareness and understanding of aspects of the non-linguistic situation that are the focus of proto-linguistic topics or arguments in, for example, joint attention-sharing routines. However, the point is not confined to infant protolanguage: predicational language also operates on non-linguistic social realities. The re-categorisation of first-order phonosemantic operators as second-order phonological and lexicogrammatical patterns and types is a further normative emergence such that second-order language is constituted not only as a means for operating on non-linguistic social conventions, but also as a means of conventionally creating and constituting social situations and realities through the grammaticalized resources of predication in second-order language. Moreover, if we see grammar not as a further level of formal linguistic encoding/decoding, we can better understand it as a further layer of reflexivity -- a fractal meta-linguistic process that views and refracts first-order languaging behaviour through the prism of second-order cultural constructs and their (always partial) perspectives. In other words, 'grammar' qua meta-language provides partial models of and perspectives on the shape of first-order languaging dynamics, usually in the service of particular social, institutional and cultural projects.

The enforced reification of segments through formal and informal teaching and learning in a given population sets up the possibility of predication. In the first instance, the selective analogizing of phonological contrasts associated with discrete organ-specific vocal tract gestures to discrete graphic notational units sets up the premises for the reification of the former as discrete segments of sound (quite contrary to the empirically determined evidence of the spatiotemporal overlap of vocal tract gestures in co-articulation). The resulting reified segments are, then, objectified as formal properties of the articulatory-auditory dynamics of speaking and listening in languaging behaviour that fulfil the criteria of error-free repetition and transmissibility discussed above. They therefore serve as normative constraints on how speakers are induced to perceive and articulate the far more complex dynamical properties of acts of speaking. In this sense, speech sound segments (phonemes, syllables) and their graphic analogues in written notation are operators that can operate on first-order languaging dynamics so as to transform aspects of the dynamics into recognizable serial orderings of static, non-overlapping, and context-independent phonemes or phonological distinctions that are, in turn, re-categorised as grammatical words, phrases, sentences, etc. The latter are abstracted away from the dynamics such that they are amenable to their transcription into written notation.

Importantly, these serial orderings of phonemes therefore take on a grammatical interpretation precisely because the serial ordering of phoneme segments corresponds to a grammatical meaning. Because the reified segments are conventional (normative) constraints that operate on first-order dynamics, we can say that the process of composing and generating serially ordered strings of such segments is productive. This is so in the sense that an unbounded set of strings or serially ordered sequences of phonemes can generate the conditions in which second-order language emerges as a conventionalized means of operating on and transforming first-order dynamics. At this point, second-order language emerges as a

conventional cultural affordance for operating on, structuring, and transforming (1) non-linguistic social conventions in situations; and (2) language itself as a conventionalized resource for operating on, structuring, and transforming itself qua system of social conventions (see Bickhard 2004) in ways that can transcend situation-bound first-order dynamics (Linell 2009).

20. Operator-Argument Dependency Relations

Once hierarchically organized serial orderings of discrete, de-contextualized formal objects are postulated as the basis of second-order language and its composition, then specialized notation systems such as writing systems, formal grammar, logic, and computer languages can be shown to have emerged as further technological and cultural elaborations of this basic principle, i.e. that language is a second-order 'code' consisting of discrete, standardized, serially ordered formal tokens on which operations of symbolic manipulation are performed. In this way, the operator-argument relation, rather than co-occurrence relations or probability of co-occurrence per se, defines the various ways in which formal tokens are organized in dependency relations such that an operator word, say, does not occur in a sentence unless a word of its argument set also occurs (Harris 1991: 332). The dependency relation expresses a meaningfulness that is more than the sum of the individual words that are combined in the dependency relation. Operators are dependent on their arguments and occur only in the presence of their argument words. Operators predicate something about the argument, i.e. they say something about it. The operator-argument dependency relation in its many forms is, then, the basis of grammatical units such as the clause. The dependency relations formed on the basis of co-occurrence relations of words yield a new kind of grammatical unit, the clause, based on the meaning-relation called predication. In this way, the clause constitutes, linguistically speaking, a unit of information about some referent situation that can be asserted, claimed, denied, argued about, refuted, and so on (Halliday 2004 [1985]: 106-121). Dependency relations are not confined to the clause and encompass morpheme- to discourse-level relations. Some examples follow:

1. The morpheme suffix *–er* in *teacher* is not just a suffix indicating the meaning 'agency', but an operator on the base morpheme *teach* such that the latter is transformed into an argument of further predication, e.g. the nominalization of the base morpheme Process. Thus, the nominalization specifies an agent that can function as the argument of a further predication, as in the clause *The teacher taught the students English*;

2. The adjective *phonetic* in the nominal group *phonetic symbol* is an operator on its argument, i.e. *symbol*. The operator predicates of the semantic class of Thing specified by *symbol* a particular sub-classification according to type;

3. In the clause, *The teacher taught the students English*, the Subject, *The teacher* and

the Finite [TEACH + PAST] are operators on the argument (predication) of the clause, i.e. *TEACH + the students English*. The Subject operator ties the argument (predication) to a deictically specified third-person non-speech participant, i.e. the referential noun phrase, *The teacher*. In other words, the predication is tied to and is about the Subject, which is contextually-sensitive to the local requirement that the Subject is correctly associated with its referent and that the referent is retrievable from the relevant local representational topology context (its reference is locally resolvable or interpretable). In other words, the entity specified by the Subject is immediately accessible to the interactants (first and second person) in the interaction (Davidse 1997). The Finite is marked for tense: it temporally grounds the argument in past time with respect to the here-now grounding of the utterance. The two operators therefore operate on and transform the argument of the clause into a fully grounded proposition, i.e. a particular dependence relation among the grammatical units composing the clause that constitute a unit of information founded on the predication relation, e.g. the Finite operator PAST operates on and transforms the (abstract) argument *The teacher TEACH the students English* into a unit of information that does not directly reflect any given real-world situation though it does relate to what we know and experience about English teachers and their students irrespective of the clause we have before us. Rather, the clause has the capacity to assert information about a situation – information that can be taken up, responded to, and transformed by the speaker or by others in various ways in the form of further operations on it. In this minimal sense, the clause has the capacity to operate on a given situation and to transform the situation and participants' understandings of the situation. It is this predicational capacity of language to operate on and to transform both non-linguistic **and** linguistic social realities that affords the emergence of virtual cultural entities that are constituted by means of the predicational capacity of language to operate on and to constitute virtual second-order social and cultural realities that exist only in off-line non-perceptual experience.

21. Languaging as Recursively Self-Maintenant Process

Languaging behaviour is a recursively self-maintenant process (Bickhard 2004, 2011; Thibault 2005a). It consists of a universe of objects that enter into constructive relations. As the examples show, albeit informally, a given trajectory through the object-space has the capacity to induce self-maintenant structures that are characterized by invariant patterns of transformation. Buss and Fontana (1996: 20) further point out that self-maintenant systems of this kind possess properties such as "regeneration, structure-dependent extension, complex substructure, capacity for hierarchical nesting" – properties that are characteristics of living systems (see also Abler 1989). The imposition of different boundary conditions on the object-space in the form of different meta-linguistic criteria generates different levels of organization, and a diversity of organizations within each level. Table 3 sets out a possible way of thinking about a flat ontology of languaging behaviour and the role of different fractal scales of meta-linguistic constraints. This can be compared to the more usual levels account in Table 4.

In Table 3, Level 0 is defined by self-copying expressions or simple ensembles of copying expressions, i.e. organ-specific vocal tract gestures. Level 1 refers to self-maintaining organizations of utterances (gestural scores) composed of Level 0 organizations. Level 2 is defined by self-maintaining meta-organizations (e.g. second-order lexicogrammar) composed of Level 1 organizations composed of Level 0 organizations.

Time scale		**Constraints**	**Relationship between Scales**	
Cultural-historical	Lexicogrammatical forms and discourse categories and patterns as second-order re-descriptions of standardized phonology and combinations as words, combinations of words, etc.	Quasi-causal constraints and attractors from cultural-time-scales: population-level dynamics		
Cultural-historical	Low-dimensional phonological forms as population-level socially distributed and normative standardizations of phonetic gestures influenced by pronunciation norms and writing	Quasi-causal constraints and attractors from cultural time-scales: population-level dynamics	Time-invariant, contextually-invariant, low-dimension (sparse)	Level 2: self-maintaining meta-organizations composed of Level 1 organizations composed of Level 0 organizations
Real-time events between persons	Rich high-dimensional phonetic gestures of first-order languaging behaviour	Real-time bio-physical processes coordinating persons-in-interaction	Time-varying, contextually-variant, high-dimensional (rich)	Level 1: self-maintaining organizations of utterances composed of Level 0 organizations
	Organ-specific phonetic gestures as primitives; the smallest kinetically persistent and self-maintaining entities of the phonological repertoire of types specific to a given language	Molecular constellations formed by combinations or syntheses of organ-specific gestures as kinetically persistent phonetic organizations		Level 0: simple ensembles of self-copying expressions

Table 3: A flat ontology of languaging behaviour and constraints on different scales

Language Strata: Hjelmslev	**Language Strata: Halliday**	**Relationship between strata**	
Content substance	semantics		
Content form	lexicogrammar		
Expression form	phonology	Abstract phonological categories and forms and their combinations	Phonological form realizes/is realized by
Expression substance	phonetics		

Table 4: A levels account: Stratification of language according to Hjelmslev and Halliday

Infant protolanguage (Halliday 1975; c.f. Reed's 1996a: 158-161 indicational language) operates on a more restricted range of conditions in the immediate situation: non-predicational or proto-linguistic utterances indicate topics or arguments without explicitly predicating something about the argument. Topics typically indicate objects, properties, events in/of situation and function to make the infant's interlocutor aware of the infant's focus of interest or attention; elementary meanings operate on situations and transform them in this way. Predicational language means that the operator is something that is said about or predicated of its argument. The operator does not occur in a clause or sentence except in relation to a particular class of argument words that are also present (Harris 1991: 332). In other words, the operator can select amongst a potentially large set of potential arguments about which it can predicate something according to the demands of a wide variety of actual and anticipated situations. This capacity means that languaging behaviour is recursively self-maintenant because it can vary how it contributes to the condition of self-maintenance of languaging agents when they engage in such behaviour. The morpheme suffix -ER in *teacher* (see above) already shows this principle at work. As we saw, -ER is an operator on its base (TEACH); it recursively operates on the base and predicates something of it. The resulting nominal is a normative emergence that itself can function as an argument in ways that can be recursively operated on by multiple potential operators in response to a wide range of environmental conditions and situations that may or may not be supported by the given environment (Bickhard 2004; Freeman 2000: 97; Section 8). Grammatical forms do not encode aspects of the environments in which they occur; they specify linguistically constituted information that can partition or differentiate the environment in ways that may or may not be judged as appropriate, correct, truthful, and so on, in the situation.

As Bickhard (2004) has shown, the interaction may or may not be supported by the given environment. If it is, it will succeed; if not, it will fail. Interactional success and failure are judged with respect to the anticipated outcomes – what could or should happen – given the situation. These outcomes are available to agents as feedback that can guide future learning. On the other hand, the judging of linguistically constituted information that is asserted about some situation as appropriate, truthful, believable, correct, sincere, etc. with respect to some current locus of cognitive or perceptual processing in the situation gives rise to a normative

notion of linguistic information. Information in this sense is concerned with what is predicated about the given locus of cognitive or perceptual processing qua linguistically constituted argument (see Allen and Bickhard 2011: 109). Languaging behaviour is one means whereby persons can vary how they contribute to their own condition of self-maintenant being-in-the-world. Language evolved above all as a further extension and augmentation of the possibilities afforded by the emergence of consciousness in many living species, i.e. the optimization and management of the life process of the self and the self's responses to its environment (Damasio 2010: 267). Damasio makes the following pertinent observation concerning the internal images that provide more precise information about the environment of the self:

> The lion's share of the advantage, I suspect, comes from the fact that in a conscious mind the processing of environmental images is *oriented* by a particular set of internal images, those of the subject's living organism as represented in the self. The self focuses the mind process, it imbues the adventure of encountering other objects and events with a motivation, it infuses the exploration of the world outside the brain with a *concern* for the first and foremost problem facing the organism: the successful regulation of life. That concern is naturally generated by the self process, whose foundation lies in bodily feelings, primordial and modified. The spontaneously, intrinsically feeling self signals directly, as a result of the valence and intensity of its affective states, the degree of concern and need that are present at every moment.
>
> (Damasio 2010: 267-268; italics in original)

Language is a socially and culturally distributed resource for maintaining the socio-cultural homeostasis of individuals and social groups. Many theories of language emphasise above all its relationship to 'mind.' Damasio's focus on the self and its role in focusing the mind process draws attention to the way in which languaging behaviour orients selves in their worlds in value-weighted ways. I have already argued that lexicogrammatical differentiators are normative, value-weighted patterns that attract first-order languaging dynamics to their basins of attraction. They enable selves to orient to normative cultural objects, events, etc. – actual and virtual – that are productively differentiated by lexicogrammatical differentiators. By the same token, they also relate selves and their actions to norms. Linguistic norms therefore have to do with the relationship of the norm both to its objects and to its subjects (selves) (see Macherey 2009: 74).

As Allen and Bickhard point out, self-maintenant systems, more generally speaking, have a normative stake in maintaining themselves (Allen and Bickhard 2011: 108). Languaging behaviour cannot be separated from languaging agents and treated as disembodied 'text' (Section 1) without jettisoning any prospect of explaining the language-person system as a class of recursively self-maintenant system in the sense defined here. In this perspective, second-order language is a culturally distributed form of 'global order' that is emergent from the local order of self-organizing first-order languaging dynamics. By the same token, the emergent organization of second-order language is a further normative emergence such that lexicogrammatical differentiators in utterance-activity specify and predicate information

about (aspects of) situations that may or may not be true or false, etc., as discussed above.

Dialogically coordinated interactional behaviour between persons elicits interpretation (Cowley 2011). Vocal tract and other gestures (facial expression, hand-arm gestures) compress information on account of the dynamical co-articulated properties. Information is used here in Gibson's (1986 [1979]) sense. Information is specificational. Thus, the information in vocal tract gestures specifies something about both the person articulating the gesture (propriospecific) and about some aspect of the situation and the person's orientation to the situation (exterospecific). Information is not the same as meaning. Meanings, on the other hand, are created by the unique, experience-grounded neural constructions of each individual person. Vocal tract and other gestures elicit and prompt for interpretation in the situations in which they occur; they do not contain or encode meanings or serve to convey or transmit these meanings from one person to another. When we concert our meanings with others in dialogically coordinated languaging activity, we make an effort, to varying degrees and with varying levels of motivation and commitment, to assimilate the meanings we interpret in others' gestures to our own viewpoints, as constructed and articulated by own bodily and neural dynamics (Freeman 2000: 15).

Seen in this light, the second-order lexicogrammatical patterns that we detect in the ambient energy fields created by vocal tract and other gestures are forms of cultural scaffolding to which interactants orient in order to augment the assimilation of each others' meanings in ways that may lead to, though they do not guarantee, common understanding. Grammar is a normative emergence in the sense discussed above that serves to solve coordination problems of precisely this kind. The normative character of lexicogrammatical constraints on languaging behaviour also serves another important purpose. The very rapid time-scales of first-order languaging activity simply are not fully available, perceptually speaking, to interactants. We do not have full awareness of the complexities of languaging behaviour, though they are usually under voluntary control and are intentional. Many aspects of the dynamics of languaging behaviour are rapid, continuous and not accessible to conscious monitoring (see also Port 2007: 352). However, we obviously do consciously attend to the articulatory and auditory properties and processes of vocal tract behaviour. We can also recall aspects of these.

Attending to and recalling vocal tract behaviour is a skill that can be improved though what Gibson (1983 [1966]) called the 'education of perception.' We are able to sample and bring to awareness for one's self and for others selected aspects of the process. The normative character of lexicogrammar thus provides abstraction amenable cultural resources for selectively sampling some aspects of the far more complex processes involved in languaging behaviour, e.g. when we use standard orthography to write down what someone has said. Such samples are second-order textual records of the primary data (Thibault 1994). Unfortunately, many scientific accounts conflate the primary (first-order dynamics) data with just such second-order samples and develop their theories accordingly (Section 1). Samples serve to bring some aspects of the process to awareness for particular, limited purposes. Moreover, these samples themselves can be embedded in and operated on in further acts of

interaction between persons.

Linguistic catalysis is, of course, a dialogical activity between persons. The capacity of the face, vocal tract, and hand-arm systems to differentiate very many fine-grained sensorimotor discriminations and their productive possibilities of combination and recombination meant that these body systems and their synergies came to be used for the selective orientation to and exploration of the environment both in collaboration with other perception-action systems and in concert with the other persons with whom one interacted. The increasingly delicate and discriminatory patterns of difference that could be articulated through these bodily resources meant that an increasingly wide diversity of meanings could be catalysed in relation to a widening range of situations. In the case of dialogically coordinated linguistic catalysis between persons, the boundary conditions include the higher-scalar ecosocial arrangements that constrain dialogically coordinated languaging behaviour between persons. Meanings emerge as a result of the interactions between the neural and bodily dynamics of individual persons and their ecosocial environments.

The emergence of higher-scalar boundary conditions entrains these dynamics to cultural patterns such that the vocal tract and other sensorimotor discriminations of a population of interacting agents are constrained by the same attractors in ways that enable the patterns of difference to be taken up and adaptively modified by others. In this way, individuals get linked to each other across time and space. These emergent patterns of differentiations are not encodings of meanings, ideas, or thoughts that are then transmitted to others, who in turn decode them (Section 4). Instead, these fine-grained discriminations and their possibilities of productive combination and recombination extend and augment the capacities of agents to explore, act on, interact with, experiment with, and analyse the emergent possibilities of their worlds (Thibault 2004a: 187).

In exploring their world in this way, agents assimilate their brain and body dynamics to it (Freeman 2000: 141). They do so by creating resonance patterns between their neural and bodily dynamics and selected aspects of their worlds through the discriminatory possibilities of this exploratory activity. As Freeman (op. cit.) points out, such activity is hypothesis testing – it seeks responses from the world that confirm or disconfirm its hypotheses, as also suggested by the notion of interactive success and failure discussed above. Languaging agents adapt to and assimilate to each other's dynamics through reciprocal exploratory activity and hypothesis testing (Freeman 2000: 141). In doing so, they "create transcendent social entities that enhance and empower the individuals." (Freeman 2000: 142).

Hypothesis testing is not based on encoding/decoding, but on the very different principle of differentiation. The many fine-grained differentiations and their productive combinations made by the face, vocal tract, and hand-arm systems are internal to the organism. In exploring and testing their environments, languaging agents interact with their environments. Differentiations do not encode/decode a meaning or content that is external to the individual (to the means of encoding??). Rather, through its interactions with its world, the agent differentiates or partitions the world by interacting with it and in the process selectively applying its own internal capacity for making fine-grained discriminations, e.g. in phonetic

gestures, in ways that can be tested against the environment itself as a differentiation that is appropriate, truthful, correct, relevant, and so on. If the agent-internal differentiation is of a kind that is appropriate to the given environment, it will be selected and confirmed by the relevant environment as one that is successful, appropriate, correct, and so on. If not, the differentiation will be de-selected and disconfirmed as unsuccessful, inappropriate, incorrect, and so on, and therefore judged to be unsuccessful, unwarranted, invalid, and so on.

The 'content' of a differentiation is not, then, arbitrarily and externally attached to a material sign-vehicle. In Halliday (2004) we saw that the content is an ecosocial experience that is coded by the coding of the semantics by lexicogrammar (Section 7). On this reading, the content – experience or social relation – derives from the external ecosocial environment and is explicitly encoded by the semantic categories of the language. Differentiations don't work like this. Phonetic gestures, for instance, are not the means for encoding/decoding experience as semantic content. Instead, they specify an environment that is presupposed as being one that is appropriate (or not) for the indicated differentiation. The phonetic gesture and its differentiations therefore specify the future-oriented interactive potentiality of the environment. Insofar as it is a hypothesis-testing act, a given phonetic gesture qua complex, multidimensional differentiator anticipates interaction possibilities that the phonetic gesture presupposes and explores and which the given environment may or may not support. The phonetic or other gesture is an explicit organization qua differentiator of interactive potential. It implicitly presupposes the contextual conditions that would support the given differentiator. By the same token, the given environment may not support these interaction potentialities such that the differentiation fails in the given environment.

Languaging Behaviour as Catalytic Process: Steps Towards a Theory of Living Language

[PART II]

Paul J. Thibault,

Faculty of Humanities, University of Agder,
Kristiansand, Norway
Faculty of Education, University of Hong Kong,
Hong Kong SAR, PR China

22. The Specificational Character of Linguistic Information

Linguistic information persists in the environment on the long, slow timescales of culture. Information in the Gibsonian perspective serves to regulate behaviour and, potentially, to change the animal's relationship to its environment. Moreover, information is a resource that must specify other environmental resources, as Reed (1996a: 48) points out, without being the same as those other resources. How does this apply to languaging behaviour? In Gibson's theory, information in the environment exists in structured ambient energy fields. The task of the animal is to locate the structure in such fields (Reed 1996a: 48). Information is specificational in Gibson's theory – it functions to specify particular environmental affordances. Structured ambient energy field remain invariant under certain transformations caused by the animal. Such invariants serve to specify persisting environmental resources. Information in a structured ambient energy field is a complex structuring of contrasts or differentiations caused by the ways in which an environmental event (e.g. a slamming door or a person's vocal tract activity) causes the acoustic or other medium to be restructured such that perceivers can become coupled to the source of the disturbance of the array, i.e. the environmental event.

Languaging behaviour, I suggest, makes available to both the self and others just such structured ambient energy fields. The detection of invariants in this field includes the phonologically salient contrasts created by organ-specific vocal tract gestures (Section 15). Linguistic information as discussed above specifies something about the relevant environment in which it is uttered. Information of this kind remains invariant irrespective of what particular individuals do. This is what Gibson (1986 [1979]) called exterospecific information. Gibson distinguishes this kind of information from propriospecific information, which depends upon what the organism is doing now. The latter specifies how the animal is engaging with its environment. Both kinds of information are simultaneously available in the structured ambient energy fields generated by languaging behaviour. Exterospecific linguistic information includes the information generated by predicational relations in language. It is

information about the environment. Propriospecific information includes information generated by the bodily dynamics of languaging behaviour. It is information about how the animal's own activity regulates its behaviour in the environment. In principle, the former can be used independently of any action on the environment.

By the same token, propriospecific information can be used independently of knowledge obtained from the first kind of information (Reed 1996a: 49). Thus, languaging agents can deploy propriospecific information to orient to and explore a structured ambient energy field caused by someone's vocal tract activity without knowing the informational content in the first sense. They can apply their sensorimotor knowledge (Noë (2004: 9; Section 13) before necessarily being able to activate a content. Persons can regulate their own perceptual exploration of such structured ambient arrays by regulating their own orienting to, attending to, and engaging with the array on the basis of propriospecific information about their own languaging behaviour – both uttering/articulating and responding. Persons monitor their own and others' bodily dynamics in ways that enable them to adjust to, calibrate with, and anticipate the perceptual exploration of the many fine-grained discrimination that enable them to regulate their own languaging behaviour.

This view differs from the standard one. According to the standard view, language is stratified into different levels of expression and content. Expression is the vehicle for the encoding, transmission, and decoding of content by sender and receiver, respectively. In this view, language gets separated both from persons and from the environment, as in Halliday's 'interface' view of the relationship between the content stratum of language (semantics and lexicogrammar) and the ecosocial environment (Section 7). On the other hand, when we see that utterance-activity carries the two kinds of information identified by Gibson, we have a principled way of showing how information in this dual sense unifies organism and environment. This is so because the simultaneous presence and pick up of the two kinds of information in languaging behaviour is grounded in active, embodied agents engaging with their worlds through their language behaviour rather than formal realizations of a linguistic computational system in which different levels of formal linguistic organization provide the coded inputs and outputs to each other. The latter kind of theory has no place for the fact that languaging agents are aware of their own activity insofar as propriospecific information is available and can be detected by them (Reed 1996a: 58) as an essential aspect of languaging behaviour that requires explanation. Halliday's account is a variant on the standard account that divides semiotic processes into three kinds: input (experience of the ecosocial environment), output (bodily activity, e.g. articulation), and higher-order integration of the two (phonology and lexicogrammar as mediators and organizers of input and output).

The current approach, on the other hand, emphasises real, living languaging behaviour of persons and how they actively explore and pick up and/or make information available and how they use that information to regulate their own and others' behaviour and awareness. This stems from the fundamental insight that action and perception are not separate, but are two aspects of the same active process of exploring the world. Instead of sensory inputs from the ecosocial world that are then reconstructed by the brain as meaningful semantic

information by a coding/realizatory process across formal levels of input/output, the ecological view sees languaging activity as a process of active exploration and detection of information that is made available in structured ambient energy fields. The task of languaging agents is to find ways of both making it available and detecting it. The task of doing so is a form of exploratory activity as agents apply what Noë called "sensorimotor knowledge" (2004: 9) in order to mediate the transitions between invariants in the processes of linguistically catalyzing their worlds.

Vocal tract behaviour of persons causes and makes available to self and to others structured ambient energy fields. These afford possibilities of information exploration and pick up. The affordances of such events constitute a complex network of interactive possibilities along many dimensions of contrast that are potentially available to perceivers in the structured ambient energy fields made available in the environment by languaging behaviour. These many and overlapping possibilities are accessible in the first instance through the application of sensorimotor knowledge and skills. The various interactive potentialities of the given utterance are therefore potentially available and accessible through the exploration of the multiple dimensions of contrasts both typological-categorial and topological-continuous. For example, one can attend to organic-specific segments while also attending to the intensity of the speaker's vocal activity. Moreover, the interactive potentialities will persist over the time frame of the utterance (Section 15).

23. Interactivity, Learning, and Languaging

The state space of vocal tract gestures is not a pre-constituted knowledge or system of possibilities that speakers realise, but a virtual learning space in relation to which one composes "the singular points of one's own body" in conformity to this problem space. Deleuze expands on this point with respect to his concept of the Idea, as follows:

> In fact, the Idea is not the element of knowledge but that of an infinite "learning", which is of a different nature to knowledge. For learning evolves entirely in the comprehension of problems as such, in the apprehension and condensation of singularities and in the composition of ideal events and bodies. Learning to swim or learning a foreign language means composing the singular points of one's own body or one's own language with those of another shape or element, which tears us apart but also propels us into a hitherto unknown and unheard-of world of problems. To what are we dedicated if not to those problems which demand the very transformation of our body and our language? In short, representation and knowledge are modelled entirely upon propositions of consciousness which designate cases of solutions, but those propositions by themselves give a completely inaccurate notion of the instance which engenders them as cases, and which they resolve or conclude. By contrast, the Idea and "learning" express that extra-propositional or sub-representational problematic instance: the presentation of the unconscious, not the representation of consciousness.
>
> (Deleuze 2004 [1968]: 241)

The state space of vocal tract gestures is a state space of virtual problems that is not opposed to the real, but to the actual (Deleuze 2004 [1968]: 263). Deleuze remarks: "The only danger in all this is that the virtual could be confused with the possible. The possible is opposed to the real, the process undergone by the possible is therefore a "realisation". By contrast, the virtual is not opposed to the real; it possesses a full reality in itself. The process it undergoes is that of actualisation." (Deleuze 2004 [1968]: 263). As discussed in Sections 8 and 26, the productive differentiations that are articulated in vocal tract gestural behaviour are not the realisation of a prior set of possibilities, but the actualization of individuated solutions to the virtual problem spaces constituted by language. Interactive differentiations just are such actualized solutions (Sections 5, 8, 17, 21, and 26).

Humans learn best in situations that promote rich, culturally saturated interactivity when they engage with and manipulate external artefacts to solve learning tasks and cognitive problems in often complex environments such as aircraft cockpits (Hutchins 1995b), interpreting fMRI brain scans by brain scientists (Alač and Hutchins 2004), and medical simulations involving senior doctors and trainee doctors (Steffensen, Thibault, and Cowley 2010; see also (Hutchins 1995a, 2010; Clark 1997, 2008; Kirsch 1995a, 1995b). Interactivity and the forms of coupling of agents to their environments that it enables is not reducible to low-level perceptual-motor skills, but is central to higher-order cognitive operations in complex environments requiring expert knowledge. Moreover, a long tradition of work in experimental psychology (Koffka 1910, 1935; Luchins 1942; Vallée-Tourangeau et al 2011) shows that what Koffka (1910) identified as *latente Einstellung* ('latent attitude'), or experience-based predispositions to learning, can influence learning negatively and therefore can guide learning in inefficient ways that delay or frustrate desired outcomes (Kirsch 1995a; Vallée-Tourangeau et al 2011).

Building on Koffka's insight, the experimental works of Luchins (1942) and Vallée-Tourangeau et al (2011 indicate the potential of interactivity to diminish negative predispositions towards learning. Unlike learning based on text-based models or mental simulation inside the individual's head, a rich, dynamical multimodal environment consisting of manipulable artefacts affords a changing array of affordances and possibilities of perception and action. These affordances and possibilities attract and shape attention and action. In doing so, they constrain action, knowledge and cognition in ways that seem more likely to promote positive learning experiences and outcomes. Interactivity with persons, artefacts, tools, and technologies in the physical and cultural environment enables learners to segment and identify the features of that environment so that they develop more effective learning strategies. Interactivity through visual scanning, haptic manipulation and exploration, sound, and movement enables learners to manipulate and re-organize the physical aspects of the learning task such that active exploration and manipulation of physical artefacts gives rise to new perceptions and new differentiations. In turn, these can transform the learning task.

These findings have important implications for the virtual forms of interactivity made possible by languaging. Generally speaking, languaging agents use the action-perception possibilities of their bodies to interact with aspects of their environments in ways that augment and enhance differentiation and therefore learning and cognition. Interacting with the

virtual environment constituted by second-order cultural patterns is an extension of this basic fact of human cognition. First-order interactivity is based on the exploration and detection of visual, spatial, kinesic, tactile, and other modes of action-perception. First-order languaging provides a rich, dynamic multimodal environment that both extends participants' grounding in the familiar physical world and makes available new kinds of manipulable perceptual input grounded in virtual worlds. Interactivity takes place in a richly multimodal interactive environment. This environment draws persons' attention to the features of problems that may help them to adopt more effective learning strategies. First-order interactivity enables agents to manipulate and restructure the nature of cognitive tasks. The cognitive task becomes a changing, dynamical multimodal configuration that reveals new affordances during the agent's time-bound interactions with the task.

The cognitive environment and thus the appeal of this kind of richly multimodal environment to many people, I hypothesise, is 'richer' and more 'intuitive' than text-based technologies and practices because it affords the manipulation and reshaping of the cognitive task through very natural, intuitive activities of our bodies such as touching, moving, pointing, visual scanning, talking, and auditory prompts and cues of various kinds (Hutchins (1995a, 1995b, 2010; Kirsch 1995a, 1995b; Clark 1997, 2008; Wilson and Clark 2009; Vallée-Tourangeau 2011). In other words, languaging is grounded in and extends the natural interactivity of human bodies. Multimodal action-perception cycles and their integration though interactivity mean that socially coordinated languaging can transform the presentation and representation of the on-line learning task in ways that regulate and shape learners' attention, perception, and cognition. These processes are based on interactive differentiation rather than the encoding of transduced inputs. Differentiation arises as a result of active exploration; differentiations are not givens and forms of interactivity that are adequate to the task of yielding such differentiations must be devised. Rather than the r\epresentation of an already settled reality by discursive modeling, interactivity couples persons to matter-energy flows and material practices so as to influence and guide them as well as to discover new things about them. This avoids the conservative essentialism whereby discourse-analytical approaches, for example, reduce everything to man-made categories of mind. The far more unruly material reality on diverse scales and their histories cannot be settled in this way. Interactivity is more concerned with intervening experimentally in the material world and seeing how it responds rather than the discursive modeling of realities already assumed to be settled in advance by the linguistic and other formalisms of the human mind and their conventions.

24. The Feel of Languaging Events: Value-weighted Dynamics

As we saw in Section 9, languaging agents exploit timing to register the value-weighted significance of visible, audible and sensible patterns in pico-scale events. They register in core consciousness the differentiated feel of such events and, drawing on a history of their participation in language practices, they produce a response based on a sense of felt agency in relation to the resulting 'feeling of knowing' (Damasio 2010: 209). That is, agents perceive that a given behaviour and its dynamical properties have an evaluative significance that has the functional capacity to influence the behaviour of self and others. How this influence

works depends on both the physical properties of the vocalization or other gesture and the perspectives on it that are adopted by speakers, listeners, and observers of such events in the dialogical array. Values in turn may be associated with particular sets of circumstances in the life of the individuals concerned. Far from being arbitrary, the acoustic and other dynamical properties of vocalizations are structured in ways that orient the attention, arousal, and motivation of listener-observers owing to the direct links between the auditory periphery and brain stem regions regulating whole-body arousal, feeling states, and action (Section 4). The dynamical bio-physical properties of vocalizations elicit immediate, whole-body responses on many different scales of organismic organization which are functional in responding to and managing different types of situations.

Prosody can be affectively and intentionally modulated to these ends. The question arises as to how prosody and responses to particular dynamical (acoustical and other) properties of different kinds of prosody in vocalizations can provide critical scaffolding for learning about different kinds of situations. This is so because the links of the auditory periphery to brainstem regions having to do with affect, arousal, and motivation also link directly to the amygdala and hippocampus (LeDoux 1996, 2000). Hence, different bio-physical properties of vocalizations may help infants to learn about particular kinds of situations and to remember the details of the action scripts that are called forth by these situations. The bio-physical properties of vocal events act directly on the neural-hormonal, perceptual, affective, and learning systems of listener-observers in ways that enhance adaptive, flexible behaviour and its development. The expressive bodily dynamics of languaging agents do not so much express or reveal the inner feelings of agents; instead, the dynamical properties of vocal, facial and other bodily expression have the capacity to induce feelings in self and others (Laird 2007).

Laird (2007: 13) uses the example of a smile to show this: a feeling of happiness is not directly caused by the smile, "but [is] rather a product of the relationship between the smile and the context in which it occurs." (Laird 2007: 13). The feeling of happiness is therefore an evaluator of the situation in core consciousness. The smile is part of a socially displayed emotion script that serves to co-regulate agents' relationships to aspects of their worlds that are made salient by the relationship between the feeling in core consciousness, the smile, and the situation. The smile is part of the coordinated dynamics of inter-individual interaction behaviour; it enables persons to see aspects of situations in value-weighted ways according to the perspectives that are triggered when the protoself in core consciousness is perturbed by the coordinated dynamics of such behaviour (Section 11). Languaging dynamics accordingly partition the world in ways that pave for way for the emergence of compressed possibility spaces, i.e. digital semantics, in symbolic neural space.

People make us of the dynamics of vocal, gestural and other forms of bodily expression or first-order languaging behaviour to coordinate perspectives and to move people – affectively, cognitively, and interactively -- in ways that are increasingly decoupled from on-line perception (Deacon 1996). Individuals respond to these dynamical properties of such events. In learning to do so, they engage in meta-linguistic practices that draw attention to aspects of

the dynamics, i.e. wordings, (Cowley 2008: 339) and how these are or may be integrated with a cultural universe of virtual entities that are evoked by lexicogrammar qua second-order cultural patterns (Thibault 2011) and which afford 'mental time travel' (Corballis 2011: 100-111). These practices put the emphasis on the primacy of interactional history and also on how the relevant brain structures evolved as adaptations for coordinated, inter-individual first-order languaging, rather than on abstract, disembodied systems that are said to process linguistic forms in the brain/mind or in text/discourse, as in mainstream formal, functionalist, and cognitive accounts of language. The question then becomes: how do individuals integrate wordings with bodily dynamics? In first-order languaging events, individuals attend to how aspects of the dynamics can be integrated to locally salient cultural meanings and values. The dynamics modulate wordings and bind them together into locally coherent units of embodied sense-making at the same time that the dynamics are constrained by wordings.

Dynamics have intrinsically unstable properties that set "the brain onto an itinerant trajectory, that is, a pathway through a chain of preferred states, which are learned basins of attraction." (Freeman 2000: 121). The hearing of wordings in the dynamics constitutes a local, momentary capture of aspects of the dynamics by a basin of attraction, thus providing an awareness of semantic closure. Our experience of the dynamics of languaging behaviour is a consequence of our time-bound exploration of the sensorimotor contingencies of the dynamical properties of utterances and of our assimilation of these properties to basins of attraction that enable us to experience wordings. Wordings qua lexicogrammatical forms cannot be directly 'decoded' from they dynamics; they were never 'encoded' into them in the first place. Freeman (2000: 120) shows how the brain learns about environmental events by organizing the dynamical properties of the brain and the body in ways that make them similar to the event that is encountered (Sections 14, 21).

As Gibson (1986 [1979]), Reed (1996a, 1996b), Hodges (2007), and other ecological psychologists have shown, persons actively explore environmental events and in the process of exploring them, they can change their perceptions of them. This does not mean that the event imprints its own features on the brain. We experience words in dynamical speech events not as features that are encoded in these events and which then impress themselves upon our bodies and brains in a passive decoding process. Instead, through repeated, habitual yet active participation in action-perception cycles, our perception of events is educated such that the sensory receptors are positioned though efference and the sensory cortices through preafference (Freeman 2000: 121). In other words, we learn actively to attend to such events and their dynamical properties so as to achieve an "optimal assimilation of the self to the object." (Freeman 2000: 121). In this explanation, there is no requirement that neurons decipher and decode perceived events. Davia comments on the alternative suggestion as follows:

> It suggests instead that the significance of a given neural event is that it occurs at a certain place and at a certain time within an overall context of neural events taking place in space and time. It is the direct correspondence between the relative position and timing of events in the perceptual field and the brain that is of primary importance.

> Indeed, there is a direct topological mapping between the retina and the V1 layer in the visual cortex. What this means is that a triangle in the visual field, for example, translates to a triangle of neural excitability and/or suppression in the V1 layer. The relative spatial and temporal relationships implicit in the triangle are preserved in terms of how the medium is structured.
>
> (Davia 2006: 278)

and

> Because the suggested mechanism of brain processing is the soliton (or travelling wave), the relationship between soliton formation and the structure (invariance or symmetry) of the environment may offer an insight into how the brain unifies aspects of the environmental stimulus into the familiar objects and events of perception. For an object or event to become an object of cognition it must embody invariance or symmetries, what is being called "structure", in space and/or time. If the potential object or event of perception embodies structure, then via the senses, that structure translates to excitable areas in the brain that embody invariance or symmetries in space and/or time. Travelling waves in the brain sustain themselves by releasing this energy, hence, they constitute a process of catalytic mediation. Thus, the particular solitonic or travelling-wave solution is dependent upon the structure (invariance, symmetry) of the objects or events of perception.
>
> (Davia 2006: 278)

The symmetries and invariances of vocal tract gestures qua environmental events structure the acoustic medium; they depend upon the symmetries and invariances within the medium and its boundary conditions (Davia 2006: 280). As Davia argues, this gives rise to a soliton-like travelling wave solution to the problem of the implicit order in the event that is encountered:

> It may seem that what is proposed is that travelling waves progress through the brain in much the same way as they progress through an excitable medium. This is not the case. As was mentioned previously, we can choose to observe a soliton such that the soliton remains in a fixed position, maintaining its organization, while mediating a set of transitions. The brain is posed with a similar problem. Whatever neural activity corresponds to the perception of an event, it must maintain its integrity, given the set of transformations in space and time imposed upon the brain as a consequence of the transformations in space and time embodied by the stimulus. As Gibson [24] and other perception psychologists have emphasized, we should think of the act of perception as dynamic, not static. There is a constant stream of stimuli, much like the constant stream of water moving through the soliton in the canal, and the essential geometric and temporal relations are preserved as an aspect of neural activity.
>
> To illustrate this point, we can visualize a simplification of a "data stream" imposed upon the visual cortex by a triangular stimulus, schematically depicted in Figure 2. If

we imagine the situation of a child who is too young to recognize a triangle, the same set of transformations is imposed upon the visual cortex as in the case of an older child who is able to recognize the triangle.

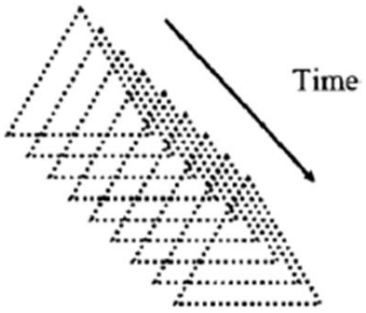


Figure 2: A simplified representation of the data stream in the visual cortex as a result of looking at a triangle

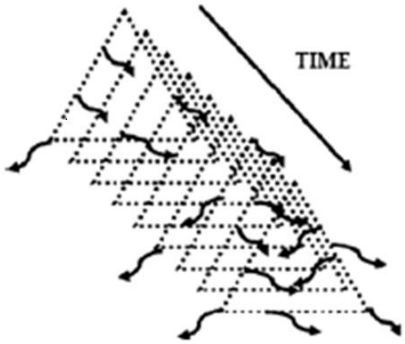


Figure 3: A simplified representation of activity in the visual cortex of a child who is too young to recognize a triangle

However, the energy that is released as a result of the stimulus disperses in a chaotic way (see Figure 3). Feedback relationships among neurons that fire as a consequence of stimulus may be formed momentarily among aspects of the triangle and other objects in the visual field. These cannot be sustained, however, because there is no ordered relationship among these different stimuli; there are no invariants or symmetries. They may be moving quite randomly with respect to one another, and so any relationships that may be formed quickly become unstable. However, solitons or travelling-wave solutions depend upon symmetries or invariance within the medium and its boundary conditions. It is the order that is implicit in the stimulus, the symmetries or invariance, that facilitates the emergence of travelling-wave solutions.

(Davia 2006: 279-280)

Solitons can maintain their organization while mediating a set of transitions. The two parity constraints identified by Fowler (2010) – i.e. (1) phonetic gestures are public events and (2) they persist in time (Section 15) – are important here. As we saw in Section 14, languaging agents manufacture a robust semantic synthesis – a stable semantic artefact – when they are

able to unify the spatiotemporal invariances of the soliton-like waves of languaging behaviour with relevant covarying aspects of situations, prior experience of and a past history of participating in such events, intentions, socio-cultural norms and values, and the relevant higher-scalar boundary conditions. Second-order cultural constructs such as lexicogrammar are boundary conditions in this sense. The travelling waves of languaging behaviour can thus maintain their organization at the same time that they mediate a set of transitions in their environment. Davia (2006: 281) explains that this is so because a soliton is "always continuous with its boundary conditions." It is always completely integrable to and constitutes a solution to its boundary conditions.

In Halliday's (2004) account of the relationship between language and the individual person's experience of the ecosocial environment (Section 7), it is somehow assumed in ways not made very clear that there exists a world of ecosocial experience with which the semantics of language 'interfaces' and reorganizes through lexicogrammar as the semantic categories of our experience. But this is a bit like putting the cart before the horse: we must first analyse the dynamical properties of vocal tract gestural events and the ways in which they both dynamically structure their environments and the awareness of languaging agents. The basic problem is the deeply rooted assumption that there just is a world out there that language interfaces with and in some way 'represents', 'encodes' or 'transduces' into the categories of language. A different theoretical solution is provided by Davia's (2006: 281-283) distinction between cognition as representation and cognition as ontology. Davia explains this with the example of a mountain:

> … we often take for granted aspects of the world revealed to us by our minds as being meaningful beyond the world of perception. For example, we have no doubt that a mountain exists whether we are looking at it or not. Although trivially true, this basic assumption must be critically examined. Exactly what do we mean by the statement that "A mountain exists"? Although there is a collection of implicit invariant relationships in space and time in the overall arrangement of rock strata, boulders, and stones, and molecules, etc., there is nothing that unifies the mountain in terms of those implicit relations and invariances. The mountain, then, is only implicit in a set of geometric invariances in space and time.
>
> (Davia 2006: 281)

In similar fashion, vocal tract gestures of speakers are presented to the perceptual systems of prospective listeners as discrete, discontinuous events. A linguistic utterance qua semantic synthesis (Section 14) is only implicit in such events when persons encounter them in their worlds. There is no pre-given linguistic object that is presented to the brain of the listener. The given linguistic object, following Davia, is only implicit in a discrete, non-unified set of statistical invariances. The vocal tract event as it occurs in its environment is a discontinuous set of spatiotemporal events. It is the neural activity corresponding to the active perceptual exploration of the sensorimotor contingencies of the event that unifies the multidimensional invariances of the co-articulated vocal tract event, i.e. the set of spatiotemporal symmetries that the event embodies implicitly (see Davia 2006: 281), into a single, unified dynamic – a

soliton-like travelling wave. The wave, as Davia points out, "is not a representation, but an ontological phenomenon." (2006: 281). The ontology of a vocal tract event in its environment may include organ-specific gestural contrasts, a pattern of wordings that are perceived in those contrasts, rhythmic, tonal, and other properties.

Moreover, the ontology of linguistic events may also include the implicit underlying representational topologies that such events can activate, the apperception of prior events with which similar events covaried in the person's experience, relevant covarying aspects of situations and situation conventions and their associated reflexivities, cultural norms and values, and so on. In far more complex ways than the mountain and the triangle in Davia's (2006: 2810282) examples, the behaviour of a vocal tract gestural event, its implicit discontinuous spatiotemporal invariances, give rise to emergent macroscopic behaviour that may **appear** to be intrinsic properties of the vocal tract event itself. However, vocal tract events do not constitute a continuous unified dynamic whose form and behaviour is a result of its own unified dynamics. Instead, our active perceptual exploration of such events gives rise to soliton-like travelling waves in the brain that correspond to (not represent) the event and which exhibit the same ontology as the event (Davia 2006: 282).

> … in this case, that ontology would be a direct result of its unified internal dynamics, a continuous solution to the boundary conditions imposed upon the excitable medium of the brain as a consequence of stimulus that embodies the ontology implicitly. Thus, perception makes explicit the implicit ontology of the objects and phenomena of the world around us.
>
> (Davia 2006: 282)

When we experience a vocal tract gestural event, we don't under normal circumstances perceive individual features per se, e.g. organ-specific gestures, but the dynamical relations among the many potential dimensions of contrast (the invariances and symmetries) in the event along with the fragments of other previously experienced events recalled in rich phonetic memory. This active, structuring process centrally includes awareness as a necessary component for the structuring of intelligent behaviour (Freeman 2000: 121) through the skilled application of sensorimotor knowledge. This is a necessary step if we are to explain the relationship between the phenomenological (qualitative) aspects of our conscious experience of languaging events and the quantitative dimension (the statistical invariances and symmetries) of such events. Davia's solution is that we correlate conscious experience with the "dynamic and relational properties of nonlinear travelling waves" (2006: 282) as we engage in the active perceptual exploration of these properties. The travelling ways generated by languaging agents compress (see Sections 4, 6, 12, 21, 24), informationally speaking, potentially many dimensions of implicit order deriving from very diverse and heterogeneous time-scales ranging from the cultural-historical to the here-&-now (Lemke 2000b) (See also the discussion of fractal patterns in Section 6).

The persistence in a population of languaging agents of the implicit order – the patterns, symmetries, invariances – of vocal tract and other related gestural or articulatory events are

the affordances that create the necessary conditions for the creation of soliton-like in the brains of these same agents. The robustness of the resulting semantic artefacts that are catalyzed during linguistic cognition between agents is a consequence of the way in which such artefacts and their forms of cognition make explicit the implicit order of the cultural-cognitive environment that they mediate as part of the process of linguistic catalysis. The dynamics of nonlinear, soliton-like waves, e.g. in vocal tract activity, persist in the cultural-cognitive environment by uniting energy and structure. On cultural timescales, a population evolves dynamical solutions to the perceptual problems that are posed by regular, repeated occurrences of such events. Persistent patterns of this kind mean that neural wiring in a population will ensure the emergence of a similar dynamic when prompted by the same kind of stimulus (Davia 2006: 283). This means that the brain will metabolize a new cognitive state according to the linguistic stimulus. On the basis of statistical learning (Kuhl 2007) in a population of languaging agents, infants learn to attune to the implicit order of the dynamical properties of habitually repeated first-order languaging events in their environment such that a travelling wave dynamic emerges in the brain that unities and synthesizes the diverse elements into a unified perceptual object.

Meaning is not the result of an encoding/decoding process. The emergence of a unified dynamic makes explicit the differentiating or discriminatory potential of vocal tract events along potentially many dimensions of contrast (both typological-categorial and topological-continuous). Making explicit a particular pattern of differentiations means that this pattern has the functional capacity to operate on and to evoke implicit representational topologies consisting of both individual and social dimensions of a prior history of covariance relations between perceived patterns of differentiations in vocal tract gestures and environmental events. The perception of a particular wording in a particular pattern of pico-scale dynamical properties in utterance-activity would be an example of how invariances are made explicit. Meaning is catalyzed when the nonlinear dynamic of a soliton-like wave makes explicit the invariances – individual, social, and cultural – in the relationship between a vocal tract gestural event and the environmental phenomena from different space and time-scales with which the event is synthesized to form a thermodynamically more stable semantic artefact (Section 14). Davia (2006: 285) relevantly points out: "Meaning is the experience that we have when that implicit ontology becomes explicit." Semantic artefacts themselves have the capacity to persist in a population as conventionalized typifications that constitute normative solutions to problems of both individual and social coordination and cognition. The catalyzing of a semantic artefact has the further functional capacity to catalyze flows of (self)-awareness, perception, cognition, feeling, affect, and action in individuals and groups, including virtual modes of experience (Verbrugge 1980: 94; Thibault 2011: 221). Lexicogrammatical patterns both stabilize and intermediate the pico-scale neural and bodily dynamics of individuals, on the one hand, and the macro-scale dynamics of society and culture, on the other.

The emphasis on affect, value-weightings and timing, rather than on form-based models of symbol processing as, for example, in discourse-analytical approaches, helps to advance our understanding of why we experience some behaviours as intelligent, meaningful and motivated (Section 4). If we take the full-bodied languaging agent as our locus rather than the

abstract, de-somatised forms typically focused on in the linguistic description of verbal patterns, we see more clearly that what matters is not so much how we interpret behaviour symbolically, but how flexibility enhances cognition, seen as an essential process of the living, embodied organism embedded in and coupled to its world. Organisms draw upon a history of multiple constraints on many different spatial and temporal scales to develop ways of regulating both themselves and others in order to benefit both the individual and the group. Actions and decisions make use of values both within the organism and in the environment to achieve homeostatic self-regulation (Damasio 2005). A focus on the processing of abstract symbolic forms remains an externalist perspective on the organism-environment relation (Maturana and Varela 1980, 1987).

As these researchers have pointed out, an externalist perspective tends to obfuscate living cognitive processes by defining them in terms of the functions that they perform – functions that are realized by abstract symbolic forms (Section 8). The expressive dynamics of living, feeling, moving bodies are thus seen as no more than the physical means of supporting and implementing these higher-order functions such that bodily dynamics and the cognitive and expressive powers of the body are downplayed or ignored. In this way, the specific characteristics of bodily dynamics and the central role they play in languaging behaviour and the cognitive and affective processes that these enact remain unappreciated and unexplained (Thibault 2011; Section 4). Language is a catalytic process grounded in agents' sensori-motor couplings to and perceptual explorations of aspects of their worlds. The cognitive and affective capacities that languaging behaviour installs in agents and in collectives are not one of a progressive ascent to higher-order symbolic capacities from lower-order embodied ones.

Instead, it a process of progressive differentiation and augmentation of sensori-motor coupling capacities to the exploration of and coupling to virtual cultural entities whose digitalized semantic content is based on the objective statistical properties of a population of interacting agents, but which a network of neurons can transform into a subjectively experienced semantic category that has the capacity to affect and orient the behaviour of agents towards entities that resemble those in the learning sample. Vertical coupling of the population-level dynamics to individual person-to-person acts of linguistic catalysis ensures that an increase in the intensity of a certain kind of bottom-up interaction may alter the population-level dynamics and therefore the balance of the higher-scalar (e.g. cultural) constraints acting on these lower-scalar interactions (Van Orden and Holden 2002: 95). Changes in the population dynamics and changes in the dynamics of real-time interactions between agents are intrinsically coupled. As Van Orden and Holden point out, "*Control parameters* index emergent, self-perpetuating, abstract relations among individuals and groups of individuals." (2002: 95). Virtual patterns of cultural (e.g. lexicogrammatical) constraints come into and go out of existence according to shifts in the overall balance of constraints across the vertical coupling relations involved.

25. Lexicogrammar as Virtual Cultural Pattern

Children often engage in bodily-based forms of intersubjective co-engagement that are scaffolded by mimetic capacities (Hutto 2008: 167; Zlatev 2008). Mimetic acts constitute and

enact intersubjective engagement based on joint perception-action such that a shared world is created in imagination. Infants exploit shared 'mimetic schemas' (Zlatev 2008) that serve as the basis of their coordinated engagement with each other. Thus actions performed on familiar body parts and environmental objects are a part of our everyday, familiar embodied ways of acting and being in the world. These body parts and environmental objects function as intersubjective anchors (Zlatev 2008). They afford non-arbitrary connections to a range of activities and modes of display that are the focus of joint attention. Such mimetic activities constitute a means of creating objects of shared attention even in the absence of the given event or object. Mimetic activities accordingly evoke absent objects and events on the basis of potentially shareable networks of sensori-motor associations that are held in working memory.

Infants learn in time to control bodily dynamics, including vocalizing, in ways that control perceptual input. Initially, the infant has little control over this input, sending random commands to the muscles. In time, he or she learns to correlate certain random commands with specific responses in the world, especially responses from other persons. This correlation is established on the basis of the consistencies that are established between the motor command and the perceptual input. In time, the infant can elicit desired responses in others by calling up the appropriate motor command (e.g. proto-imperatives; see Halliday (1975)). He or she can control vocal and other bodily behaviours in concert with others (and later solo) that establish a consensual domain of consistent motor-sensory relationships. These motor-sensory relationships bias perception in value-weighted ways that lead to higher-order behavioural control. Thus, control of vocal tract and other gestural activity mean that the gesture can be used to get others to fulfil one's needs and wants. The use of the gesture for higher-order control of this kind just is, from the child's point of view, the meaning of the gesture.

By the same token, the child learner also discovers that others in the same cultural milieu interpret his gestures on the basis of their beliefs, including beliefs concerning behavioural regularities or patterns that are shared and to which the participants are entrained. Moreover, the child discovers not only that such regularities – e.g. a syllabic pattern - can be harnessed to control his interactions with others, but that the use of such patterns is rewarded with positive affect from others. Initially, the child's gestures were directed to obtaining desirable input and avoiding undesirable input. The seeking of desirable input and the avoiding of undesirable input means that the behaviour is motivated – at first by what Trevarthen (1992, 1998) and Cowley (2008: 339-340) have discussed as the infant's *intrinsic motivation system*. However, the child's discovery that others use certain regular and preferred behavioural patterns in order to manage their interactions with each other in time gives the child the **cultural** motivation to align his or her own internal and behavioural dynamics with the dynamics of others in the same cultural matrix.

In this way, the child learns that such behavioural regularities based, in the first instance, on higher-order physical invariants (Section 10) serve as normative standards which people aim for and which they use as a reference point for evaluating their own and others' behaviours, in

the process learning how the syllabic or other pattern is grounded in a cultural matrix of dialogically coordinated bodily activity. It is not a matter of the syllabic or other pattern encoding a meaning, but of how the pattern is integrated with other aspects of interactional events across place and time scales. This integration occurs on the basis of local cultural beliefs and values that give rise to increasingly conventionalised patterns that serve to solve problems of coordination between individuals in their social encounters with each other. The increasing ability to reify and therefore to lift out of the flow of activity some patterns (e.g. the syllabic patterns and their combinations that the child later learns to call 'words' and, eventually, to write them) and to use these patterns to control events across an increasing diversity of space and time scales may foster the illusion that these patterns are arbitrary instantiations of symbolic tokens from a language-code (Sections 3-4). But this view loses sight of the bigger picture.

Such patterns are grounded in apperception in our neural and bodily dynamics: they have the power to induce an apperception of previously experienced (and remembered) events with which the vocal, facial, or other pattern covaried (Verbrugge 1980: 94). Statistical learning is relevant here too. Infants sensitize to the distributional patterns in a given population of (aspects of) vocalizations, gaze vectors, pointing, and the specific event types with which these covary in the situations in which they are embedded. It is in this way that infants learn how language structure attunes them to virtual experience. The emergence of predication as a semantic relation between combinations of words meant that word choice was no longer dependent on the indexical association between word and extra-linguistic experience (Harris 1991). This meant that the predication had meaning by virtue of the word co-occurrence relations in the absence of correlated real-world events (Harris 1991: 354). Grammar emerges as a system of required constraints on these co-occurrence relations.

Words contain physical information about the other words with which they are likely to co-occur. Thus, the meaning of one word could predicate about the meaning of another word in ways that were de-coupled from on-line perception and experience. In this way, operator-argument structures can predicate of and confer structure upon virtual, off-line perceptions that serve to partition some aspect of virtual experience and to orient us in a cultural world of virtual things and events that are salient and important in a given population of agents. Virtual entities of this kind are grounded in stable patterns of neuronal activity whose content is a prototype that agents extract from a population of similar entities experienced by the agent. Their virtual content is grounded in the objective statistical properties of a population -- properties that the network of neurons can activate as a category that is experienced as an off-line virtual structure whose patterns of neuronal activation are de-coupled from on-line perception. These virtual semantic entities have the functional capacity to affect how the agent behaves towards entities that approximate the prototype that has been learned. Moreover, wordings are grounded in and constrained by the functional character of the non-linguistic action systems from which they arise and which they never entirely transcend.

The experience of hearing wordings or lexicogrammatical patterns in languaging events only arises as a result of the agent's participation in the cultural activities that promote a belief in

and the capacity to adopt a 'language stance' (Cowley et al 2004; Cowley 2011) on such events as the child's languaging activity comes under more and more layers of historical-cultural constraints that afford access to virtual cultural patterns. The experience of hearing wordings in dynamical language events only arises after participation in the cultural routines that promote and reward this kind of experience in the early years of the child's life. Wordings are highly abstract, compressed patterns that are implicit in the patterns of physical energy of phonetic events at very high levels of organization and knowing. The perceptual exploration of linguistic events therefore allows for the exploration of and access to increasingly abstract, virtual patterns that are nonetheless grounded in real-time bodily dynamics of languaging behaviour and can be heard in them by those who have learned the cultural tricks for recognizing and attending to them.

Cowley (2008: 340) demonstrated that our metalinguistic practices guide us to seek out words as what we think we hear in vocal activity. For instance, mothers prompt infants to repeat the words that they (mothers) believe to have a reality that is separate from the material dynamics of first-order languaging behaviour. Infants are induced to repeat the verbal pattern *qua* word and to ignore the behavioural flow of facial, vocal and bodily expression in which this pattern is embedded. The infant is thus made aware of words as reified verbal patterns which can be at least partially abstracted from the behavioural flux and which can be repeated and transcribed into a written notation as if they were the same replicable entities from occasion to occasion, irrespective of the unrepeatable and unique dynamical characteristics of every vocalization qua historical singularity. This reification of verbal patterns leads to the idea that meanings are transmitted from one person to another and that they exist independently of the behavioural and experiential flux of coordinated inter-individual languaging activity.

However, the flow of behaviour in such activity does not encode and transmit ideas or information from a sender to a receiver by means of words. Words and wordings, as we learn to articulate and to attend to them in languaging behaviour, are value-weighted constraints on that behaviour that bias and shape action and perception around digital semantic categories; they are not coded messages that are sent and received. Rather, a person's vocalization, for instance, makes available information that perceivers in the relevant dialogical array can attend to and extract from the vocalization through their participation in the relevant language practices and associated beliefs. The caregiver (e.g. mother, father) makes the infant aware of the verbal patterns by guiding the infant's attention to those patterns rather than others in vocal dynamics and thereby helps the infant to learn how to use these patterns to achieve higher-order control of actions (Thibault 2011). From the rich array of auditory information that is potentially available in the vocalizations that the infant hears, he or she is encouraged by caregivers to attend to those higher-order acoustic invariant patterns that, according to western cultural beliefs, specify words.

In the process, the infant also learns that the metalinguistic practices of talking about and recognizing reified verbal patterns (e.g. words and wordings) as having an independent existence and the language that is so talked about are inextricably intertwined. Thus, infants are taught how to attend to and to interact with some of the information in the auditory signal

in order to detect and make use of verbal patterns while disattending to other aspects of the dynamics or relegating them to the background of their awareness. It is on the basis of this folk theory of wordings – Harris (1981) has called it the 'language myth' – that our western notion of verbal patterns (words, etc.) as freestanding entities constituting a code of form-meaning correlations is based. The problem with the code view is that language is treated as a second-order system of invariant forms or structures that can be isolated from the behavioural flux rather than a heterogeneity of interacting populations of dynamical events spread across different time scales of neural, bodily, interactional, and cultural processes.

26. Virtual Lexicogrammatical Patterns as Productive Differentiators

Utterances have intentional semantic contents that constrain the possibilities for action, perception, and cognition in particular situations. Semantic differentiators set boundary conditions on behaviour that limit the possibilities for behaviour of participants. Intentional semantic contents self-organize on cultural-historical timescales, which means that in a population of persons they persist in time relative to and are available to control processes in real-time interaction. Intentional contents thus have the functional capacity to reduce degrees of freedom in real-time interaction between persons. Semantic differentiators qua intentional contents have the capacity to effect a qualitative shift from one stable state to another in a system close to one of its critical states (Hollis et al 2009: 217); Section 9). They modify the system's phase space and limit the number of potential trajectories through that space. According to assemblage theory, an assemblage, in accordance with our understandings of nonlinear, far-from-equilibrium systems, is an emergent dynamical system that exerts causal powers over its component parts. Semantic differentiators, in selecting and delimiting some region of socially organized matter-energy flows and states, set boundaries and restrict the options for action, perception, affection, and cognition. Participants' responses self-organize as multimodal flows of events within the boundaries set by the differentiation This capacity for sustaining the trajectory of the intention in such performances emerges within a control hierarchy of vertically coupled constraints (Van Orden and Holden 2002: 95-96).

It is commonplace to claim that language conveys information. This notion of information is very vague and ill-defined and seems to rest on the idea that the amount of information conveyed by an utterance depends on the amount of linguistic content (e.g. content words) that are packed into a given sentence ('the more content words, the more information'). Let us try and think of it in different terms. I'll start with a made up (invented) example for the purposes of a little Gedankenexperiment: *You can come in now*. Obviously, this sentence could be uttered in very many different situations. At first glance, this invented example is understandably difficult to understand: what is the meaning of *you*, what is the situation in which it is spoken, what are the respective spatial perspectives and locations of speaker and listener, who speaks and to whom, when was it spoken, and so what is the meaning of *now*?

We do not know the answers to any of these questions. In order to understand the utterance, we would need to know all of these factors. However, the sentence is very interesting in a nontrivial way: it already constrains for us possible situations in which it might occur and which we can easily imagine. This fact says something fundamental about the nature of

language. We need to see language-in-action as guided by our perceptions and understandings of the events in which languaging activity is embedded. This means that we need to connect our perceptions of language structure to our perceptions of unfolding event structures, including spatial and temporal relations. I would suggest that even our invented example above shows us how language structure constrains and attunes our perceptions and understandings of the events with which language is connected.

I would argue that the 'information' in the sentence is specificational in Gibson's (1986 [1979]) sense. The utterance does not construe or encode a content that is already there. Instead, the catalytical potentiality of the utterance qua bodily event acts on and transforms the addressee's body's potential for action. The utterance does this by differentiating (not re-presenting) and focusing on that aspect of the virtual totality of the situation that the speaker wishes to transform through the interaction into an actuality. Specifically, the experiential structure of the utterance (ACTOR + PROCESS: ACTION + CIRCUMSTANCE: TIME) along with the modal operator 'can' specifies a value that serves as the desired interaction outcome in the situation. It individuates a future interactive potential in the situation. The experiential structure of the utterance sets up a form of extended, virtual action-perception that is specified by the semantic differentiators that comprise the utterance. This virtual structure modifies the content of the prior uncertainty gradient and its matter-energy flows. In so doing, it constrains the action-perception possibilities of the addressee by specifying an action format, an action oriented to the addresser's perspective ('come in'), an action role ('You' as designated performer of the action), and a time ('now'). The perceived language structure is one of a potential set of actions that seek to influence the current critical state of the addressee in the situation with reference to a concrete space of possibilities (Section 9). The virtual, as Deleuze (2004 [1968]: 263) explains, is a fully real component of the hypothetical situation in which our example might occur. The virtual, Deleuze (op. cit.) further explains, is not opposed to the real, but to the actual. This is so in two ways.

First, the utterance qua vocal tract action is a physical-material process at the same time that its real-time dynamics are constrained by virtual patterns emanating from longer, slower cultural dynamics. These dynamics give rise to the perception 'in' the utterance of wordings (lexicogrammatical patterns) as extensive crystallized formal structures. These structures are the stuff that linguistics has reified as 'language' whilst at the same time ignoring or failing to see the intensive ecological flows of matter, energy and information that are catalyzed along gradients of uncertainty and which make these crystallized formal structures possible. Secondly, these virtual-yet-real patterns set up and constrain, as noted above, potential interaction outcomes that can be actualized in the further development of the situation as different classes of multimodal events, e.g. the addressee 'comes in', that are semantically differentiated by wordings. Multimodal events are the catalyzed outcomes of linguistic catalysis. They do not necessarily correspond to action performances though they always entail changes in the physiological arousal and neural activation of the human bodies that are affected by languaging behaviour. Seen in this light, multimodal events can correspond to flows of action, perception, cognition, and feeling, seen as the actualizations of the forms of virtual experience that language has the capacity to evoke. Lexicogrammatical patterns in

language can constrain an attunement to virtual forms of experience along with an orientation to the cultural rewards that derive from aligning to these patterns. These virtual linguistic patterns are very different from the patterns specified by the immediate physical environment (Verbrugge 1980: 93-94; Thibault 2011: 220-223). Virtual linguistic patterns are like Deleuze's concept of the Idea (see also Section 23):

> A multiple ideal connection, a differential *relation*, must be actualized in diverse spatio-temporal *relationships* at the same time as its *elements* are actually incarnated in a variety of terms and forms. The idea is thus defined as a structure. A structure or an idea is a 'complex theme', an internal multiplicity –in other words, a system of multiple, non-localiseable connections between differential elements which is incarnated in real relations and actual terms.
>
> (Deleuze 2004 [1968]: 231; italics in original)

Moreover, the differential elements and their relations are not defined negatively in terms of oppositions as in the structuralist tradition of linguistics that harks back to de Saussure. These elements are, in the words of Deleuze, constitutive of a "pure positivity, in the form of a problem to which are assigned relations and points, places and functions, positions and differential thresholds which exclude all negative determination and find their source in the genetic or productive elements of affirmation" (Deleuze 2004 [1968]: 258). These genetic or productive aspects of differentiation have the functional capacity to individuate actualized solutions to the virtual problem spaces that are created in language. The virtual linguistic pattern induces a focus on the relevant or salient properties of the environment that the speaker wishes to harness at the same time that irrelevant ones are filtered out or damped down so that the desired causal chain of elements and events can be actualized and productively channelled as an assemblage. The differential elements and their relations in the linguistic utterance assign a value-weighted distribution of the important and the unimportant in the defining of the problem space, e.g., the relevant degrees of freedom, what is salient, what does not matter, who does what, when, etc. The singularities and affective capacities of the addressee's body are meshed with those of the relevant spatial environment affording locomotion and a stable spatial orientation, the location of the addresser, the virtual pattern of differentiations in the utterance, etc. in order for an embodied multimodal event to be actualized as the solution to the virtual problem space constituted by the linguistic utterance.

The utterance, by virtue of being said in the given situation, exercises a capacity to affect the intended addressee just as that person has the capacity to be affected, i.e. the utterance can bring about material bodily changes in that person's body by prompting a change in the internal organization of the person when his or her body is taken by the catalytical action of the utterance to one of that person's thresholds of self-organization. Capacities are virtual until they are actualised. A person can be taken to a particular threshold of self-organization when a set of patterns and thresholds is installed through habit and routine in that person or a population of persons by means of the catalytic capacities of utterances. The utterance is an order parameter in the sense that its grammatical structure specifies through a process of semantic differentiation a particular locus of cognitive or perceptual processing. The utterance

is a control parameter in the sense that an intentional semantic content flows from the source (the speaker) throughout the entire action trajectory so as to shape and modulate it until the trajectory reaches its conclusion (Section 16). The end result is a modification in the internal organization of the addressee, whose potential for action is transformed. It is sustained, habitual engagement with utterances and the physical and other events with which they covary that sets up through apperception the patterns and thresholds of physiological arousal and neural activation that utterances have the capacity to amplify or inhibit through their catalytical action.

In languaging behaviour, these parameters and their variations can trigger a switch in the distribution of attractors that define the state space of the situation. In this way, a given distribution of attractors is transformed into a different, topologically non-equivalent one that gives rise to transition points between phases. A situation or situation-type is a region of stability and transitions between regions defining a structured space of possibilities. A situation is, to use Deleuze's term, a virtual multiplicity (Deleuze 2004 [1968]: 230-231). Multiplicities are virtual, but not actualized. They can be actualized in a variety of ways. The tendency of the addressee in our hypothetical example to 'come in' is, in the situation, real at all times even if the person is not actually undergoing a phase transition triggered, say, by the addresser's utterance. Moreover, the phase transitions can be actualized in a variety of ways, not necessarily linguistic. The specific means of actualization of the tendency to 'come in' may vary. Languaging, as noted above (Section 9), is a culturally shaped mode of exploratory activity that changes the content of uncertainty gradients.

The fundamental question is: how to characterize the information that supports the perceiving, understanding, and acting of the individual or group in relation to its macroscopic (ecological) environment – the world in which we live. Information in the specificational sense is not about and does not stand for the things in the world, but is specific to them. Grammar can be seen as a scaling up of the agent-environment system so that the many fluctuations and variables on the micro-scale are subject to more and more constraints on the macro-scale (ecological scale). This means that language is geared to the perceiving and understanding of both the persisting and changing properties of the world. This further means that macroscopic 'grammatical' information is dependent upon and specific to the natural properties of the agent-environment system at the ecological scale.

Rather than viewing information as an abstract commodity to be exchanged, I would pose a different question: what kinds of event structures (spatial, temporal, etc) must be brought into coordination with the structure of language in order to achieve the required perception or understanding? Consider the situation in the official FAA recording of the conversations between the pilots of Lufthansa 405 and Speedbird 226 and ground controllers about the night-time UFO sighting by the pilots that occurred at 03.24 on 18th November, 1995 over Long Island, NY. The entire recording can be listened to on YouTube: http://video.google.com/videoplay?docid=-4398747992022736462

Consider the following preliminary exchange between the pilot of Lufthansa 405 and Boston Center immediately after Lufthansa has first sighted the anomalous traffic:

> **Lufthansa 405**: Ah Boston, Lufthansa 405
> **Boston Center**: Lufthansa 405 go ahead
> **Lufthansa 405**: Ah … we just ah … past traffic ah … on our left wing ah about two, three thousand feet above us ah … what kind of traffic was it?

None of the speakers (the Boston ground controller and the pilot in their respective aircraft) are co-present. They have to coordinate their spoken descriptions with their respective understandings of spatial and visual fields and their readings of specific instruments concerning altitude, heading, speed, and so on. The episode takes place at night. In the situation, the properties of the language become important determinants of the kind of socially coordinated and distributed perception and thinking that takes place about the phenomenon – the UFO – observed by the pilots in their respective aircraft (Lufthansa 405 and Speedbird 226) and discussed by the pilots and the ground controllers. Rather than seeing this unique workplace environment as one in which individuals 'use' language to produce and process information about the situation, we can, following Hutchins (1995a, 1995b), put the focus on the socially distributed nature of the thinking that takes place and the way language functions to support and shape this thinking. The language interacts directly with the properties of the event and in ways that affect the properties and capacities of the larger thinking system and its components.

Language plays a crucial orienting role in which all participants (pilots and ground controller in our example) experience virtual semantic structures that the available lexicogrammatical formats assemble and evoke as a form of extended virtual perception. Verbrugge (1980: 83) argues that language affords virtual perception and action. According to Verbrugge, language has the capacity to induce us "to experience events, to view them from fixed and moving points of observation, to move about in social and geographic environments." (Verbrugge 1980: 93). Virtual experience of this kind is distributed over the various participant-observers in the event at that same time that the experience is deictically grounded in the subjectively felt experience of each of the participants relative to their specific locations in this particular dialogical array. Thus, "on our left wing" and "two, three thousand feet above us" depend on sensori-motor knowledge for their deictic grounding in the embodied perspectives of the various observer-participants in the event at the same time that the lexicogrammatical patterns of the Lufthansa pilot's utterance evoke virtual structures of perception that are assimilated to past experience on the basis of apperception. It is in this way that the participants in an event of this kind can reciprocally attune to structure in ways that guide action, perception, cognition, and feeling (Thibault 2011: 220).

27. The Language Stance

Cowley (2011) argues that we learn to adopt a 'language stance' on the dynamics of first-order languaging. In ascribing linguistic behaviour to ourselves, to other persons, to robots, he suggests that we adopt a 'language stance' that serves to predict and account for such

behaviour in terms of our culturally shaped ideas as to what constitutes 'language'. On this view, we ascribe wordings to patterns we see and hear in the dynamics of first-order languaging behaviour. Does this mean that all languaging behaviour is a matter of ascription? Can we ascribe languaging to a talking ATM machine? Is this on a par with talking to another human? I think not and I am sure that Cowley would agree. My worry is that we risk conflating the clearly derivative languaging of the ATM with the languaging that goes on between persons. According to Cowley, nature's trick is that humans learned to recognize and to respond to second-order cultural patterns such as wordings in the dynamics of first-order languaging behaviour. First-order languaging is naturally expressive and elicits interpretations between persons. Recognizing and interpreting second-order patterns can be an important aspect of this process. Rather than explaining it as the ascribing of a language stance, I think we can also tie it to the developmental emergence of reflexive awareness and the forms of knowledge this makes available. In my view, this helps us to build a clearer explanatory link between wordings as second-order cultural constraints on first-order dynamics and the fact that we perceive these wordings as intentional semantic contents in actual occasions of interaction when we adopt the 'language stance' towards our own and others' utterance activity.

Wordings and the perception of wordings in languaging dynamics are a form of reflexive awareness – awareness of second-order 'linguistic' properties of the dynamics. First-order languaging already constitutes awareness of an ongoing interaction with one's environment on account of the sensorimotor coupling of agent to environment in the formation of an agent-environment system. First-order languaging is a mode of catalyzing (experiencing) the agent's world. First-order languaging is a first-level interactive system with which agents can explore and differentiate their worlds. Second-order language, on this view, is a second-level interactive system that takes as its environment the dynamical properties of first-order languaging and its interactions with its environment and interacts with that environment. First-order dynamics thus become an object of reflexive awareness that adds its own layer of differentiations to those made by first-order dynamics. The resulting system manifests newly emergent properties not fully reducible to first-order dynamics at the same time that second-order processes feed-off and cannot exist without the environment of first-order ones.

These include various possibilities: (1) selectively reflecting on aspects of first-order dynamics and their effects; (2) developing meta-linguistic criteria for making explicit aspects of linguistic structure and making these objects of meta-linguistic analysis and reflection; (3) reifying or objectifying aspects of linguistic organization or structure and referentially projecting these out of language onto features of the world as in standard correspondence theories of reference; (3) re-describing aspects of the dynamics as phonological features and then re-categorising these as morphosyntactic or lexicogrammatical features and patterns that have 'semantic' values assigned to them; and (4) making explicit the categories, knowledge claims and presumptions concerning the relations of agents to their worlds that provide a folk-theoretical resource for accounting for and keeping track of one's self and others across times and places with respect to a normative fabric of commitments, obligations and responsibilities

that constrain these same interactions and the folk-theoretical basis of the semantics used to account for and understand them.

Grammar is often assumed by linguistic theories to be a theory or model of the world. Instead, it is a folk-theory that provides normative, culturally shaped semantic resources for making sense of the human world. It makes available an explicit framework for talking about and rationalizing human behaviour in terms of a narrative-like structure of values and dispositions. These can include, for example, what something meant, what a term refers to, the intended meaning of a word, etc., the psychological states and dispositions of interactants, their reasons, intentions, motivations, feelings, attitudes, evaluations, and so on. Linguists, psychologists, and philosophers err when they ascribe to grammar ontological categories in either the world or in mind. The resulting linguistic reflexivity can potentially yield further levels of knowing in a hierarchy of such levels as each new level interacts with the environment of the lower levels and yields new differentiations with respect to the new level. Such differentiations would afford the possibility of knowledge of higher levels of abstraction, e.g. the formulation of the constraints that yield the meanings of sequences of clauses and sentences and their relations and meanings (Harris 1991: 336); or the formulation of the grammar and semantics of a distinct scientific sublanguage that are not necessarily sublanguages of any one natural language (Harris 1991: 290-291).

Third-order (discourse?) level interactions take as their environment not just second-order interactions. Rather, third-order interactions interact with the environment of second-order interactions interacting with first-order ones. This conception is totally different from the standard levels accounts of language, see as consisting of a number of inter-related levels or cycles of coding, e.g., phonology, graphology, lexicogrammar, and discourse. In such accounts, each level is either the means of encoding some other levels or of decoding some other level where each level is either the coded input or output to some other level. On the present account, each level constitutes processes of interactive differentiations with respect to the other levels with which it interacts, i.e. processes that catalyze the experience or meaning for the agent of these multiply constrained organizations of potentialities for agent-environment interactions.

28. Grammatical Operators as Catalytical Constraints on Socially Organized Flows of Matter and Energy

A person's vocal tract activity is an environmental event that has the functional capacity to trigger or catalyse material changes in a body or a social assemblage of bodies. A conversation between persons is a socially organized assemblage of bodies in this sense. In conversation, humans co-orient to each other and to relevant aspects of the worlds in their milieu at the same time that the conversation has a well-defined spatial structure and temporal duration (see Scheflen and Ashcraft 1976; Thibault 2008; DeLanda 2011: 16). Utterances can prompt a change in the internal structure of the individual when that body is pushed to one of its thresholds of self-organization. What we need to look at are not the separate bodies occupying different locations in Cartesian space (Section 32), but the thresholds of intensity

that can cause cognitive, affective, and other transformations of these bodies, both individually and as socially organized assemblages of bodies.

Deleuze's concept of the intensive refers to physical properties such as temperature or pressure that can form gradients containing potentially productive flows of energy. For example, a temperature gradient results from the coupling of hot and cold masses of air or water. Similarly, pressure gradients result from the coupling of high and low pressure masses. Chemical gradients result from the coupling, for example, of materials with different Ph, e.g., acid and alkaline (DeLanda 2011: 95). Affect gradients result when two bodies with different feeling states and different levels of arousal are coupled. Cognitive gradients are created when two or more bodies with different informational states are created. All of these differences represent differences of intensity that have the capacity to direct flows of matter, energy, and information. Such differences are productive because of the role they play in directing and organizing the flow of matter and energy in the assemblage and the critical thresholds, e.g. of arousal, that need to be activated for the intentional structures of different individuals to be coordinated and/or changed (Freeman 1995: 90-91). Meaning and value emerge interactively as agents orient to and assimilate themselves to the affordances of their environment (Freeman 2000: 89). The choice of one action trajectory over some other is then tipped in favour of the most economical or efficient means of dispersing the locally available energy gradient as agents adapt available resources to their immediate circumstances.

Just as ice turns to water and therefore its manner of spatial organization when it reaches a critical threshold that is supplied by an external source of energy that flows into the ice and affects its temperature (DeLanda 2011: 121), utterances can trigger or catalyze the destabilization of the patterns of activity in the entire sensory cortex, giving rise to a new spatial pattern of activity, which is nonlinear and chaotic (Freeman 1995: 66-67; Sections 14-15). This works by installing in a population of interacting individuals a set of persistent patterns and thresholds that are catalyzed by utterances on the basis of routine or habit. Habitual, repeated patterns of interaction between agents make some patterns more standardized and repetitive such that they become crystallized as social routines and culturally constrained or 'grammaticalized' procedures for making semantic structures routine. Grammaticalization is therefore a culturally constrained scaffolding of semantic structure that provides normative criteria for the interpretation of utterances. The question arises: how do patterns and thresholds of physiological arousal and neural activation based on new triggering relations with utterances get installed in an assemblage of bodies in a given population?

The crucial point here is that language is a form of value-realizing behaviour that coordinates the interactions between persons, seen as an individuating process. As Maturana and Varela (1987: 218) point out, the small-scale interpersonal transactions that bonded early hominids in small groups of community and reciprocity based on food sharing, stable sexual relations, cooperation, child rearing, and close emotional bonds opens up an enlarged possibility space of individuating transactions as social individuals explore their relations and their capacities to affect and be affected by each other. The more intensive emotional bonds and solidarity gradients that bind individuals in close-knit communities also afford enhanced possibilities

for exploring each other as individuals. Bipedalism and the enhanced possibilities this affords for distributing and sharing food among the members of the group, frontal coitus, face-to-face child-rearing, and the sharing of parenting roles between males and females all select for "a biology of cooperation and linguistic coordination of actions." (Maturana and Varela 1987: 222; Freeman 1995: chap. 6; Thibault 2004b; Savage-Rumbaugh 2011).

Interpersonal relations and interactions that afford the intensive reciprocal exploration by agents of the very many fine-grained sensorimotor discriminations and their expressive potential in vocal tract, facial, and gestural modes of co-articulation lead to the emergence of languaging agents with the capacity to narrate their selves and to keep track of the other selves with whom one is bonded in relations of reciprocity and community across place and time scales. Languaging is the individuation of positive, productive differentiations that have the functional capacity both (1) to operate on and transform social reality and (2) to individuate the persons who engage in these transactions as selves that function as valued distinctions in a domain of socially coordinated linguistic interactions. Such interactions afford the recursive exploration and narration of selves through the making of finer and finer expressive distinctions whereby selves and their stories are operated on and transformed when expressive distinctions become an object of reflection in coordinated languaging behaviour between agents. Languaging does not directly reveal or express the inner states of individuals. Insofar as it both stabilizes and intermediates between the pico-scale dynamical properties of individuals and the larger-scale social and cultural dynamics of populations of individuals (Section 24), languaging affords semantic differentiators that are decoupled from online perception-action. Decoupling means that semantic-pragmatic strategies of concealment of motive and intention can be entertained and enacted so as to further the interests of individuals in their dealings with one another. Recursively self-maintenant languaging agents have learned to exploit and adapt social and cultural norms for the invention and narration of selves that serve the strategic purpose of enacting and negotiating modalised intentional stances with others (Ross 2007). The 'language stance' (Cowley 2011; Section 27) thus serves the self-maintenance of the agent (Section 8) by enabling the agent to manipulate the reflexivities of situations and their conventions along with the interactive potential of these situations for its own strategic purposes.

Persistent patterns of vocal tract gestures, not individual behavioural units, change the individual's relationship to its environment. The selection pressure is on combinations of units in co-articulated vocal tract behaviour and populations of resources. The 'grammaticalization' of these unit-combinations is then a population-level adaptive response by way of standardization of and reduction in complexity of sensori-motor procedures that enables languaging agents to adapt to and reflect on language (Argyropoulos 2008; Section 5). Lexicogrammatical patterns are socially distributed and culturally transmitted routines that select for certain kinds of learning, cognition, and organized behaviour in the interactions between the persons who belong to the same interpersonal network. The pico-scale behavioural units of first-order languaging behaviour are measurable on timescales ranging from milliseconds to fractions of seconds (Section 4). That is, units on these time scales have the functional capacity to be salient in some way in the interactions between persons and

between persons and their environment. Because they have the capacity to be salient, they are capable of making a change in the individual's or the group's relationship to its environment. The physical, social, and cultural affordances of the environment are, typically, far more persisting than are the pico-scale dynamics of first-order languaging behaviour on any given occasion (Reed 1996a: 30).

Persistent patterns of vocal tract behaviour are 'grammaticalized' in the sense that they become routine and norm replicating. Lexicogrammatical patterns are norm replicators in this sense. Such patterns express one's belonging to a group or interpersonal network. They do not cause behavioural responses in others; instead, they act as triggers or catalysts for behavioural responses of conspecifics that are presupposed to possess complex internal organization (DeLanda 2006: 22). This internal complexity, DeLanda points out, is not explainable solely in terms of material causes such as being in possession of a nervous system of a certain kind. Linguistic catalysts imply the existence of reasons for acting, predispositions to act, and beliefs as to the existence of a particular social order or set of implicit underlying representations and associated subjective gradients (DeLanda 2002: 194). The routine nature of lexicogrammatical patterns can lead to explanations that obscure their true functionality.

For example, Langacker (1987) views grammar in terms of cognitive routines that model internal cognitive or mental processes (Section 3). Halliday (2004) views them as the outputs of an abstract systemic potential that is social in origin (Section 1). Both of these views obscure the ways in which lexicogrammatical patterns are catalysts that are not reducible to either of these views. Moreover, both views fail to show how languaging behaviour often involves the successful manipulation of causal chains of events that are not reducible to either abstract mental or cognitive faculties or the formal outputs of choices from systems in the form of 'text'. The routine and norm replicating characteristics of lexicogrammatical patterns have been honed by their application to problem solving situations over many generations (DeLanda 2011: 24). These patterns may appear to be habitual formal patterns while also being able to catalyze successful causal interactions between material entities and to direct flows of matter-energy in the social world.

Lexicogrammatical patterns are cultural replicators. They arise in a given population of interacting agents largely as a statistical consequence of the many efforts of individuals to coordinate their intentional interactions with each other on the basis of stable conventions. On the individual level, utterances are intentional acts at the same time that they are shaped by and conform to norm replicating patterns that are an unintended statistical consequence of the many efforts of individuals to coordinate their actions with each other in mutually comprehensible ways. On the other hand, utterances can catalyse beliefs corresponding to underlying presupposed networks of representations that are the result of socialization through the agencies of family, peer group, school, etc. However, linguistic catalysts act as triggers, not causes, of these representations. In a given population, a given linguistic operator can trigger or catalyze variable representations according to socialization effects that are variable and are accordingly distributed in variable ways throughout a population.

Linguistic utterances do not replicate in the biological sense – they do not replicate using a self-assembly code as does DNA. Instead, we interpret them on the basis of apperception. In this way, utterances get integrated to networks of representations that are stored in memory on the basis of prior experience. Utterances are grounded in apperception (James 1899; Bickhard 1998; Thibault 2005a, 2005b): they have the capacity to induce an apperception of previously experienced (and remembered) events with which the vocal pattern covaried (Verbrugge 1980: 94; Thibault 2011: 220-223). Linguistic utterances have the functional capacity to trigger or catalyze networks of associations that can, in turn, evoke associations with other events and thereby add to, subtract from, or otherwise modify the network of associations in ways that catalyze flows of cognition, feeling, and behaviour. Networks of associations are constituted by stable patterns of neuronal activity called distributed representations. Their content is a prototype based on a sample of a population of linguistic utterances that are detected by the individual. Their prototypical content is derived from objective statistical properties of a population (of utterances), which an assemblage of neurons can transform into a subjective experience of an apperception. Apperceived experiences are not stored in the physical connections of neurons. The only thing that is stored is a configuration of connection weights or values in a particular neuronal assemblage. A given utterance (or other stimulus) can set off the original, stable pattern of activation. The neuronal network stores a set of properties in the form of value-weighted connection strengths that give the neuronal network the capacity to activate the prototype as a subjectively experienced conceptual-semantic structure when the right stimulus occurs. Subjectively experienced and felt conceptual-semantic networks that are activated in this way can affect the individual's behaviour.

I shall now look at a brief transcription taken from a video recording of a conversation between Paul, the husband, and Dion, Paul's husband concerning the mess that Paul's pet cockatoo, a species of parrot indigenous to Australia, has scattered from his cage onto the kitchen floor. Paul has just returned home from work. The conversation takes place in the presence of Barney, the caged cockatoo, in the kitchen. The transcript of the verbal data is as follows:

> Dion: (1) yeah (as if answering for Barney) (2) what's he spilt it already? (3) look at the mess, Paul (4) you're gonna have to clean up in here (5) it's a pigsty
>
> Paul: (6) it is rather a pigsty
>
> Dion: (7) yes (8) well (9) it's your animal
>
> Paul: (to Barney): (10) you make a mess Barney (11) you make more mess than I do everyday

The following analysis is of clauses 3-5 only. Dion grounds her intentional orientation to the situation in an appeal to community norms concerning matters such as who is responsible the cleaning up of the mess caused by the pet cockatoo as well as norms of tidiness and cleanliness. In doing so, her utterance is empowered with the capacity to affect the addressee, her husband, Paul, on the basis of such community-held norms that bind its members to each

other in relations of obligation and responsibility. It is on this basis that persons have the capacity to affect others and to be affected by them through their languaging activity. In the verbal transcript and the video recording on which it is based, Dion seeks to secure a commitment from her husband that he will take responsibility for cleaning up the cockatoo's mess. Cockatoos are excellent talkers and yet it is interesting to know that Dion does not seek to hold Barney (the cockatoo) responsible, but her husband. The reason for this lies in the fact that Barney is not presupposed to possess the necessary criteria of personhood: he does not the capacity to be affected by the utterance in the way intended by Dion. Barney is but a peripheral player in this language game. In other words, Paul, but not Barney, is endowed with the normative status of a social person who has the capacity to be affected by his wife's utterance in the intended way. Utterances have the functional capacity to transform the social status of persons by, for example, holding them responsible for their actions or conferring specific expectations, commitments and obligations on them. Dion's utterance serves this function (Section 8). Such capacities are exercised in and through the utterance-activity of persons. In this way, utterances have the functional capacity to transform the socio-cognitive status of someone. When this occurs, we can say that these capacities have been actualized.

Dion's utterance assigns a value-weighted distribution of the important and the unimportant in the defining of the problem space (Section 26). The utterance specifies, from her point of view, as analytically re-constructed by me from my own observational perspective, the relevant degrees of freedom of the participants, what is salient, what does not matter, who does what, when, etc. The singularities and affective capacities of the bodies of the participants are meshed with those of the relevant aspects of the situation (e.g. the mess caused by Paul's cockatoo's scattering his seed on the floor) affording perceptual co-orientation, with the respective locations and points of view of the participants both with respect to each other and also the relevant aspects of the scene, with the virtual pattern of differentiations in the utterance, the proposed course of action, Dion's proximate intention, Paul's willingness to commit himself, the relevant community norms that are appealed to, etc. in order for an embodied multimodal event to be actualized as the solution to the virtual problem space constituted by the social-affective-cognitive assemblage which Dion's utterance helps to create.

Clause 1 is imperative mood of the subcategory [JUSSIVE: SUBJECT IMPLICIT]. More specifically, it serves the function (Section 8) of directing Paul's attention to the mess in the kitchen, as specified by the choice of Predicator *look at* and the Complement *the mess*, along with the vocative element *Paul*. The three clauses comprising Dion's utterance are analyzed as follows, drawing upon and adapting some principles in Hasan 2009 [1992]):

- ***Express Proximate Intention***: Make a claim on Paul's attention and make him aware of the mess ('look' in clause 3); get him to clean it up ('you're gonna have to clean up in here') (clause 4);

- ***Give a Reason-Explanation***: *it's a pigsty* (clause 5);

- ***Ground the utterance consisting of the three clauses in relation to a more general principle***: 'clean and tidy kitchens are healthy, desirable, etc.; messes caused by pets are not';

- ***Appeal to a Community-held Validity Claim as the ultimate grounds of the utterance***: Norm Conforming Attitude towards Community Norms of 'tidiness' and who is responsible for cleaning up messes.

Clause 4 selects declarative mood combined with a number of other lexicogrammatical selections. These features include the putting of the second person pronoun *you* in first position, the future orientation in *going to*, the modal operator of obligation *have to*, the lexical specification of the desired course of action *clean up*, as well as the locative circumstance *in here*. Significantly, clause 5 follows these two clauses. Its function may be glossed as [GIVE A REASON FOR/MOTIVATE THE DESIRED COURSE OF ACTION], i.e. the proposed action specified in clauses 3 and 4. It also functions to position both interlocutors in a particular attitudinal relationship to the referent situation. The three clauses in Dion's overall speech act function in their various ways to construct and negotiate both an interpersonal relation with Paul by attempting to engage him in a particular way as well as to create a specific, shared attitudinal orientation to the situation.

With specific reference to clause 3, *it's a pigsty*, I shall now say a few words about the experiential structure of the clause as a semantic operator on the implicit presupposed representations in the situation (again, as re-constructed by me from my observational perspective). Clause 5 can be analyzed as follows:

i. Differentiate the current locus of attention by selecting a variable in the local network topology: the pronoun 'it' exophorically refers to the generalized external situation or milieu in which the utterance occurs;
ii. In so picking out and focusing on the external surroundings (the mess caused by Paul's cockatoo), the pronoun 'it' is conventionally understood to be contextually sensitive to the local requirement that the pronoun is correctly associated with its referent and that the referent is retrievable from the context of utterance (i.e. its reference is locally resolvable or interpretable);
iii. The pronoun 'it' is selected as the local point of departure – the Theme (Halliday 2004: chap. 3) – for the development in the remainder of the clause (the Rheme) of a goal-directed action-structure;
iv. The transitivity (operator-argument) structure of the clause enables cognitive operations on the topology that partition and filter it to a smaller subset of possibilities to which attention is directed according to the specific contrast set that acts as a basis on attraction for the action-trajectory;
v. The action structure is a time-locked intentional trajectory that is semantically modulated by the speaker's intention all along its trajectory until its conclusion;
vi. The utterance induces apperceptions of representations of pigsties and related matters (see below) by setting up in its transitivity structure a relationship of attribution ('is') between 'it' and the nominal group *a pigsty*.

Dion's utterance in clause 5 is perceptually grounded in a situated context on which the utterance operates and which it transforms. Her utterance picks out the external situation that 'it' refers to as the point of departure for a message about that situation and predicates something of it. The verb Process 'is' functions to profile a temporal relation (Langacker 1987: 214-215) between two nodes in the local topology, viz. the situation referred to by 'it' and the nominal group *a pigsty*. The selection of the simple present tense establishes the speaker, Dion, as the agent who is responsible for the orienting role of the utterance. This stands in contrast to the co-orienting function of the progressive tenses (past and present), which invites both addresser and addressee to attend to the process together. Processes profile temporal relations or interconnections between things at the same time that relations are conceptually dependent on the things that are interconnected (Langacker 1987: 215). In the present example, the Process profiles a temporal relationship between two nodes in the local representational topology, i.e. the mess in the room caused by the cockatoo and evoked or apperceived virtual representations of pigsties, other messy rooms, messy pets, careless pet owners, etc. to which the state of the room is unflatteringly compared and which draw upon a stock of such representations as grounded in prior experience.

Representations are defined by their presuppositions – what is implicit and not explained and by their contrast spaces. (They define what the relevant options for the representation are). The presuppositions are the underlying representational topologies that make up the dimensions of a representational state space, while the state space is defined by a distribution of particular differentiations in that state space, a particular partition of possibilities into distinct basins of attraction. Linguistic utterances are norm-governed actions. This is so in the sense that the members of a common culture regularly associate a particular norm of socially appropriate behaviour with a given class of action. This means that social actors expect each other to make (approximately) the same association and to behave appropriately in relation to the norm that is specified. Experiential meaning can therefore be understood as normative differentiators that have the functional capacity in some social situation that is commonly understood by participants to be associated with or to induce apperceptions of underlying presupposed representations on the part of the various participants in the situation. Experiential differentiators do not encode representations of mental or social contents. Instead, they are normative patterns that have the power to evoke or to induce situationally appropriate representations that are common to all the participants in the situation.

Categories of experiential grammar (Halliday 2004 [1985]: chap. 5) are differentiators: they do not encode some pre-existing content or state of affairs. Instead, they differentiate (partition) the world (1) by inducing apperceptions of previous experienced situations and events with which the linguistic differentiator covaried in the agent's first-person experience; and (2) in doing so they indicate a usually quite delimited contrast space of possibilities in the form of relevant equivalence classes (Garfinkel 1981: 51-66) of presupposed underlying representations. Statistical learning is relevant here too. Infants learn to sensitize to the distributional patterns of vocalizations, gaze vectors, pointing, and the specific event types with which these covary in the situations in which they are embedded (Section 3). It is in this way that infants learn how language structure attunes them to virtual experience. It is in this

sense that utterances can be said to operate on a semantically structured space of implicit underlying representations which they transform by adding to them, modifying them, subtracting from them, and so on (see Bickhard 1998: 206). Apperception means that utterances are integrated with previous experience in memory in ways that get channelled in particular directions by the contrast sets of representations that the utterance evokes for participants. Because lexicogrammatical patterns are conventional resources, they have the functional capacity to evoke underlying or implicit presupposed networks of representations that are mutually accessible to all participants in the situation. Thus, utterances are conventional means of specifying implicit systems of representations to all participants on the basis of cultural learning. Participants learn how specific utterance-types indicate more information that is available in the form of the utterance itself by learning the normative conditions under which utterances are conventionally interpreted.

In this way, complex associative networks of representations get built up in memory. The associative connections that are built up through experience form integrated cognitive maps of the person's world as more and more associations are created through processes of adding to, subtracting from, and modifying the underlying representational topology along with its subjective feeling gradients. Habitual or routine experience of vocal and other stimulus events serves to trigger or catalyse apperceptions of similar experiences of such events and the environmental experiences with which they covaried in prior experience. Apperceptions set up cascades or flows of associations that increase in duration and density at the same time that they catalyse further connections as well as complex hierarchies of increasingly abstract terms in conceptual-semantic networks. A vocalization sets off a soliton-like wave in the excitable medium of the brain. The wave forges and accelerates a phase transition when a critical threshold is crossed. In this way, a catalytic pathway to each item in the network is created thereby setting off a self-replicating, auto-catalytically closed network. Every item in the entire network is accessible and retrievable through a catalytic pathway of associations that serve to evoke each other (Carpenter and Davia 2006: 1082; Sections 13-14).

In overcoming the structural constraints of the excitable medium in this way, each item in the associative network is accessible through a catalytic pathway that links all the items in an auto-catalytically closed, semantically coherent network of relations. A network of this kind both gives rise to flows of cognition and is triggered by them in a self-replicating and self-expanding network of associations that constitutes a thermodynamically more stable structure – a globally coherent assemblage -- than do any of the separate items taken singly. The overall associative network is therefore a thermodynamically more stable product than the individual component items. The catalytic process overcomes the structural constraints of the diverse items' properties and connection strengths and in so doing it facilitates the forging of a catalytic pathway that links all of the items in an overall conceptual-semantic assemblage. The soliton-like wave generated by vocal tract activity delivers the energy that is needed to change the structure of the excitable medium of the brain in order to facilitate the phase transitions required for the network to be activated (Carpenter and Davia 2006: 1082).

Gabora (2004) argues persuasively that the entire conceptual-semantic network, or "worldview" in her terms, is a cultural replicator, not the individual ideas that comprise it. Her argument is that "a gesture or idea is how a worldview reveals itself in a particular context" (Gabora 2004). This would be so because only a fragment of a person's worldview gets expressed or revealed at any given time. This would mean that the conceptual-semantic network is a cultural replicator, not the individual items that comprise it. However, phonetic gestures (the stimulus) do not express or reveal ideas. This suggests a vestige of the code view that I have rejected. Utterances, as argued above, trigger the activation of patterns of neuronal activation that have the capacity to evoke conceptual-semantic networks. Utterances are culturally stabilized differentiators that serve as indicators of future interactive potential that may or may not be successfully activated (Section 8). Successful worldviews get repeatedly activated and elaborated because they work. Therefore, they get selected and replicated through processes of cultural transmission on account of their potential future success. However, as pointed out above (this Section), phonetic gestures are cultural replicators; they replicate through processes of cultural learning and transmission and enforced social obligation in the sense that languages constitute norms of behaviour to which individual speakers of the language are expected to conform (DeLanda 2011: 56). As external resources that can trigger or catalyze the activation of conceptual-semantic networks in individuals, vocal tract gestures therefore rely upon regular, habitual, norm replicating patterns so that the required population-level entrainment effects can occur (Sections 6, 17).

Linguistic replicators are catalysts that have the functional capacity to amplify or to inhibit the self-organizing processes made possible by intense matter-energy flows. It is these flows that characterize the range of stable states and forms of organization available to a particular social system, not abstract forms variably located in the heads of individuals or seen as the outputs of systemic choices. Lexicogrammatical forms are catalysts in the sense that they are operators that can act as control switches or knobs, so to speak, to select one stable state over another. Lexicogrammatical differentiators are semantic control parameters that specify, informationally speaking, the ways in which a socially organized flow of matter and energy can be affected. They can trigger a qualitative shift from one stable state to another in a system close to one of its critical states (Hollis et al 2009: 217); Section 9). The digitalization of these semantic parameters keeps them constant and discrete. However, other parameters can be varied quantitatively, e.g. semantic grading of, say, adjectives (*very good*) or prosodic amplification and intensification that increase the amount of energy flowing into the system (a person's body or assemblage of bodies) until a critical threshold is reached, resulting in a cognitive, affective, or behavioural change. Such changes are qualitative, not quantitative. A body or a social assemblage of bodies (human and nonhuman) is qualitatively transformed from one cognitive, affective, behavioural or juridical, etc. state to some other. Moreover, linguistic catalysts, like catalysts in general, mean that "low expenditure of energy can bring about high energy transformations." (DeLanda 2002: 147). Lexicogrammatical patterns are cultural replicators that can bring about phenotypic effects, for example, when large flows of energy are triggered by a single directive emanating, for example, from the CEO of a large company.

Such was the case when, on Saturday, 29th October 2011, Alan Joyce, CEO of the Australian airline QANTAS ordered with immediate effect a worldwide grounding of the entire QANTAS fleet and the lockout of QANTAS employees in response to a series of strikes and failed negotiations between the QANTAS Board of Directors and unions representing three different groups of QANTAS employees. (http://www.smh.com.au/travel/travel-news/live-fwa-orders-qantas-dispute-terminated-20111030-1mq0o.html#comments). An airline like QANTAS is a form of social organization that exists in a constant state of metastable readiness. This is only possible because of the enormous expenditure of energy required to hold a certain form of social organization in place in such a state. In other words, the CEO's directive is transformed in a myriad of ways on many different temporal and spatial scales of organizational and individual agency into a large number of smaller flows affecting more specific assemblages that are the parts that make up the whole of QANTAS. As Foucault and DeLanda have shown, this requires the constant drill and discipline of bodies. It therefore requires the control and organization of human and other nonhuman bodies (e.g. aircraft) in social arrangements that enable the flows of matter and energy that define the given metastable states that are characteristic of a particular form of social organization such as an international airline like QANTAS (DeLanda 2002: 147). An act of linguistic catalysis is not necessarily a single flow on a single spatial or temporal scale, but can involve a multiplicity of flows on many different spatial and temporal scales. These can range from an intimate encounter between two persons to the interactions between transnational corporations, nation states, and so on.

Lexicogrammatical differentiators operate on socially organized matter and energy flows in ways that may or may not be successful. A successful differentiation is able to inhibit and/or amplify these socially organized matter-energy flows. A linguistic differentiator acts as a catalyst in the sense that it functions as an abstract operator on, in the first instance, pre-linguistic social reality; it can switch these matter-energy flows and arrangements from one stable state to another, i.e. from one attractor to another without itself changing. Second-order grammatical patterns arise in the first instance from first-order languaging dynamics as the mainly unplanned statistical consequences in a population of the constant efforts to coordinate and stabilize intentions, thinking, decision making, and so on around social norms. However, these second-order patterns act back on and constrain the first-order dynamics in various ways. In other words, second-order patterns function as order parameters that set limits on and stabilize the formal shape of first-order patterns. Once this occurs, catalytic constraints start to combine with each other to form operator-argument type dependency and other relations. These grammatical constraints have the capacity to operate on a semantically differentiated world of stable, socially organized matter-energy states and flows, forcing a metastable dynamical system of this kind from one stable attractor to another (DeLanda 2002: 291).

The increasing standardization of these efforts through processes of norm replication and grammaticalization may be seen as a way of making more explicit these intentions, thoughts, decisions, and so on, by allowing for their second-order articulation as grammatico-semantic categories. However, we should not overlook the ways in which these same processes of standardization entail a whole disciplinary apparatus comprising enforcement and

legitimation practices that are deployed and propagated throughout a population to achieve large-scale forms of coercive coordination required by governmental and bureaucratic institutions and corporations and their agencies for the economic, political, and cultural organization of whole societies.

29. Affective Capacities and Intensive Processes of Differentiation

Body movements relate persons to each other. Bodies have active capacities to affect and passive capacities to be affected by the actions of others. During an interactive encounter between persons, body parts undergo controlled transformations in the form of articulated movement patterns that have the capacity to affect others. At any given time, a body comprises capacities to act and to be acted upon to varying degrees or intensities. An interactive assemblage of persons on an occasion of talk acts on these capacities and transforms them in ways that yield emergent effects by virtue of their forming an assemblage with other bodies. Interactional synchrony (Condon and Sander 1974) resulting from the cross-modal entrainment of the bodily rhythms of two or more persons in dyadic interaction is a most basic, hard-wired capacity in this sense. These emergent effects can either mesh productively with bodies to form other connections in the assemblage as a whole, or fail to mesh productively in ways that diminish the power of the body to mesh productively with the assemblage as a whole. First-order languaging is a fluid or mobile assemblage of bodies along with aspects of their worlds (e.g. features of situations, artefacts, technologies, etc.). The micro-temporal bodily dynamics of first-order languaging behaviour tap into the core predisposition of humans to synchronize their bodily rhythms with each other. The reciprocal effects of this synchronization serve to generate the required arousal once a certain threshold of intensity is reached, leading to a spontaneous transformation or reorganization of the inter-individual dynamics that couple persons to each other in a dynamical state of interactional synchrony. Persons can entrain to each other's neural and bodily dynamics once the required threshold of arousal has been crossed and prior symmetries are broken.

On this basis, one person is able to affect another person in ways that orient the second person to the first person's locus of cognitive or perceptual processing. In this way, a shared perceptual or cognitive focus is achieved. Our local phenomenology biases us to look at the discrete elements that compose the end result. We see and hear etc. the gestures, the utterances and the achievement of shared cognitive or perceptual focus that results. These elements are, in reality, only the end products of intensive processes of differentiation, i.e. the flows and thresholds that are the necessary processes for the emergence of these end products (DeLanda 2011: 130). The meshing of bodies to form an assemblage sets up the conditions for the progressive differentiation of a topological continuum of intensive differences into a more metric, measurable and segmented material reality corresponding to our familiar everyday world and its 'segmentation' into the familiar molar objects, bodies, events, and so on.

The segmentation of utterances into words and other patterns based on the perception of wordings puts the focus on finished products. These products amount to what Deleuze recognized as 'rigid segmentarities' (1994: 222). Wordings are the second-order codifications

of what were, in the first instance, the unintended statistical consequences of very many efforts of individuals to coordinate with each other in a population or interpersonal network of interacting individuals. Historically, these emergent second-order consequences of first-order interactions became more and more codified by literacy practices. That is, they were reified as a second-order code of normative typifications that became the basis of descriptive typologies. These typologies are based on descriptions and classifications of the static properties of fixed essences that get reified by a phenomenology of abstract forms, e.g. wordings. Various forms of disciplinary procedures enforce conformity to these norms, including the agencies of symbolic control in the media and education systems.

Wordings are formal abstracta that typically form the basis of many text-based transcriptions of interactional events. This means that such events are treated as reified final products that can be segmented into abstract units of variable spatial and temporal extension. Moreover, the dynamical properties of these products and the units that comprise them are defined as low intensity equilibria, i.e. as formal instantiations of higher-order systemic properties that eliminate the intensive gradients that are the drivers of interaction (Deleuze 2004 [1968]: chap. 4). On the other hand, first-order languaging behaviour is comprised of populations of interacting pico-scale events that are defined by intensive differences that are maintained in a far-from-equilibrium condition by a continuous throughput of matter and energy throughout the lifetime of the particular interactive event. It is only in first-order interactivity that gradients are maintained; the flows and fluxes of intensive differences and critical thresholds of material dynamics define the processes of first-order languaging behaviour. The latter is irreducible to second-order products such as text-based transcriptions and the units into which they are segmented on different levels, e.g., turns, discourse moves, sentences, clauses, phrases, words.

On the other hand, first-order languaging behaviour always takes place between individuals – both human and non-human – in a population of individuals. The pico-scale dynamical properties of first-order languaging behaviour are akin to Deleuze's "molecular fluxes with thresholds or quanta" (1994: 222) that are far more fluid, dynamic, and less stratified. They are far closer to our biology and to what make us living, feeling, animate beings. As Deleuze observes of these molecular fluxes, "many things happen on this second line – becomings, micro-becomings, which don't even have the same rhythm as "our" history … " (Deleuze 1994: 222). Moreover, these micro-becomings, perhaps corresponding to micro-perceptions and micro-affects, fine sensorimotor differentiations that elude their homogenization to second-order constructs, can be harnessed in socially coordinated first-order languaging in ways that give rise to new assemblages, new forms of distributed cognitive systems, new forms of interactivity, and new forms of human intelligence.

On this view, the emphasis is less on the fixed properties of these bodies than on their capacities to affect and be affected in a field of intensive differences (Deleuze 2004 [1968]: 292-302; DeLanda 2002: 62). An interactive encounter between persons puts persons and aspects of their worlds into new functional relationships with one another. Individual persons have a variety of means for forming assemblages with others. Linguistics has tended to focus

on the intrinsic properties of language, usually seen in typological terms. An enormous amount of descriptive detail has been accumulated in this sense. However, we know less about the capacities of individuals to form an assemblage when persons engage in first-order languaging. Capacities are not the same as intrinsic properties. An utterance can be described in terms of its intrinsic properties, e.g. its phonological, lexicogrammatical, and other properties. However, its capacity to afford interaction with another person is not a matter of yet another intrinsic property (DeLanda 2002: 63). Instead, we need to investigate the productive role that utterances can play in the formation of assemblages when humans exploit distributed cognitive systems to make things happen, to solve problems, etc.

Like affordances in Gibson's theory, capacities are relational; they need to be exercised and therefore they may depend on the co-presence and co-orientation of other individuals, both human and non-human. Capacities are capacities to affect and to be affected. In this sense, they are symmetric. Someone's utterance directed my way affords an opportunity to engage with the other person just as I may say something to him or her thereby affecting the other person and engaging his interest or attention. Languaging behaviour is the "glue and grout of social co-ordination" (Steffensen, Thibault, and Cowley 2010: 208) in such encounters because it articulates heterogeneous elements (e.g. persons, artefacts, tools, technologies, etc.) as social-cognitive-affective assemblages without in any way sacrificing the heterogeneity and autonomy of the component parts, which retain their capacity to be detached from any given assemblage and attached to others. When such a system arises, the capacities of the different components interact with each so that the new whole has emergent properties. By the same token, the new whole – the assemblage -- constrains and enables its components, damping or inhibiting some capacities, amplifying and extending others, as well as adding still others.

30. Mutual Attunement, Meaning, and Entrainment to Population-level Cultural Dynamics

In co-ordinated first order languaging between persons, individuals attune to and track in space and time pico-scale bodily events of each other that take place on timescales of the order of milliseconds to fractions of seconds. Individuals adjust their sense organs on these very rapid timescales to explore, to track and to modify the sensory input obtained from diverse sources, e.g. vocalizing, gesturing, head movements, eye gaze, facial expressions, bodily orientation, and so on, in a changing multimodal array that is integrated over time and located in space as a multimodal interactive event. Brains attune to and feed off bodily events on very fast timescales (Freeman 2000: 105). This is an active and intentional process of adaptive exploration of these events. It is a commonplace of many linguistic and semiotic theories to say that meanings exist between persons, not 'in' them, although these same theories are not very forthcoming with causal explanations of the mechanisms that make this possible. No shared physical substrate is shown to exist that would make this possible.

Generally speaking, reified generalities like abstract social or cognitive codes are postulated as the means whereby individuals semiotically mediate their relations with their worlds. We have already seen that vocal tract gestures and the stimulus information that they project into

the environment do not contain and transmit meanings. The subjective experience of meaning can be activated in the brains of individual persons on the basis of their unique experience. Individual persons have the capacity to affect and to be affected by each other. This includes the capacity to induce meanings in other persons. Assemblages of persons draw on biological and cultural resources to co-ordinate their meanings and activities in ways that can lead to newly emergent structures and possibilities for thinking, feeling, and acting in the assemblage. This is how a distributed cognitive-affective system works. Both bodily activities such as vocalizations and other bodily movements and external resources enable persons to be coupled to each other and to aspects of their external environments.

According to both the motor theory of speech perception (Liberman and Mattingly 1985; McNeilage et al 1967) and the gestural theory of co-articulation (Browman and Goldstein 1992, 1995), listeners entrain to the vocal tract gestures of speakers and reconstruct them in their brains. Fowler (1986, 2010) uses Gibson's (1986 [1979]) theory of event perception to show that speech sounds provide information about the vocal tract events that cause the sounds and that speaker's attune to each other's vocal tract gestures. The listener thus reconstructs and attunes to the speaker's gestures. This has clear emotional and affiliative consequences for the kinds of social bonding that occurs when social-affective-cognitive assemblages of individuals are formed (Freeman 1995: 123-134). This includes in my view non-cognitive emotional or affiliative language – what Malinowski called "phatic communion". Phatic communion "serves to establish bonds of personal union between people brought together by the mere need of companionship and does not serve any purpose of communicating ideas." (Malinowski 1936: 314-316). Attunement also facilitates entrainment to the semantic differentiations in the human forebrain. The perception of the speaker's vocal tract gestures provides the stimulation that can recreate previous patterns of activation that are stored as a configuration of connection strengths in a neural network. The stimulus does not contain a message that is decoded by the brain. Freeman (1995: 66) shows that the sensory cortices at the interface between brains and the external world work very differently from the encoding/decoding model.

First, the stimulus is transduced by the receptor layer into a pattern of action potentials and "then into the cerebral cortex, through the thalamus to cortex" (Freeman 1995: 66). At this point, the stable pattern destabilizes the entire sensory cortex so that the previous state, expressed in a spatial pattern of activity, is now expressed in a different spatial pattern, which is nonlinear and chaotic. Nonlinear and chaotic patterns, Freeman explains, create novel patterns (Freeman 1995: 67). Freeman (1995: 66) emphasises that this new pattern is "triggered, not selected, by the stimulus, and it is determined by prior experience with this class of stimulus." (1995: 66). The stimulus – the vocal tract gesture and the auditory stimulus information it causes to be propagated through the medium of the surrounding air – is not the encoding of a pre-existing message that is transmitted to the listener, who in turn decodes it. Rather, the new pattern that is triggered expresses the "nature of the class and the meaning for the subject" (1995: 66). That is, the stimulus places the cortex in one of its basins of attraction. These are grounded in prior experience and work on the basis of apperception (Section 9).

The sensory data in the form of the auditory stimulus information provide the listener with information about the speaker's vocal tract gestures (Section 15). In entraining to these patterns and in reconstructing the other person's gestures in the speaker's brain, this data triggers the structuring of the neuroactivity that leads to the activation of socially distributed classes of semantic differentiators in a population of languaging agents entrained to the population-level dynamics. The resulting semantic differentiators are a construction of the forebrain, not the decoding of a pre-existing message that was carried by the stimulus from a 'sender' to a 'receiver'. The perception by the listener of the speaker's gesture and its reconstruction in the listener's brain thus affords the recreation of the stable semantic differentiators that have been sampled from a population of similar patterns which the individual has detected on the basis of previous contextualized encounters with these gestures.

31. First-order Languaging Dynamics: An Analysis

First-order languaging dynamics and the reciprocities they create promote habits and routines that builds upon and extends the process which Stern (1997, 1998, 2004) called "moving on", leading to the creation of a succession of now-moments (Section 14). Participants in languaging behaviour move along together, they enact and create a series of now-moments that have their origins in the narrative processes generated in core consciousness by the proto-self (Section 11). These brief narrative units lasting no more than a few seconds give rise to intersubjective moments of affect-charged "meeting" which "increase the ability of the nervous system to intensify and co-regulate with someone else's activity, and they occur only when the participants' nervous systems have been able to engage in mutual adjustment and self-regulation -- the condition that Stern refers to as moving along." (Hart 2011 [2006]: 27). These now-moments reciprocally bind the neural and bodily dynamics of dialogically coordinated participants in states of intersubjective entanglement (Section 32), resulting in mutual recognition and shared experience. Moments of dialogically coordinated meeting intensify the flow of affect through the entangled dynamics of dialogically coordinated agents and "promotes the nervous system's capacity for self-regulation and attention control" (Hart 2011: 27), for co-orientation, modification of feeling states, extended perception, and so on. Moments of meeting, as defined by Stern and Hart, are, to quote Hart, "a dyadic expansion of the consciousness." (Hart 2011 [2006]: 27). The intersubjective processes of moving on and the resulting now-moments described by Stern and Hart are, in infant development, the precursors of conversational dialogue. Dialogue can thus be seen as successive loci or pulses of intersubjectively coordinated and deictically grounded experience that loop between participating agents in dialogue and their environment(s). Dialogue is a succession of now-moments experienced in this intersubjective space-time. These now-moments are metricised rhythmic pulses or intervals of intersubjectively coordinated and deictically grounded experience that interactants co-orient to. Seen in this light, second-order lexicogrammatical resources are forms of cultural scaffolding that enable agents to solve problems of coordination as they attempt to mutually explore each other and move along together in dialogue.

In the following paragraphs, I analyze a brief example to show first-order languaging

dynamics in action. Using techniques based on a combination of both multimodal interaction analysis and Praat software to analyze speech sounds (Thibault 2011: 223-232), I focus in detail on six micro-temporal or pico-scale phases, which are presented in Figures 4 to 9 below. The episode is transcribed from a television documentary concerning the life of a family in Sydney (Australia). The two interactants, Laurie and Noeleen, who are husband and wife, respectively, argue about Laurie's inability to make any of the pens he is trying to write with work. The somewhat energetic exchange takes place around a bar in the kitchen-dining room area of their home. The analyzed material is no more than a brief example from the longer episode in which it occurred.

Pico-phase 1

This micro-phase consists of two pulses of energy in which N's vocalizations and body movement are synchronized. In this micro-phase she seeks to secure L's attention and to change L's perception of one of the pens that he couldn't get to work in an earlier phase of the interaction (not analyzed). In the first pulse, the utterance 'yeah' is synchronized with N. turning her attention back to her husband from the adjacent pen rack (on her left) from which she has just taken one of the pens that L. couldn't get to work. She shows the pen, which she is holding in her right hand, to L. The pitch of her vocalization 'yeah' is rising-falling-rising, starting on 1961 Hz, reaching a peak of 2687 Hz, falling to 2015 Hz before finally rising to 2892 Hz. The rising tone anticipates (Section 4) both her own next action at the same time that it anticipates the re-orientation of L. to the pen in her hand. The second pulse synchronizes the action of her briefly writing with the pen on a piece of paper lying in front of her on the bar top at the same time that she says 'look'. The pitch melody of this utterance is in sharp contrast to 'yeah'. This utterance initiates on a similar pitch range to 'yeah', i.e. 2037 Hz, rises to a peak of 3315 Hz, only to fall sharply to around 1246 Hz. The concluding fall signals closure, not anticipation: she shows her husband that the pen really does work.

N: yeah look
N: turns towards L. holds pen in front of him on 'yeah', leans across bar towards him, starts writing on 'look'

N:	jæ	lʊk

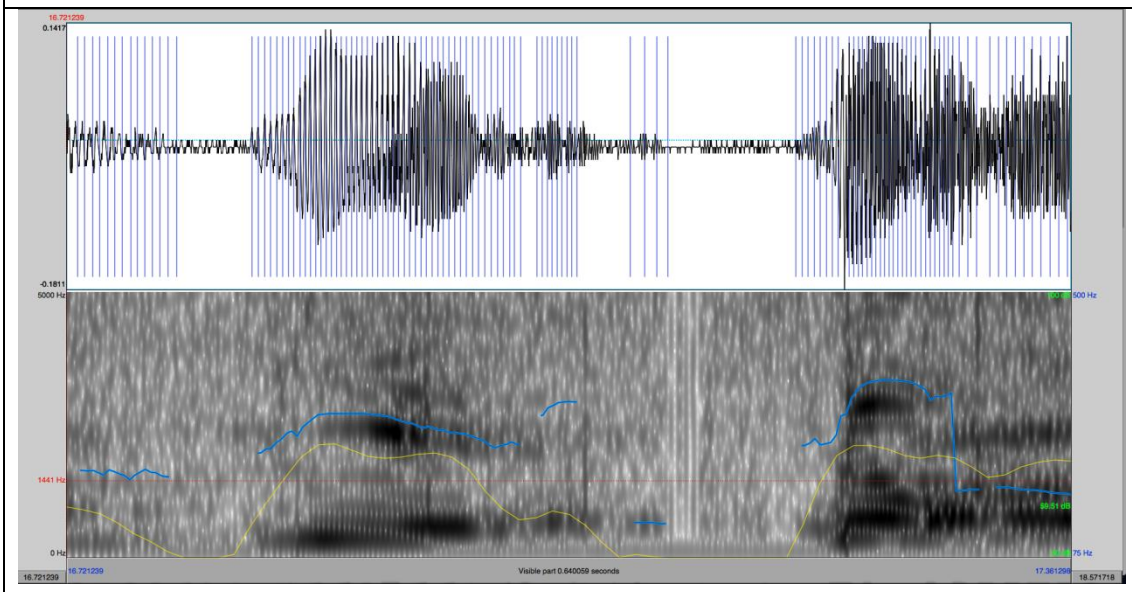


Figure 4: Pico-phase 1

Pico-phase 2

L's response as shown in Figure 5 consists of the utterance 'one of those I picked up first' during which N. holds the pen in her hand in front of L. The pitch ranges from an initial 1214 Hz to a final 109.1 Hz. Again, the falling tone signals closure. L is evidently not open to N's demonstration and persists with his own view of the matter.

L: one of those I picked up first

L:		wən	əv	ðoz		aj	pɪkt	əp	fərst
N:								ðɪs	ɪz

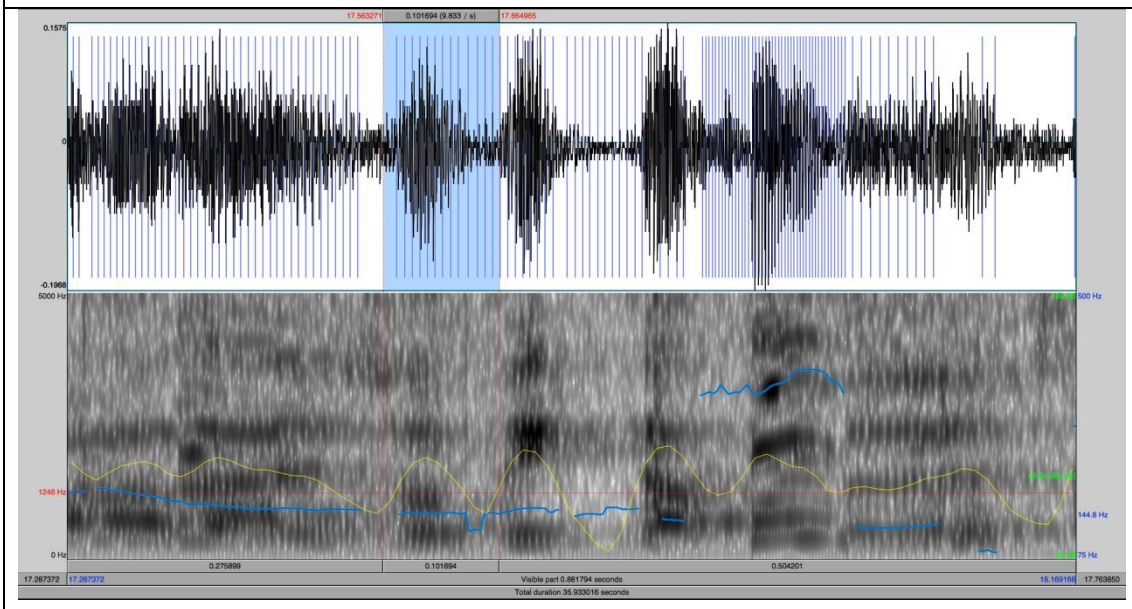


Figure 5: Pico-phase 2

Pico-phase 3

The demonstrative pronoun *this* in N's utterance grounds the pen she is holding in her right hand to show to L. in deictically felt bodily experience. This can be explained as follows. The English demonstrative *this* consists of three sub-morphemic markers: /TH/ + /I/ + /S/ (Bottineau 2007: 55). Bottineau (op. cit.) shows that, in demonstrative *this*, the phono-semantic operator /I/ assimilates past memory (/TH-/) to the item identified in the present (/S/). These three items are sub-morphemic markers or phono-semantic operators that depend on sensori-motor knowledge. The observer's relation to an articulatory event such as /TH/ is mediated by patterns of sensorimotor dependence, as discussed by Noë (2004) (Section 13).

In the case, of demonstrative *this*, the sensori-motor knowledge constitutes in experience the representation of the concept /IDENTIFICATION OF INDICATED OBJECT WHICH YOU NOW REMEMBER/. Thus, the speaker, N, invites the listener, L, to locate in the physical world the object, i.e. the pen N is holding and referring to. Moreover, this object is presumed to have been already referred to, perceived, or otherwise known, in the situation and therefore to be held in concurrent working memory. The nominal group *the one you had*, which is coupled to the pen N is holding in her hand to show to L, accordingly assimilates the present act of showing to the past memory of the object. In this way, deictic *this* anchors the situation in felt bodily experience and in the spatio-temporal frames of reference connected to this. Figure 6, Frame 1 shows how N's utterance, her right arm-hand + held pen, upper body orientation, and gaze constitute a vector of interest and attention that is directed to her

husband, L. as she shows him the pen and connects this present perception to his immediately prior experience of it.

L's utterance cannot be separated from this whole-body orientation in space and time and concomitant whole-body sensing, but is anchored in it and is inextricably a part of it. In this way, the two speakers establish and sustain a co-orientation frame that lasts for approximately 0.978 s during which time interval the vocalization is synchronised with the other body movements described above before the shift in orientation indicated in Figure 6, Frame 2 occurs. Figure 6, Frame 2 shows a shift in bodily orientation as N directs her gaze downwards and begins writing with the pen on a sheet of paper resting on the kitchen bench that lies between the two speakers while L continues writing on his own sheet of paper.

N's utterance *this is the one you had* partially overlaps with L's prior utterance. Specifically, *th-* begins slightly before *fi-*: *th-* begins at 17.863 s and *fi-* starts at 17.979 s. The entire unit *this is* starts at 17.863 s, begins the overlay with *first* at 17.971 s, with *first* concluding at 18.165 s, and overlapping and concluding in synchronization with the second syllable of *this*. These moments of overlap are pico-scale temporal events, having, in the present case, a duration of 0.186 s. The pitch of N's utterance starts on 3055 Hz, peaks at 3531 Hz and concludes on 293.2 Hz on *had*. L's overlapping *first* is, of course, the conclusion of the utterance analyzed in Figure 5. During the overlap with the start of N's utterance, the pitch range varies very little – 563.9 Hz to 650.6 Hz. The overlap, the synchronization of the rhythms of the two speakers and the rapid, excited tempo of their speaking demonstrate entrainment effects as they track, explore and modulate each other in micro-time. On initiating *this*, N ceases writing (Figure 5) and again holds the pen in front of L, only to resume writing on conclusion of *had* at the end of her utterance. Throughout the duration of this utterance, N's gaze is directed to L, before then being directed towards the writing task on saying *had*. L's gaze is similarly directed to N in the same time interval. Again, this micro-phase shows how language works synergistically with perception to orient to the environment as a whole-body process in order to adjust to and explore the environment, including environments that are not available to online perception.

N: (this is) the one you had

N. shows pen in her hand to L
N. gaze to L., which is reciprocated very briefly then L.'s gaze to rack on 'so'
N begins writing on paper; L continues writing on his own sheet of paper

N: ðə wən ju hæd

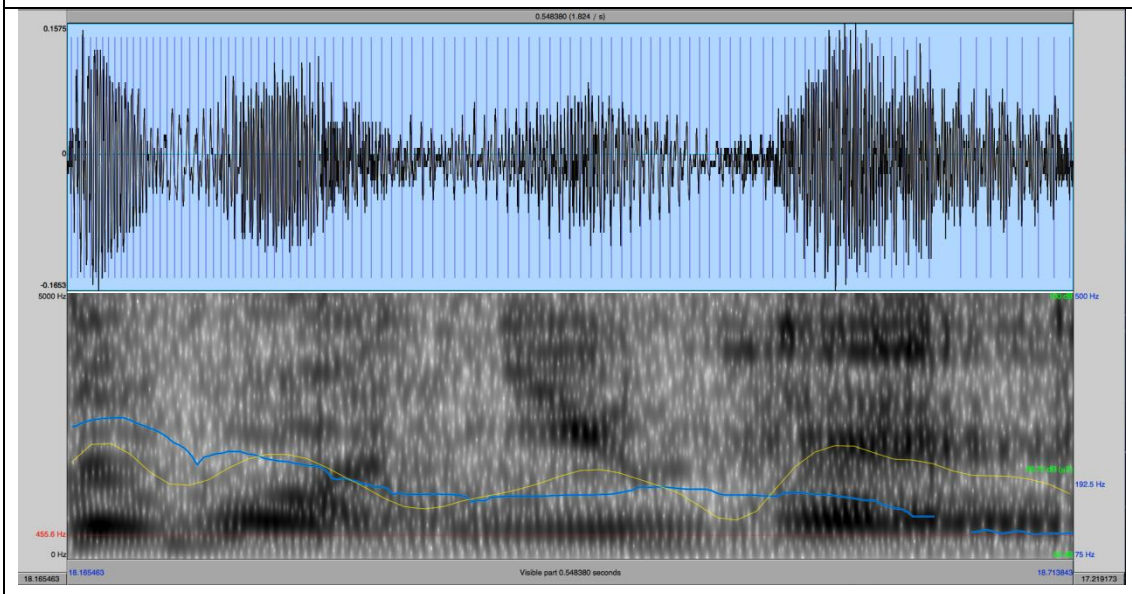


Figure 6: Pico-phase 3

Pico-phase 4

In this micro-phase, L's utterance *so I picked up one of those first* synchronizes with the following body movement. In the initial phase of the utterance, L raises his left arm-hand, which is holding a pen in order briefly to show it to N. His hand-arm performs a pronounced downward beat gesture that is co-temporal with *first*, which also receives a great deal of emphatic stress. During this time interval, N continues writing on the paper; her gaze is oriented to that task. When L utters *first* this coincides with N's uttering of *right*, which is synchronized with her head nod. L's gaze is fixed on N throughout the entire utterance. N's body articulates two principal orientations simultaneously. The primary one is expressed by her overall body position, her body posture, her gaze vector, and the position of her hands-arms – all of these elements are co-synchronized and oriented to indicate her orientation to the writing task. Without modifying this orientation, N's uttering of *right* and her co-temporal head nod as L utters *first* indicates a second, nested orientation, which is to what her husband is saying.

The pitch range of L's utterance is from an initial 1257 Hz to a closing pitch of 1420 Hz. The rise, as before, has an anticipatory function. In this case, L signals his intention to continue to hold the floor as he continues to develop his utterance (Figures 8 and 9). The brief overlap of L's *first* with N's *right* lasts approximately 0.170 ms and is characterized by a noticeable fall in pitch from 2676 Hz to 1809 Hz. N's non-committal tone and the falling pitch signal her intent to allow L to continue. Again, the overlap signals the ways in which the dynamics of languaging behaviour between persons reciprocally bind persons in states of intersubjective

entanglement as they explore each other's dynamics on these rapid pico-scales and in doing so co-regulate each other's attention, perception, involvement, and awareness.

L: so I picked up one of those first
N: right + head nod
L. raises left- hand-arm holding pen and holds it in front of N. ... then sharp downbeat on 'first'
L. gazes at writing task she is performing

L: so aj pɪkt əp wən əv ðoz fərst
N: rajt

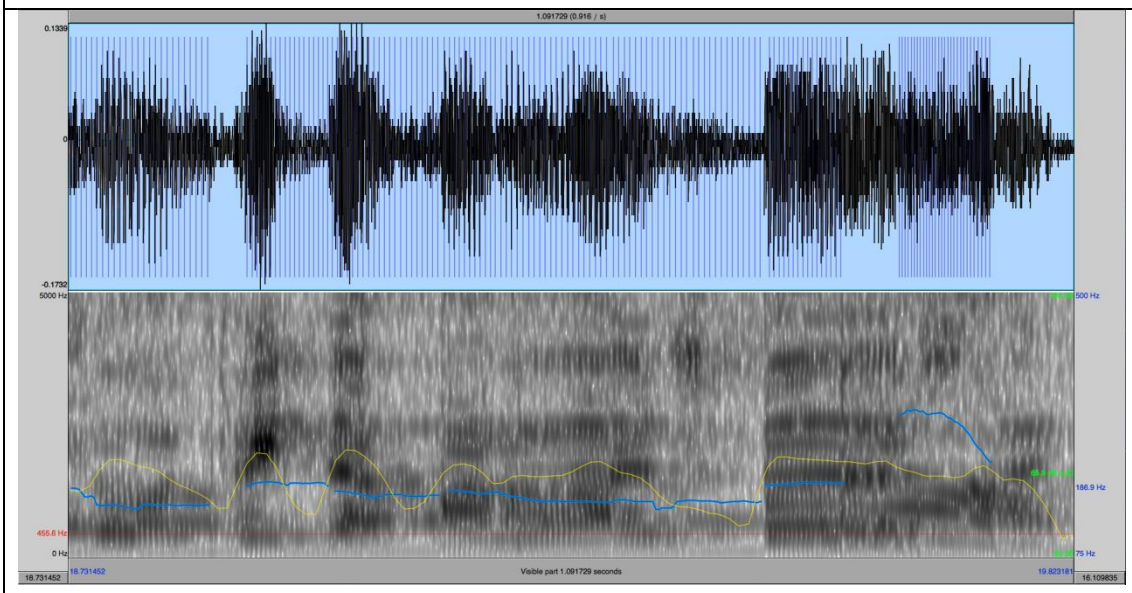


Figure 7: Pico-phase 4

Pico-phase 5

L's utterance is characterized by a strong beat gesture, which coincides with *third*, which also receives emphatic stress. His gaze is directed at N during his emphatic, rapid delivery of this utterance. L's attention remains focused on the writing task, thereby sustaining the same postural orientation described in relation to Figure 7 above. On conclusion of his utterance, L begins a shift in orientation towards to the pen rack on his right.

L: I said that's the third one
L. arm downbeat, culminating on 'third' + gaze held on L. on 'third' before being directed to pen rack on his right;
L: gaze re-engages N.

L: aj sɛd ðæts ðə θərd wən

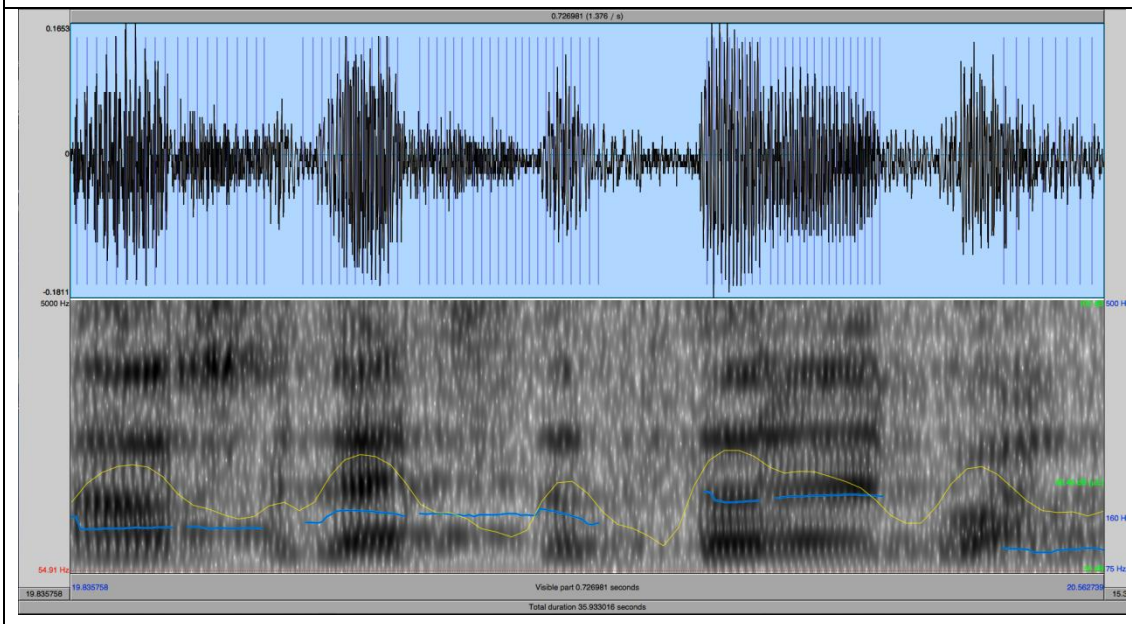


Figure 8: Pico-phase 5

Pico-phase 6

During this utterance, L shifts orientation to the pen rack as his right arm, head, gaze and upper body orient to the pen rack whilst he reaches towards it to replace the pen he was holding. N's head and gaze turn left to track this shift in L's orientation. During the latter part of his utterance (*hear me say that*), his head turns briefly back to L in order to re-establish more direct interpersonal contact with her in a brief surge in intensity that this head movement captures. Also noteworthy is the higher energy level and therefore the increased salience of *say* along with the surge in pitch that L's uttering of *say* evidences. The initial pitch of this utterance is around 856.4 and at the end drops to 444 Hz after reaching a high of 1311 Hz at the beginning of *say*.

L: didn't you hear me say that?
L. puts pen he had taken back on rack
L. gaze to rack then back to paper he is writing on; N's gaze tracks L's action

L: dɪdənt ju hɪr mi se ðæt

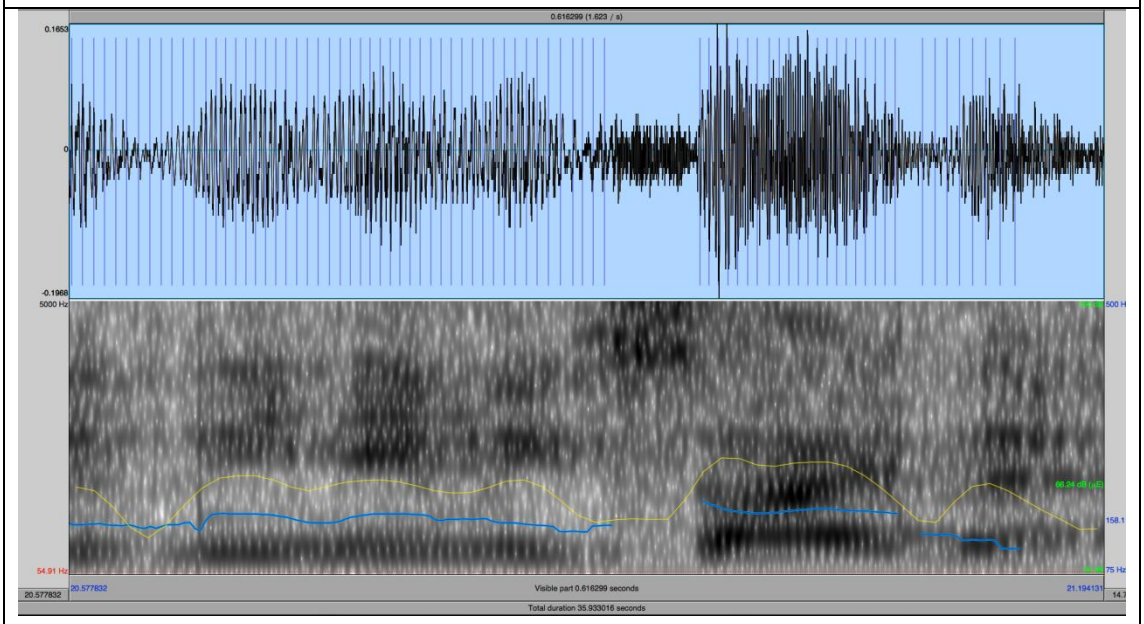


Figure 9: Pico-phase 6

Some Further Implications of the Analysis

We do things with our voices and other bodily movements that have the capacity to affect others. In minding and en-minding others with our voices we perform actions. The voice gives us access to others because we can change them with our voice and its many possibilities for modulation and differentiation. In the code view, we think of the voice as an encoder of ideas or meanings that we only have indirect access to. However, vocalizing is, above all, action that can move and change others. The direct effects upon the world that our voices can exercise are thus able to influence the indirect cognitive access we have that is dependent on neural coding. Voices extend our agency and our worlds because their direct, material agency brings about changes in others that in turn have the capacity to change our understandings of the world. It is this direct contact (e.g., the voice, the hands, and so on) that gives us our grip on the world and renders our ideas and understandings veridical. It is through this direct contact with and acting on the world that our more abstract ideas and understandings are enhanced.

First-order languaging as in vocalizing and related behaviours is a process of orienting both addressee and addresser. Both interpreting agents orient to and experience the integration of bodily activities and lexicogrammatical formats and patterns which are assembled and harnessed as a dynamic time-bound trajectory. The resulting meanings are distributed between all of the agents (addressers, addressees, observers) and vary according to the feeling states of the individual's body, the perspectives and psychological situations of the

participants, their personal histories and concomitant experiences, and their perceptions and understandings of the circumstances in which the languaging behaviour takes place. Meaning is an emergent and variable process which does not correspond to an external reality, but is constantly negotiable and able to be adjusted in order that some kind of working agreement, however contingent and context-bound, about what is meant, can be arrived at. Languaging is a means of attending to experience in concerted fashion with either others or with the self in ways which afford ways of responding to and transforming the situation, including felt bodily responses to it, by providing the possibility for increased specification of the experience to be articulated. These processes take place across many different bodily and extra-bodily scales.

For example, the dynamics of the speaker's voice may provoke in the listener physiological responses (e.g., tensed muscles) which are felt by the addressee as feeling states of the body that provide implicit evaluations of and potential responses to the situation. Voice prosodies such as perceived patterns of rhythm, tempo or pacing, intonation (tone) and degree of loudness afford ways of synchronizing with the rhythms of the speaker and in achieving arousal, affective attunement, interpersonal harmony or convergence, and so on. Inter-individual patterning of this kind thus constitutes a mutual form of body-sensing and body-attunement that can bring about a qualitative shift in the experiences of the interlocutors. Prosodies also interact with and are integrated to lexicogrammatical units in various ways to signal salience, contrast, modal stances, and so on. They can, therefore, bring about a qualitative shift in participants' perceptions and understandings of lexicogrammatical units along with the referent situations that they construe. Phono-semantic operators (morphemic sub-markers) (Bottineau 2007:55) enact forms of deictic sensing and orientation that anchor the situation in felt bodily experience and the reference points attached to that. Wordings couple deictically grounded felt experience to the historically accumulated cultural experience of a population and the networks of associations, norms, and values (individual, social, and cultural) that wordings evoke by apperceptively matching current experience with previous experience, including felt bodily experience.

Meaning is not the starting point of this process, but is its emergent and always distributed achievement. Languaging is a form of exploratory activity that is oriented to the achieving of meanings, to making them emerge in the temporal flow of the activity itself. Languaging is a way of completing what was partially implicit, in the process transforming the nature of the situation and the values and understandings of its participants, though by no means in identical ways for all the persons involved. Emergent meanings do not reflect the situation, but dynamically transform it in the process of activating concerted thinking between participants. Meanings are crystallized in this process rather than being the preformed and static inputs to operations of encoding and decoding.

32. The Quantum Coherence of Whole-body Sense-making and the Ontology of Entanglement

Through first-order languaging we become entangled with and interwoven with each other's bodies and bodily feelings. The study of conversation has for the most part focused on verbal patterns and other abstracta such as discourse moves, conversational turns, and so on. Oddly,

bodies and their articulatory capacities are seen as the encoders and transmitters of these abstracta. Linguistic and discourse-analytical approaches talk in terms of addressers/addressees, senders/receivers, or at best in terms of speakers/listeners. In these conceptions, addressers/addressees, senders/receivers, and so on, are local points in space-time, reflecting the metaphysical premises of classical physics. These approaches remove the life from languaging.

With the exception of the latter pair of terms, none of these terms accounts for the ways in which languaging is not simply or only heard and seen; it is also *felt*. First-order languaging is a felt whole-body experience. Languaging is underpinned by the always dialogical, enkinaesthetic, entanglements of agents with each other (Stuart and Thibault Forthcoming). Bodily feelings, neurohormonal flows, blood pressure, galvanic skin response, physiological arousal and response, bodily displays, and action tendencies, to varying degrees of intensity, are synchronized as organized affective and bodily responses of whole living beings to each other. Such responses are reciprocal though not usually or necessarily symmetric. Your display of anger towards me does not mean ipso facto that I too will feel anger towards you. I may do so, but I may also instead feel shame, embarrassment, humiliation, etc.

Experimental psychologists have investigated the link between language comprehension and activation of the neural substrates of action and emotion (Gallese, et al., 1996; Gallese, Keysers & Rizzolatti, 2004; and Rizzolatti, et al., 1996). When it comes to language, these approaches have focused on the same kinds of abstracta referred to above, viz. lexicogrammatical patterns. In one such experiment conducted by Mouilso et al (2007), participants were asked to read "angry, sad or neutral sentences to decide if the sentences made sense" (Mouilso et al 2007: 1326): "Sense judgments are made by moving a lever (see Figure 4) so that the time taken to move the lever can be used as a behavioral measure of the level of bodily activation associated with the emotion produced by the sentence." (Mouilso et al 2007: 1326) These approaches may tell us something about the ways in which second-order language has the potential to constrain and trigger enkinaesthetic responses. Thus, the cultural values and connotations that are evoked by particular lexicogrammatical patterns and their associated value-stances can trigger bodily feelings (Thibault 2005a: 288-299).

However, my concern in this paper lies in articulating an alternative to mechanistic views of talk that abstract away from body dynamics. Such views have predominated in conversation analysis and in discourse analysis. In these approaches, discourse-analytical or conversation analytical units such as 'moves' and 'turns' are said to be combined into sequences on the basis of rule-governed regularities that remove language from felt experience and personal history (Section 1). From the present perspective, these approaches fail to see that languaging is grounded in dialogically coordinated pico scale bodily events that are not merely heard and seen. They are also felt and sensed as whole-body experiences and feeling states (Cowley 2006). The fact that languaging is intimately and intrinsically connected to bodily feelings and sensations and the evaluations of situations that persons make about their own and other's feelings is of course entirely unaccounted for in these approaches. However, languaging just *is* whole-body sense-making. I argue that this felt dimension is absolutely fundamental to our

experiencing and interpretation of first-order languaging (and much more). This view is supported by the experimental research of Gick and Derrick (2009), which showed that the neural processing of speech integrates naturalistic event-relevant tactile information during auditory speech perception.

Schrödinger's (1983 [1935]) used the term 'entanglement' to describe how two physical systems that enter into physical interaction with each other can never be the same again, even after they have separated, owing to the reciprocal influence of the one on the other (Schrödinger 1935: 555). The two systems thus become, according to Schrödinger, "entangled". In making this point, I am drawing on one of the fundamental lessons of contemporary physics to the effect that the coupling of microscopic quantum events to macroscopic ones means that the microscopic will affect the macroscopic. This is very different from saying that speakers and listeners, in their languaging, are atomistic monads (individuals) that are reducible to more and more micro states and properties on ever small scales, or that they encode, 'transmit' and 'decode' information or meanings to and from each other.

In the case of human interaction, the earlier work of Condon and Ogston (1966) and Condon (1970); see also Condon and Sander 1974), using high speed sound film techniques to analyse the flux of interactional behaviour between persons, showed that the assumption of discrete behavioural units that exist or occur "within and between the behaviour of individuals" (Condon and Ogston 1966: 338) does not hold. Instead, the very material nature of interaction dynamics transformed this into a recognition of what these authors called ""patterns of change" within ongoing behavior" (1966: 338). To quote Condon and Ogston:

> Intensive analysis revealed harmonious or synchronous organization of change between body motion and speech in both intra-individual and interactional behavior. Thus the body of the speaker dances in time with his speech. Further, the body of the listener dances in rhythm with that of the speaker!
>
> (Condon and Ogston 1966: 338)

The theory of mirror neurons is pertinent here. Mirror neurons connect observation and behaviour. Italian researchers Gallese, Rizzolati and Arbib, Fogassi, Fadiga, and others discovered mirror neurons, at the University of Parma, Italy in the 1990's. These researchers discovered that mirror neurons are activated both when someone observes another individual carrying out an action perceived to be intentional and when the same person carries out the same action. Mirror neurons are, in part, the neurological substrate for understanding and predicting another's observed actions. They suggest a relationship between observation of the behaviour of others and our understanding of it, including the ability to imitate others. In suggesting that the other is virtually present in our brains, we are able to relate to and empathise with the embodied perspectives of others as well as understand and anticipate others' intentions. Infants quickly pick up the vocal, facial, and other body actions associated with languaging. In observing and attending to the languaging behaviour of others, infants' neuronal activity builds connections between the motor cortex, perceived positions and

movements of the articulators (lips, tongue, etc.) and the pick up through observation of auditory, visual, kinesic, and other information of the other person's vocal tract, facial and other movements.

First-order languaging is a form of concerted activity between persons. It promotes the forms of observation and attendant emotional involvement or empathetic attachment that lead to the activation of the observer's premotor system in the frontal cortex. In this way, the patterns of neuronal activation that occur prime or ready the observer for the action though without necessarily entailing that the observer carries out the observed action. Mirror neurons allow us to co-participate in, or to co-author, virtually speaking, the other's actions without actually performing the action in question (Bråten 2007). In this sense, they may play a role in the forms of concerted rehearsal discussed in Section 3. The observation of another's actions does not, therefore, automatically translate into the observer's execution of a similar action. Languaging is a form of rehearsal in which the extremely fine-grained differentiations made by co-articulated vocal tract gestures and prosodies enable others to recreate in their neural activity similar patterns. Such patterns enable observers to perceive or to simulate an understanding of the other's meanings and perspectives from the observer's own perspective. The concerted nature of languaging also means that languaging agents are entrained to similar, not necessarily identical, patterns. Agents accordingly develop shared expectations as to what the observed behaviours are.

The focus on verbal abstracta and the concomitant failure to provide a naturalistic grounding to languaging behaviour has led to the neglect of the temporal dimension of language behaviour. Timing not just time (let alone clock time) is crucial here, as is rhythm. Co-constructed interaction between two (or more) agents crucially involves timing. The neurological dynamics of the brain function as a system of oscillating processes on very many time scales. Many aspects of observable language dynamics occur on very fast scales. The notion of pico-scale bodily dynamics has been developed to demonstrate the centrality of very short or rapid time-scales that are involved in many aspects of vocalizing, eye movements, rhythm, tone, gesture, and much more. These phenomena are not readily captured by conventional analysis and transcription. These descriptive methods focus on second order abstracta of the kind that linguistics over the past 100 years has specialized in documenting.

Crucially, the central nervous system is a system of oscillators and the modulatory relationships between oscillators. Neurons and neuron circuits are oscillatory and are modulated by influences from other neurons and neuron circuits in distributed networks of brain activity. These, in turn, coordinate the pico-scale bodily dynamics that are co-orchestrated as emergent, inter-individual dynamical patterns when two or more agents co-synchronize their neural and bodily dynamics in dialogically coordinated interaction. Oscillators have the functional capacity to synchronize with and entrain to one another's behaviour (DeLanda 2002: 92). In a population of interacting agents, the capacity to synchronize and entrain to other agents and to other, non-human individuals (e.g. environmental affordances) on different spatio-temporal scales allows for the coordination of internal biological rhythms and external social ones (see also Lefebvre 2004 [1992]: 38-45).

In a population of interacting agents, synchronization is catalyzed by signals -- e.g. optical, chemical, mechanical (DeLanda 2002: 93), auditory, etc. – whose intensity must be maintained at a critical threshold for the synchronization to be maintained over a given time span.

To successfully manage their interaction, agents-in-interaction must co-entrain their neural circuitry and bodily dynamics into a synchronous time-locked behavioural dance. Interactants attune to each other's body dynamics on the pico-scale as a coupled system. What Buzsáki (2006: 8) calls the "feeling of time" of human agents ranges from tens of milliseconds to tens of minutes. It is this time span, according to Buzsáki, which corresponds to the "temporal range of brain oscillators, which may serve as an internal metric for time calibration" (Buzsáki 2006: 8). The best human time resolution is in the sub-seconds range, which corresponds to the pico-scale of the co-constructed bodily dynamics that are central to first-order languaging. Neuronal oscillations related to pico-scale bodily dynamics are crucial to the functioning of languaging (Section 31).

The brains and bodies of individual persons have the capacity to enter into co-synchronous patterns by changing their rhythm so as to converge over a given time span on a common or inter-individual pattern of rhythmic activity. Neurons have different frequencies, which are utilized to pull each other to a synchronized pattern. The stimulation of particular brain regions gives rise to synchronous patterns as various neuronal groups converge on the common pattern. In this way, individual brain structures are created when they synchronize to form coherent neural circuits that create rhythmic oscillations known as brain waves. These observations apply on the intra-individual level and lend support to the earlier research of Condon and Ogston (1966) regarding the "harmonious or synchronous organization of change between body motion and speech" on the intra-individual level. As pointed out above, these researchers also observed these same effects on the interactional or inter-individual level when, for example, two or more persons engage in conversation. It is now known that in interpersonal interaction, persons co-adapt and co-adjust to each other's bodily and neuronal rhythms.

Again, as Condon and Ogston (1966) showed, participants in conversation synchronize their bodily rhythms and movements with one another. They enter into a "resonance pattern" (Hart 2008 [2006]: 83). Research on mirror neurons suggests that oscillating neurons coordinate the interactive dance of neural and bodily dynamics that occurs between persons in conversation. Thus, two or more individuals who are participating in interactional synchrony give rise to resonance patterns that are inter-individual in character. These can vary in duration and intensity. The flow of energy between the body-brain systems of the two (or more) individuals means that novel patterns arise as the two agents co-regulate each other. Researchers in this area have used terms such as "affective attunement" or "intersubjective awareness" to explain the co-regulation of feeling, affect, perception, action and cognition that takes place when inter-individual synchronization occurs. Moreover, mirror neurons explain the ability to connect observation and behaviour in ways that indicate that the other

persons with whom one interacts are "represented" in one's patterns of neuronal activity (Hart 2008 [2006]: 90).

In the present account, I would argue that we do not so much 'empathise' with others in the way that Hart and others have suggested; we become entangled with each other's neural, affective, and bodily dynamics. The notion of 'empathy' continues to promote an individualistic model in which one freestanding person empathises with another, as shown in the usual discussion of the theory of mirror neurons. In this theory, other persons' feelings, intentions, and states of mind are said to be re-created in the neural activity of the observer through processes called mirroring (Hart 2008 [2006]: 91). Hart (2008 [2006]: 92) cites a study by Ekman (2003) who conducted experiments on mood changes in relation to the activation of certain facial muscles. This research shows that if "someone contracts the same facial muscles as the other person, he or she senses the other's emotion" (Hart 2008 [2006]: 92). Hart points out that this is not imitation. Instead, empathetic attunement is achieved on the basis of external gestures, vocalizations, and body movements that activate the mirror neuron circuits in the observer such that the observer can sense another's feelings, moods, intentions, and so on. Infants thus learn to integrate the other's feelings, moods and states of mind through the observation and perception of the other's gestures, facial expressions, and vocalizations in order to achieve affective and other forms of attunement

One problem with this view – for example, that other persons are "represented" in patterns of neuronal activity – is that it does not account for the ways in which interpretation of each other arises as agents perceive, feel and engage with each other's feelings and neural and body dynamics in real-time. Moreover, as stated above, the emphasis remains on individuals who attune to, empathise with, and represent each other. Individuals are still assumed to be separately existing localized entities that are externally related to each other by space and time. I would rather say that, ontologically speaking, individual persons are not definable independently of the nexus of relations and dynamical fields in which they are entangled with others, neurally and bodily. To use a Heiddergerian turn of phrase, we are *with* each other in our *Mitwelt* (Stuart and Thibault Forthcoming). In particular, for the purposes of this paper, agents are entangled in dialogically coordinated social events and communities that are controlled by the brains and bodies of the diverse participants in those events. Such events are not controlled by rules that govern the sequential unfolding of interaction, seen as a series of moves or turns of the kind that are typically featured in discourse-analytical and conversation-analytical approaches to talk (Hodges 2009; Thibault 2008, 2011). These analytical constructs cannot capture the very small temporal scales of the bodily dynamics that are important here. Persons in interaction actively engage with the other's feelings at the same time that the actions (e.g. vocal dynamics, facial expressions, gestures) of the other prompt us to act and to feel in ways that alter the 'feeling of what happens' (Damasio 1999, 2005) between and within persons.

Dialogically coordinated social events make use of pico-scale bodily dynamics that connect people in ways that we interpret as meaningful. Rather than viewing the relationship between self and other as an opposition between subject and object, as in the Cartesian tradition, I

would invoke Heidegger's concept of *Dasein*. In Heidegger's conception, the crucial feature of *Dasein* is it's being **with** the world, being involved with it, being concerned for its world. In agreement with Heidegger, Stuart and Thibault (Forthcoming) argue that the enkinaesthetic enfoldings of agents with their worlds constitute the primordial pre-conceptual understandings that are the very "thrownness" – the term is Heidegger's -- of agents in terms of which all understanding and interpretation between agents and their worlds are created. It is the primordial way of encountering and experiencing the world of new-born infants that enables then to find their way in and with the world, as evidenced in the earliest forms of enkinaesthetic enfolding between parents and new-born infants.

Pico-scale events are bodily dynamics that typically occur on time-scales in the order of fractions of seconds to milliseconds. They are real patterns of inter-individual behaviour though they are not representational patterns: on the time and space scales of human interaction, they enable agents to keep track of both others' and the agent's own concernful engagements with aspects of their worlds. The fact that both pre-linguistic babies and some other species are able to track aspects of selves and others on the basis of such patterns shows that they are not linguistic patterns in the sense that 'language' is usually understood though, as I argue in this paper, they are dynamical patterns that are absolutely fundamental to the workings of first-order languaging. For instance, we humans use words like "embarrassment", "anger", "desire", and so on, as second-order (linguistic) descriptors to index aspects of these patterns which we take to be cognitively, affectively, and interpretatively salient at the community level.

On both the phylogenetic and ontogenetic time scales, such linguistic descriptors are second-order patterns with respect to the first-order patterns of the pico-scale dynamics for contingent historical reasons. On the scale of real-time interaction between persons, patterns on the scale of the pico-scale dynamics are first-order in the sense that they are not representational with respect to any other level. They are, however, real patterns that people respond to, interpret and talk about even though they are, for many people, often much harder to talk about than second-order (linguistic) patterns though this does not mean that they cannot talk about them. What is perhaps most significant is that our concernful enactment of and engagement with these (first-order) dynamical patterns, in enabling persons to move each other and to be moved by others, prompts or gives rise to interpretation. For example, in saying that we feel the anger in someone's voice, or that someone spoke in a friendly voice, and so on, we are responding to some aspects of a real inter-individual pico-scale pattern and attributing it to the properties of individuals (e.g. properties of self and/or other in some interactive encounter).

According to the metaphysic lurking in the ontological shadows of this account, we are treating some aspects of the real patterns so identified as properties of individual agents (e.g. 'Harry is angry'; 'I feel sad'). Following Dennett (1991) and Ladyman and Ross (2007: 243), I would further argue that insofar as and to the extent that these second-order descriptors correctly approximate real patterns we can say that the descriptors endorse a standard metaphysics of individuals as real entities while at the same time failing to correctly account for the ontological status of the real patterns. In other words, individual agents in this

metaphysic are "locators" – the term is from Ladyman and Ross (2007: 243) – for aspects of interactional patterns that we use to keep track of always changing inter-individual interactional dynamics that cannot be correctly correlated with individual behaviours or perspectives per se owing to a history of entanglements between the temporal trajectories of selves-in-interaction (Thibault 2004a, 2004b). This shift in perspective requires an ontology of enfolding or entanglement (Stuart and Thibault Forthcoming). Persons are not skin-bound discrete locators occupying different localized regions of time and space. They are instead temporally unfolding trajectories in a field of mutually overlapping and enfolded affective and other dynamical patterns of reciprocal engagement between trajectories-in-time.

Oscillatory neuronal networks (Sections 3, 6, 14) also imply nonlinear time: simultaneity of oscillations, rather than one event preceding another and hence 'causing' it. Many open questions can be raised at this point: How do human interactants perceive and chunk the perceived body dynamics of self and other? These processes depend on time and timing. How are brain-constructed time and external time different from the points of view of different persons? What are the implications of this for the ways in which humans interact and become 'entangled' with one another's neural and bodily dynamics? How do the (different or convergent?) ways of constructing past experience of persons impinge upon the ways in which co-synchronous brains in interaction co-construct experience? How do the complex, multilevel circular processes of causation involved in the co-constructed dynamics of two (or more) brains in reciprocal interaction modify and complicate our understandings of the oscillatory and modulatory dynamics of brains in real-life biological agents interacting with each other and with aspects of their worlds?

Stuart and Thibault (Forthcoming) propose the neologism 'enkinaesthetic' because trajectories of agents, unlike "chunks of Cartesian space", can, to varying degrees of felt intensity and duration, 'occupy' the same temporal-spatial-energic-semantic field. In this, we are at one with the concept of eidetic field that has been put forward by biologist, Anton Markoš (2002: 69):

> Understanding, orientation, discovering new possibilities is knowledge gained by interpretation –that is, a genuine hermeneutic feat. Eidetic biology understands life processes precisely in these terms. It views living beings as eidetic spaces or semantic fields of formative causation. Like knowledge and learning, development and evolution are based on the interpretation of the embodied living experience stored as *eidos*, that is, as a specific formative strategic way of formation –information, interpreted by imitation within the current context.
>
> At first glance, such an approach to living beings as mindlike existences might seem to be in sharp contrast to the traditional view, according to which the bodily existence is reducible to a particular composition of material parts filling the inert geometrical space. For eidetic biology, the elements of life are elements of meaning rather than chunks of Cartesian space: a mechanism is merely a projection of *eidos* into a geometrical space. But any mechanism is by definition an *eidos* (purpose, strategy, instruction) implemented in a special way of mechanistic causation (as an abacus is an

> implementation of numbers into strings of beads). To make use of a device –mechanical or not—means to understand it.
>
> (Markoš 2002: 69)

Rather than "chunks of Cartesian space" that send coded messages to each other across the empty space that separates the respective chunks, persons are living, moving, and feeling beings who inhabit a dynamical field of energies, relations, reciprocities, meanings, and interpretations. It is through meaning and interpretation that they relate to and connect to their worlds through being enfolded with the neural, affective, and bodily dynamics of each other in a community of reciprocal enfoldings between agents and between agents and objects. Living beings are coherent systems that are "synchronized through many levels of organization" (Markoš 2002: 105) from the most microscopic to the most macroscopic. The essence, if you like, of whole-body-sense-making lies in the way in which meaning-making is not localized at particular articulators such as the vocal tract, but involves, on the intra-individual level identified by Condon and Ogston (1966), many different size scales from the molecular to the whole organism. The travelling wave of a phonetic gesture unifies feeling, perception and action in the whole organism such that response time is immediate (Hollis, Kloos and Van Orden 2009: 213); feeling, perception and action occur on such rapid timescales extending to the most primordial layers of core consciousness (Section 11) that there can be no time for information processing or encoding/decoding. The fractal coupling of travelling waves at all levels of the organism (Davia 2006) suggests that organisms are coherent systems at all levels of organization in ways that allow for and are characterized by "high-efficiency energy transfers with minimum loss." (Markoš 2002: 105). Markoš discusses the work of another biologist, Mae-wan Ho (1993, 1994), in order to shed light on the coherence, synchronization and nonlocal nature of the organization of living beings as "self-structuring fields" (Markoš 2002: 104) across all scalar levels of their organization in contrast to the chemical kinetics of nonliving systems. To quote Ho:

> Intercommunication can proceed so rapidly through the liquid crystalline continuum of the organism that in the limit of the coherence time and coherence volume of energy storage—the time and volume respectively over which the energy remains coherent—intercommunication is instantaneous and nonlocal. There is no time-separation within the coherence volume, just as there is no space-separation within the coherence time. Because the organism stores coherent energy over all space-times, it has a full range of coherent space-times, which are furthermore, all coupled together. (p. 93)
>
> (Ho 1988: 93; quoted in Markoš 2002: 106)

The coherence in living beings is, as Ho (1988, quoted in Markoš) and Markoš (2002: 106) argue, quantum in nature. Living organisms are accordingly interpreted as "highly coherent systems interconnected through many orders of space (10^{-10} to 10^{1}m) and time (10^{-14} to 10^{7} s)" (Markoš 2002: 106). According to Ho:

> One comes to the startling conclusion that the coherent organism is a macroscopic quantum object. It has a macroscopic wavefunction that is always evolving, always changing as it entangles its environment. This wavefunction is the unique, significant form of the organism.
>
> (Chang et al 1998: 94; quoted in Markoš 2002: 106)

Sense-making is a macroscopic energic-semantic wave function of the *whole* living being –- i.e. across all scales of its being --in concert with that being's world. Whole body-sense-making involves the quantum coherence of the entire organism. Whole body-sense-making cannot be localized as the outputs or inputs of particular performatory or receptive systems of the body such as, for example, the vocal tract and the sense receptors. The idea of entanglement, mentioned above, suggests that individuals become entangled with each other at the organismic level on the basis of their "macroscopic wave function that is always evolving, always changing as it entangles its environment" Chang et al 1998: 94; quoted in Markoš 2002: 106). The wave function of living beings is a dynamical temporal pattern or "kinetic melody" (Luria 1973: 32, 36; Stuart and Thibault Forthcoming) of bodily action and feeling that is the "unique signature" of the individual. Ho (1988) and Chang et al (1998) thus interpret organisms as macroscopic quantum objects.

Ho therefore postulates that the coherence of living beings is quantum in nature (Markoš 2002: 106). This means that the characteristics of quantum physics are not confined to the microscopic domain; they are coupled to the macroscopic world, meaning that quantum effects also pervade the macroscopic domain, as Ho's postulate concerning the quantum nature of living beings suggests. There can be no return to the notion that living beings are just "chunks of Cartesian space" that can be defined in terms of the ever smaller components out of which they are composed, as in the classical model. All organizational scales of living beings are implicated in whole body-sense-making. Whole-body-sense-making is a projection of the quantum coherence of the whole living being in concert with its world.

The earlier findings of Ho resonate well with those of Davia, discussed above (Section 13). The essentially instantaneous and nonlocal character of the intercommunication across all fractal levels of the organism's quantum coherence means that the whole organism is an excitable medium of coherent space-times on different fractal levels, all of which are fully coupled. Language activity is whole-body-sense-making. It is truly immediate and works over the whole range of the wavelengths of the body's quantum coherence to yield a global, nonlocal bodily response as agents catalyze their worlds in and through their dialogically coordinated languaging behaviour.

References

Abler, William (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures* 12: 1-13.

Alač, Morana and Hutchins, Edwin (2004). I see what you are saying: action as cognition in fMRI brain mapping practice. *Journal of Cognition and Culture* 4:3: 629-661.

Allen, Jedediah W. P. and Bickhard, Mark H. (2011). Normativity: A crucial kind of emergence: Commentary on Witherington. *Human Development* 54: 106-112.

Argyropoulos, Giorgos P. (2008). The subcortical foundations of grammaticalization. In: A. D. M. Smith, K. Smith, and R. Ferrer i Cancho (Eds.), *The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language*. Singapore: World Scientific Press, pp. 10-17.

Arndt, Horst and Janney, Richard Wayne (1987). *Intergrammar: Towards an integrative model of verbal, prosodic and kinesic choices in speech*. Berlin and New York: Mouton de Gruyter.

Baldry, Anthony and Thibault, Paul J. (2006). *Multimodal Transcription and Text Analysis: A multimedia toolkit and course book* (with associated online course). Foreword by Jay Lemke. London, UK and Oakville, CT: Equinox.

Barbieri, Marcello (2003). *The Organic Codes: An introduction to semantic biology*. Cambridge, UK: Cambridge University Press.

Bickhard, Mark H. (1998). Levels of representationality. *Journal of Experimental and Theoretical Artificial Intelligence* 10: 179-215.

Bickhard, Mark H. (2004). The dynamic emergence of representation. In: H. Clapin, P. Staines, P. Slezak (Eds.), *Representation in Mind: New Approaches to Mental Representation*. Amsterdam, The Netherlands and Boston, MA: Elsevier, pp. 71-90.

Bickhard, Mark H. (2008). Social ontology as convention. *Topoi 27*(1-2): 139-149.

Bottineau, Didier (2007). The cognemes of the Spanish language: Towards a cognitive modelization of the submorphemic units in the grammatical words of the Spanish language. *The Public Journal of Semiotics* I(2), July 2007: 50-74.

Bouissac, Paul (2006). The optic, haptic, and acoustic dimensions of gestures: evolutionary significance and methodological implications. Berlin Gesture Center, Interdisziplinäres BGC-Kolloquium Freitag, 10. November 2006; [http://www.semioticon.com/people/articles/Gesture.htm]; downloaded 2nd October 2010.

Bouisssac, Paul (2010). *Saussure: A guide for the perplexed*. London and New York: Continuum.

Bråten, Stein (2007). Altercentric infants and adults: On the origins and manifestations of participant perception of others' acts and utterances. In: Stein Bråten (Ed.), *On Being Moved: From mirror neurons to empathy*. Amsterdam/Philadelphia: John Benjamins, pp. 111-135.

Bråten, Stein and Trevarthen, Colwyn (2007). Prologue: from infant intersubjectivity and participant movements to simulation and conversation in cultural common sense. In: Stein Bråten (Ed.), *On Being Moved: From mirror neurons to empathy*. Amsterdam/Philadelphia: John Benjamins, pp. 21-34.

Browman, Catherine P. and Goldstein, Louis (1992). Articulatory phonology: an overview. *Phonetica* 49: 155-180.

Browman, Catherine P. and Goldstein, Louis (1995). Dynamics and articulatory phonology. In: Robert P. Port and Timothy van Gelder (Eds.), *Mind as Motion: Explorations in the dynamics of cognition*. Cambridge, MA and London, UK: The MIT Press, pp. 175-193.

Browman, Catherine P. and Goldstein, Louis (2002). *Articulatory Phonology*. Draft: Chapter 1, Speech production actions as phonological primitives. [http://sail.usc.edu/~lgoldste/ArtPhon/Papers/Week%201/AP_ch1.pdf], downloaded: 22nd November 2011.

Buzsáki, György (2006). *Rhythms of the Brain*. Oxford and New York: Oxford University Press.

Bybee, Joan L. (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication* 2: 249-278.

Bybee, Joan L. and Thompson, Sandra A. (2000). Three frequency effects in syntax. *Berkeley Linguistic Society* 23: 65-85.

Cassell, Justine and Tartaro, Andrea (2007). Intersubjectivity in human-agent interaction. *Interaction Studies 8*(3): 391-410.

Chang, Jiin-Jiu, Fisch, Joachim, and Popp, Fritz-Albert (Eds.), (1998). *Biophotons*. Dordrecht: Kluwer.

Carpenter, Patricia A. and Davia, Christopher J. (2006). A catalytic theory of embodied mind. http://csjarchive.cogsci.rpi.edu/proceedings/2006/docs/p1080.pdf. Downloaded: 12th May 2011.

Clark, Andy (1997). *Being There: Putting brain, body and world together again*. Cambridge, MA: The MIT Press.

Clark, Andy (2008). *Supersizing the Mind*: *Embodiment, action and cognitive extension*. Oxford: Oxford University Press.

Condon, W. S., and Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Diseases* 143: 338-457.

Condon, William S. (1970). Method of micro-analysis of sound films of behavior. *Behavior Research Methods, Instruments & Computers* 2(2): 51-54.

Condon, William S. and Sander, Louis W. (1974). Neonate movement is synchronized with adult speech: interactional participation and language acquisition. *Science* 183 (January 1974): 99-101.

Corballis, Michael C. (2011). *The Recursive Mind: The origins of human language, thought, and civilization*. Princeton and Oxford: Princeton University Press.

Cowley, Stephen J. (2006). Language and biosemiosis: a necessary unity? *Semiotica* 162(1/4): 417-444.

Cowley, Stephen J. (2007). How human infants deal with symbol grounding. *Interaction Studies* 8(1): 81-104.

Cowley, Stephen J. (2007). Cognitive dynamics and distributed language. *Language Sciences* 29(5): 575-583.

Cowley, Stephen J. (2008). The codes of language: turtles all the way up?. In: Marcello Barbieri (Ed.), *The Codes of Life: The rules of macroevolution*. Berlin: Springer, pp. 319-345.

Cowley, Stephen J. (2011). Taking a language stance. *Ecological Psychology* 23(3): 185-209. To link to this article: http://dx.doi.org/10.1080/10407413.2011.591272.

Cowley, Stephen, Moodley, S., and Fiori-Cowley, Agnese (2004). Grounding signs of culture: primary intersubjectivity in social semiosis. *Mind, Culture, Activity* 11(2): 109-132.

Damasio, Antonio (1999). *The Feeling of What Happens: Body, emotion and the making of consciousness*. London and Sydney: Heinemann.

Damasio, Antonio (2005). The neurobiological grounding of human values. In: J.-P. Changeux, A. R. Damasio, W. Singer, and Y. Christian (Eds.), *Neurobiology of Human Values*. Berlin and New York: Springer Verlag, pp. 47-56.

Damasio, Antonio (2010). *Self Comes to Mind: Constructing the conscious brain*. London: Heinemann.

Davia, Christopher J. (2006). Life, catalysis and excitable media: A dynamic systems approach to metabolism and cognition. In: Jack Tuszynski (Ed.), *The Physical Basis for Consciousness*. Heidelberg: Springer-Verlag, pp. 255-292.

Davidse, Kristin (1997). The Subject-Object versus the Agent-Patient asymmetry. *Leuvense Bijdragen* (Leuven Contributions in Linguistics and Philology) 86(4): 413-431.

Deacon, Terrence W. (1996). Prefrontal cortex and symbol learning: why a brain capable of language evolved only once. In: Boris M. Velichkovsky and Duane M. Rumbaugh (Eds.), *Communicating Meaning: The evolution and development of language*. Malwah, NJ: Erlbaum, pp. 102-138.

DeLanda, Manuel (2002). *Intensive Science and Virtual Philosophy*. London and New York: Continuum.

DeLanda, Manuel (2011). *Deleuze: History and Science*. New York and Dresden: Atropos Press.

Deleuze, Gilles and Guattari, Félix (1987). *A Thousand Plateaus*. Minneapolis, MN: University of Minnesota Press.

Deleuze, Gilles 2004 [1968]. *Difference and Repetition*. Trans. Paul Patton. London and New York: Continuum.

Deleuze, Gilles (1994). *Difference and Repetition*. New York: Columbia University Press.

Demos, E. Virginia (1988). Affect and the development of the self: a new frontier. In: Arnold Goldberg (Ed.), *Frontiers in Self Psychology: Progress in self psychology, Volume 3*. Hillsdale, NJ and London, UK: The Analytic Press, pp. 27-53.

Dennett, Daniel C. (1991). Real patterns. *The Journal of Philosophy* 88(1): 27-51.

Edelman, Gerald M. (1987). *Neural Darwinism: The theory of neuronal group selection*. New York: Basic Books.

Edelman, Gerald M. (1989). *The Remembered Present. A biological theory of consciousness*. New York: Basic Books.

Ekman, Paul (2003). *Emotions Revealed*. New York: Times Books.

Fontana, Walter and Buss, Leo (1996). The barrier of objects: From dynamical systems to bounded organizations. In: John Casti and Anders Karlquist (Eds.), *Boundaries and Barriers*. Reading, MA: Addison-Wesley, pp. 56-116.

Fowler, Carol A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics* 14: 3-28.

Fowler, Carol A. (2010). Embodied, embedded language use. *Ecological Psychology* 22(4): 286-303. To link to Carol Fowler's Article: DOI: 10.1080/10407413.2010.517115; URL: http://dx.doi.org/10.1080/10407413.2010.517115

Freeman, Walter J. (1995). *Societies of Brains: A study in the neuroscience of love and hate*. Hillsdale, NJ and Hove, UK: Lawrence Erlbaum.

Freeman, Walter J. (2000). *How Brains Make Up their Minds*. New York: Columbia University Press.

Gabora, Liane (2004). Ideas are not replicators but minds are. *Biology and Philosophy* 19(1): 127-143.

Gafos, Adamantios I. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20(2): 269-337.

Gafos, Adamantios I. (2006). Dynamics in grammar: comment on Ladd and Ernestus and Baayen. In: M. L. Goldstein, D. H. Whalen, and C. Best, (Eds.), *Laboratory Phonology 8: Varieties of Phonological Competence*. Berlin/New York: Mouton de Gruyter, pp. 51-79.

Galantucci, Bruno (2005). An experimental study of the emergence of human communication systems. *Cognitive Science* 29: 737-767.

Gallese, V., Luciano F., Leonardo F., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119: 593–609.

Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences* 8: 396-403.

Garfinkel, Alan (1981). *Forms of Explanation: Rethinking the questions in social theory*. New Haven and London: Yale University Press.

Gibson, James J. (1977). *The Theory of Affordances.* Hillsdale, NJ: Erlbaum.

Gibson, James J. (1983 [1966]). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

Gibson, James J. (1986 [1979]). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Erlbaum.

Gick, Bryan and Derrick, Donald (2009). Aero-tactile integration in speech perception. *Nature* Vol. 462, November 2009: 502-504.

Goldstein, Louis M., Fowler, Carol (2003). Articulatory phonology: a phonology for public language use. In: Antje S. Meyer and Niels O. Schiller, (Eds.), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, Berlin: Mouton de Gruyter, pp. 159-207. http://sail.usc.edu/~lgoldste/ArtPhon/Papers/Week%201/Goldstein-Fowler.pdf (Downloaded: 13 August 2011)

Goffman, Erving (1983). The interaction order. *American Sociological Review* 48: 1-17.

Halliday, M. A. K. (1975). *Learning How to Mean: Explorations in the development of language.* London: Arnold.

Halliday, M. A. K. (1991). Towards probabilistic interpretations. In: Eija Ventola (Ed.), *Functional and Systemic Linguistics: Approaches and uses*. Berlin: Mouton de Gruyter, pp. 39-61.

Halliday, M. A. K. (2004 [1985]). *Introduction to Functional Grammar*. 3rd edition. Revised by Christian M. I. M. Matthiessen. London and Melbourne: Arnold.

Harris, Roy (1980). *The Language Makers*. Ithaca, New York: Cornell University Press.

Harris, Roy (1981). *The Language Myth*. London: Duckworth.

Harris, Roy (1990). On redefining linguistics. In: Hayley G. Davis and Talbot J. Taylor (Eds.), *Redefining Linguistics*. London and New York: Routledge, pp. 18-52.

Harris, Roy (2000.) *Rethinking Writing*. London: Athlone Press.

Harris, Zellig (1952). Discourse analysis. *Language* 28(1): 1-30.

Harris, Zellig (1991). *A Theory of Language and Information: A mathematical approach*. Oxford and New York: Clarendon Press.

Hart, Susan (2008 [2006]). *Brain, Attachment, Personality: An introduction to neuroaffective development*. London: Karnac.

Hart, Susan 2011 [2006]. *The Impact of Attachment: Developmental neuroaffective psychology*. New York and London: W. W. Norton.

Hasan, Ruqaiya (2009) [1992]). Rationality in everyday talk: from process to system. In: Jonathan J. Webster (Ed.), *Semantic Variation: Meaning in society and in sociolinguistics, The collected works of Ruqaiya Hasan, Volume 2*, London, UK and Oakville, CT: Equinox, pp. 309-352.

Ho, Mae-Wan (1988). On not holding nature still: evolution by process, not by consequence. In: M-W. Ho and S. W. Fox (Eds.), *Evolutionary Processes and Metaphors*. Chichester, U.K.: Wiley, pp. 117-114.

Ho, Mae-Wan (1993). *The Rainbow and the Worm*. Singapore: World Scientific.

Ho, Wae-Man (1994). What is (Schrödinger's) negentropy?. In: E. Gnaiger et al (Eds.), *What is Controlling Life?*. Innsbruck: Insbruck University Press, pp. 50-61.

Hockett, Charles D. (1960). The origins of speech. *Scientific American* 203: 89-96.

Hodges, Bert H. (2007). Good prospects: ecological and social perspectives on conforming, creating, and caring in conversation. *Language Sciences* 29: 584-604.

Hodges, Bert H. (2009). Ecological pragmatics: values, dialogical arrays, complexity, and caring. *Pragmatics & Cognition* 17(3): 628-652. doi 10.1075/p&c.17.3.08hod issn 0929–0907 / e-issn 1569–9943.

Hollis, Geoff, Kloos, Heidi, and Van Orden, Guy C. (2009). Origins of order in cognitive activity. In: S. J. Guastello, M. Koopmans, and D. Pincus (Eds.), *Chaos and Complexity in Psychology: The Theory of Nonlinear Dynamical Systems.* Cambridge, MA: Cambridge University Press, pp. 206-241.

Hutchins, Edwin (1995a). *Cognition in the Wild*. Cambridge, MA: The MIT Press.

Hutchins, Edwin (1995b). How a cockpit remembers its speeds. *Cognitive Science* 19: 265-288.

Hutchins, Edwin (2010). Cognitive ecology. *Topics in Cognitive Science* 2: 705-715.

Hutchins, E., and Klausen, T. (2000) Distributed cognition in an airline cockpit: Cognition and communication at work. In: *Cognition and communication at work*, Y. Engström and D. Middleton, (Eds.). Cambridge University Press: New York, pp. 15-34.

Hutto, Daniel D. (2008). First communions: mimetic sharing without theory of mind. In: Jordan Zlatev, Timothy P. Racine, Chris Sinha, and Esa Itkonen (Eds.), *The Shared Mind: Perspectives on intersubjectivity*. Amsterdam/Philadelphia: John Benjamins, pp. 245-276.

Ihde, Don (1979). *Technics and Praxis*. Dordrecht and Boston: D. Reidel.

James, William (1899). Apperception. In: *Talks to Teachers*, Chap. 14, Cambridge, MA. [http://www.des.emory.edu/mfp/ttpreface.html]; downloaded 23rd November 2011.

Juarrero, Alicia (1999). *Dynamics in Action: Intentional behavior as a complex system*. Cambridge, MA and London, UK: The MIT Press.

Keijzer, F. A. (2003). Self-steered self-organization. In: W. Tschacher, and J. -P. Dauwalder (Eds.), *The Dynamical Systems Approach to Cognition*. River Edge, NJ: World Scientific, pp. 243-259.

Kinsbourne, Marcel 2005. Imitation as entrainment: brain mechanisms and social consequences. In: Susan Hurley and Nick Chater, (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science. Volume 2: Imitation, human development, and culture*. Cambridge, MA and London, England: The MIT Press, pp. 163-172.

Kirsh, David (1995a). The intelligent use of space. *Artificial Intelligence* 73: 31-68.

Kirsh, David (1995b). Complementary strategies: why we use our hands when we think. In: J. Moore and J. Fain Leman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* . Mahwah: NJ: Lawrence Erlbaum, pp. 212-217.

Koffka, Kurt (1910). *Über latente Einstellung.* Bericht *über* den IV. Kongress für exper. Psychol. in Innsbruck.

Koffka, Kurt (1935). *Principles of Gestalt Psychology*. New York: Harcourt, Brace & World.

Kravchenko, Alexander (2003). *Sign, Meaning, Knowledge: An essay in the cognitive philosophy of language*. Frankfurt am Main and New York: Peter Lang.

Kress, Gunther and Van Leeuwen, Theo (2001). *Multimodal Discourse: The modes and media of contemporary communication*. London: Arnold.

Kuhl, Patricia K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science* 10(1): 110–120.

Kuhl, Patricia K. and Meltzoff, Andrew N. (1982). The bimodal perception of speech in infancy. *Science* 218: 1138-1141.

Kuhl, Patricia K. and Meltzoff, Andrew N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development* 7: 361-381.

Ladyman, James and Ross, Don (with Spurrett, David and Collier, John). (2007). *Every Thing Must Go: Metaphysics naturalized*. Oxford and New York: Oxford University Press.

Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar, Volume I: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Laird, James D. (2007). *Feelings: The perception of self*. Oxford and New York: Oxford University Press.

LeDoux, Joseph E. (1996). *The Emotional Brain: The mysterious underpinnings of emotional life*. New York: Simon & Schuster.

LeDoux, Joseph E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience* 23: 155-184.

Lefèbvre, Henri (2004 [1992]). *Rhythmanalysis: Space, time and everyday life*. Trans. Stuart Elden and Gerald Moore. London and New York: Continuum.

Lemke, Jay L. (2000a). Across the scales of time: artefacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity* 7(4): 273-290.

Lemke, Jay L. (2000b). Material sign processes and emergent ecosocial organization. In: Peter Bøgh Andersen, Claus Emmeche, Niels Ole Finnemann, and Peder Voetmann Christiansen (Eds.), *Downward Causation: Minds, bodies and matter*. Aarhus: Aarhus University Press, pp. 181-213.

Liberman, Alvin M. and Mattingly, Ignatius G. (1985). The motor theory of speech perception revised. *Cognition* 21: 1-36.

Lieberman, Philip (1967). *Intonation, Perception and Language*. Cambridge, MA: The MIT Press.

Linell, Per (2005). *The Written Language Bias in Linguistics: Its nature, origin and transformations.* Oxford: Routledge.

Linell, Per (2009). *Rethinking Language, Mind, and World Dialogically: Interactional and contextual theories of human sense-making*. Charlotte, NC: Information Age Publishing, Inc.

Love, Nigel (1990). The locus of languages in a redefined linguistics. In: Hayley G. Davis and Talbot J. Taylor (Eds.), *Redefining Linguistics*. London and New York: Routledge, pp. 53-117.

Love, Nigel (2007). Are languages digital codes?. *Language Sciences* 29: 690-709.

Luchins, A.S. (1942). Mechanization in problem solving: The effect of einstellung. *Psychological Monograph* 54, whole No. 248.

Luria, A. R. (1973). *The Working Brain: An introduction to neuropsychology*. Trans. Basil Haigh. London: Allen Lane.

Macherey, Pierre (2009). *De Canguilheim à Foucault: La force des normes*. Paris: La Fabrique Éditions.

MacNeilage, P. F., Rootes, T. P., and Chase, R. A. (1967). Speech production and perception in a patient with severe impairment of somesthetic perception and motor control. *Journal of Speech and Hearing Research*, 10: 449-467.

MacNeilage, MacNeilage, P. F., Studdert-Kennedy, M., and Lindblom, B. (1987). Primate handedness reconsidered. *Behavioral and Brain Sciences* 10: 247-303.

Malinowski, Bronislav (1936). (1936). The problem of meaning in primitive languages. In: CK. Ogden and I.A. Richards, *The Meaning of Meaning*, [1923], Supplement I. London: Kegan Paul, pp. 296-336.

Markoš, Anton (2002). *Readers of the Book of Life: Contextualizing developmental evolutionary biology*. Oxford and New York: Oxford University Press.

Maturana, Humberto R. (1970). *Biology of Cognition*. BCL Report 9.0. University of Illinois, Urbana.

Maturana, Humberto R. and Varela, Francisco J. (1980). *Autopoesis and Cognition*. Dordrecht: Reidel.

Maturana, Humberto R. and Varela, Francisco J. (1987). *The Tree of Knowledge: The biological roots of human understanding*. Revised Edition. Boston: Shambala.

Mayr, Ernst (1976). Typological versus population thinking. In: *Evolution and the Diversity of Life*. Cambridge, MA: Harvard University Press, pp. 26-29.

Melser, Derek (2004). *The Act of Thinking*. Cambridge MA: The MIT Press.

Melser, Derek (2009). Verbal communication: from pedagogy to make-believe. Pre-publication draft of paper in *Language Sciences* 31(5), Sept 2009. http://www.derekmelser.org/essays/essayverbal.html [Downloaded 8th January 2010].

Meltzoff, Andrew N. and Moore, M. Keith (1977). Imitation of facial and manual gestures by human neonates. *Science* 198: 75-78.

Mouilso, Emily, Glenberg, Arthur M., Havas, David, and Lindeman, Lisa M. (2007). Difference in action tendencies distinguish anger and sadness after comprehension of emotional sentences. In: D. S. McNamara and J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society*. Austin, TX: Cognitive Science Society, pp. 1325-1330.

Noë, Alva (2004). *Action in Perception*. Cambridge, MA and London, England: The MIT Press.

O'Regan, J. Kevin and Noë, Alva (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24: 939-1031.

Pattee, Howard H. (2000). Causation, control and the evolution of complexity. In: Peter Bøgh Andersen, Claus Emmeche, Niels Ole Finnemann, and Voetmann Peder Christiansen (Eds.), *Downward Causation*. Aarhus: Aarhus University Press, pp. 63-77.

Pattee, Howard H. (2001). The physics of symbols: bridging the epistemic cut. *Biosystems* 60: 5-21.

Port, Robert (2007). The graphical basis of phones and phonemes. In: Ocke-Schwen and Murray J. Munro (Eds.), *Language Experience in Second language Speech Learning: In honor of James Emil Flege*. Amsterdam/New York: John Benjamins, pp. 349-365.

Port, Robert (2008). All is prosody: phones and phonemes are the ghosts of letters. Key Note Address for *Prosody 2008*, Campinas, Brazil, May 2008. To appear in the conference proceedings. [http://www.cs.indiana.edu/%7Eport/HDphonol/Port.Prosody2008.keynote.pdf] Downloaded 27th September 2010.

Port, Robert F. (2010). Language as a social institution: why phonemes and words do not live in the brain. *Ecological Psychology* 22: 304-326.

Rączaszek, J., Tuller, B., Shapiro, L. P., Case, P., and Kelso, S. (1999). Categorization of ambiguous sentences as a function of a changing prosodic parameter: a dynamical approach. *Journal of Psycholinguistic Research* 28(4): 367-393.

Reed, Edward S. (1996a). *Encountering the World: Toward an ecological psychology*. New York and London: Oxford University Press.

Reed, Edward S. (1996b). *The Necessity of Experience*. New Haven: Yale University Press.

Rizzolatti, G., Fadiga L., Gallese V., and Fogassi L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3: 131–141.

Rizzolatti, G. and Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences* 21: 188-194.

Rosen, R. (1985). *Anticipatory Systems: Philosophical, mathematical and methodological foundations*. Oxford: Pergamon.

Ross, Don (2007). *H. sapiens* as ecologically special: what does language contribute?. *Language Sciences* 29: 710-731.

Salthe, Stanley N. (2011). Naturalizing Information. *Information* 2: 417-425. doi:10.3390/info2030417; ISSN 2078-2489; www.mdpi.com/journal/information.

Sancier, M., and Fowler, C. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese. *Journal of Phonetics* 25: 421-436.

Sander, Louis (1977). The regulation of exchange in the infant-character system and some aspects of the context-content relationship. In: M. Lewis and L. Rosenblum, (Eds.), *Interaction, Conversation, and the Development of Language*. New York: Wiley, pp. 133-156.

Saussure, Ferdinand de (1993). Eisuke Komatsu (Ed.), *Cours de Linguistique Générale: Premier et troisième cours d'après les notes de Reidlinger et Constantin*. [*Cours* III, I^{ère} *partie*: 1910-1911]. Collection Recherches Université Gaskushuin n^{o} 24. Tokyo: Université Gakushuin.

Savage-Rumbaugh, Sue (2011). Human language – human consciousness. [http://onthehuman.org/2011/01/human-language-human-consciousness/]; On the Human: a project of the National Humanities Center.

Scheflen, Albert E. (with Ashcraft, Norman). (1976). *Human Territories: How we behave in space-time*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Schögler, Ben and Trevarthen, Colwyn (2007). To sing and dance together: from infants to jazz. In: Stein Bråten (Ed.), *On Being Moved: From mirror neurons to empathy*. Amsterdam/Philadelphia: John Benjamins, pp. 281-302.

Schrödinger, Erwin (1935). Die gegenwärtige Situation in der Quantenmechanik, *Naturwissenschaften* 23: 807-812; 823-828; 844-849. [English translation: 'The present situation in quantum mechanics: a translation of Schrödinger's "Cat Paradox Paper", originally published in *Proceedings of the American Philosophical Society*, 124, 323-38. [And then appeared as Section I.11 of Part I of *Quantum Theory and Measurement* (J.A. Wheeler and W.H. Zurek, Eds.). Princeton University Press, New Jersey 1983). [http://www.tu-harburg.de/rzt/rzt/it/QM/cat.html, downloaded 16th June 2011]

Shankweiler, Donald, Strange, Winifred, and Verbrugge, Robert (1977). Speech and the problem of perceptual constancy. In: Robert Shaw and John Bransford (Eds.), *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Hillsdale, NJ: Erlbaum, pp. 315-346.

Simondon, Gilbert (2010). *Communication et Information: Cours et conférences*. Ed. Nathalie Simondon. Chatou: Les Éditions de la Transparence.

Stark, Rachel E., Rose, Susan N., and McLagen, Margaret (1974). Features of infant sounds: the first eight weeks of life. *Journal of Child Language* 2: 205-221.

Steels, Luc (2006). Semiotic dynamics for embodied agents. *IEEE Intelligent Systems* 21(3): 32-38. doi:10.1109/MIS.2006.58.

Steffensen, S.V. (2009). Language, languaging, and the extended mind hypothesis. *Pragmatics & Cognition* 17(3): 677-698.

Steffensen, Sune V., Thibault, Paul J. and Cowley, Stephen J. (2010). Living in the social meshwork: the case of health interaction. In: Stephen Cowley, João C. Major, Sune V. Steffensen and Alfredo Dinis (Eds.), *Signifying Bodies: Biosemiosis, interaction and health*. Braga: The Faculty of Philosophy of Braga Portuguese Catholic University, pp. 207-244.

Stern, Daniel N. (1977). *The First Relationship*. Cambridge, MA: Harvard University Press.

Stern, Daniel N. (1984). Affect attunement. In: J. D. Call, E. Galensch, and R. L. Tyson, (Eds.), *Frontiers of Infant Psychiatry, Volume 2*. New York: Basic Books, pp. 3-13.

Stern, Daniel N. (2004). *The Present Moment in Psychotherapy and Everyday Life*. New York: W. W. Norton.

Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17: 3–45.

Stuart, Susan and Thibault, Paul J. Forthcoming. Enkinaesthetic polyphony as the underpinning for first-order languaging.

Studdert-Kennedy, Michael and Goldstein, Louis (2003). Launching language: the gestural original of discrete infinity. In: M.H. Christiansen and S. Kirby, (Eds.), *Language Evolution: The states of the art*. Oxford University Press, pp. 235-255.

Swenson, R., and Turvey, M. T. (1991). Thermodynamic reasons for perception-action cycles. *Ecological Psychology 3:* 317-348.

Thibault, Paul J. (1994). Text and/or context?. State-of-the-Art article. In: *The Semiotic Review of Books* (Toronto) 4(1) (May 1994): 10-12.

Thibault, Paul J. (1997). *Re-reading Saussure: The dynamics of signs in social life*. London and New York: Routledge.

Thibault, Paul J. (2004a). *Brain, Mind, and the Signifying Body: An ecosocial semiotic theory*. Foreword by M. A. K. Halliday. London and New York: Continuum.

Thibault, Paul J. (2004b). *Agency and Consciousness in Discourse: Self-other dynamics as a complex system*. London and New York: Continuum.

Thibault, Paul J. (2005a). What kind of minded being has language: Anticipatory dynamics, arguability, and agency in a normatively and recursively self-transforming learning system, Part 1. *Linguistics and the Human Sciences* 1(2): 261-335.

Thibault, Paul J. (2005b). What kind of minded being has language: Anticipatory dynamics, arguability, and agency in a normatively and recursively self-transforming learning system, Part 2. *Linguistics and the Human Sciences* 1(3): 355-401.

Thibault, Paul J. (2005c). The interpersonal gateway to the meaning of mind: unifying the inter- and intraorganism perspectives on language. In: Ruqaiya Hasan, Christian Matthiessen, and Jonathan J. Webster (Eds.), *Continuing Discourse on Language: A Functional Perspective, Vol. 1*. A volume presented to Michael Halliday on the occasion of his 80th birthday at the International Systemic Functional Linguistics Congress, Sydney, July 2005. London and Oakville, CT: Equinox, pp. 117-156.

Thibault, Paul J. (2008). Face-to-face communication and body language. In: Karlfried Knapp and Gerd Antos (Eds.), *Handbooks of Applied Linguistics (HAL) Linguistics for Problem-Solving: Perspectives on Communication Competence, Language and Communication Problems, and Practical Solutions*. Volume 2: *Interpersonal Communication*. Gerd Antos & Eija Ventola (eds.). Berlin. Mouton, pp. 285-330.

Thibault, Paul J. (2011). First-order languaging dynamics and second-order language: the distributed language view. *Ecological Psychology* 23(3): 210-245. To link to this article: http://dx.doi.org/10.1080/10407413.2011.591274

Trager, George L. (1958). Paralanguage: A first approximation. *Studies in Linguistics* 13(1): 1–12.

Trevarthen, Colwyn (1992). An infant's motives for speaking and thinking in the culture. In: Astri Heen Wold (Ed.), *The Dialogical Alternative: Towards a theory of language and mind*. Oslo: Scandinavian University Press.), pp. 99-137.

Trevarthen, Colwyn (1998). The concept and foundations of infant intersubjectivity. In: Stein Bråten (Ed.), *Intersubjective Communication and Emotion in Early Ontogeny*. Cambridge: Cambridge University Press, pp. 15-46.

Turvey, M.T. (1977). Preliminaries to a theory of action with reference to vision. In: R. Shaw and J. Bransford (Eds.), *Perceiving, Acting and Knowing: Toward an ecological psychology. Hillsdale*, NJ: Lawrence Erlbaum Associates, pp. 211-265.

Turvey, M. T., and Shaw, R. E. (1995). Toward an ecological physics and a physical psychology. In: R. Solso and D. Massaro (Eds.), *The Science of Mind: 2001 and beyond*. New York: Oxford University Press, pp. 144-172.

Vallée-Tourangeau, Frédéric, Euden, Gemma and Hearn, Vanessa (2011). Einstellung defused: Interactivity and mental Set. *Quarterly Journal of Experimental Psychology* 64(10): 1889-1895.

Van Orden, Guy C. and Holden, John G. (2002). Intentional contents and self-control. *Ecological Psychology* 14(1-2): 87-109.

Van Orden, G.C., Moreno, M.A., and Holden, J.G., (2003). A proper metaphysics for cognitive performance. *Nonlinear Dynamics, Psychology, and Life Sciences* 7: 49–60.

Van Orden, G.C., Holden J.G., and Turvey, M.T. (2005). Human cognition and 1/f scaling. *Journal of Experimental Psychology: General* 134: 117–123.

Van Orden, G.C., Kello, C.T., and Holden, J.G. In press. Situated behavior and the place of measurement in psychological theory. *Ecological Psychology*.

Verbrugge, Robert R. (1977). Resemblances in language and perception. In: Robert Shaw and John Bransford (Eds.), *Perceiving, Acting, and Knowing: Toward an ecological psychology*. Hillsdale, NJ: Lawrence Erlbaum, pp. 365-389.

Verbrugge, Robert R. (1980). Transformations in knowing: a realist view of metaphor. In: R. P. Honeck and R. R. Hoffman (Eds.), *Cognitive and Figurative Language*. Hillsdale, NJ: Lawrence Erlbaum, pp. 87-125.

Verbrugge, Robert R. (1985). Language and event perception: Steps toward a synthesis. In: W. H. Warren, Jr. and R. E. Shaw (Eds.), *Persistence and Change*. Hillsdale, NJ: Lawrence Erlbaum, pp. 157-194.

Vihman, Marilyn M. (1996). *Phonological Development: The origins of language in the child*. Cambridge, MA: Blackwell.

Von Uexküll, J. 1928. *Theoretische Biologie*. Frankfurt am Main: Suhrkamp.

Warren, William H. Jr. and Verbrugge, Robert R. (1981). Auditory information for breaking and bouncing events: a case study in ecological acoustics. *Haskins Laboratories: Status*

Report on Speech Research SR-67/68.

Werner, Heinz and Kaplan, Bernard (1984 [1963]). *Symbol Formation: An organismic-developmental approach to the psychology of language*. Hillsdale, NJ: Lawrence Erlbaum.

Wheeler, Michael (2005). *Reconstructing the Cognitive World: The next step*. Cambridge, MA and London, England: The MIT Press.

Wilson, R. A., and Clark, A. (2009). How to situate cognition: Letting nature take its course. In: P. Robbins and M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition.* Cambridge: Cambridge University Press, pp. 55-77.

Zlatev, Jordan (2008). The co-evolution of intersubjectivity and bodily mimesis. In: Jordan Zlatev, Timothy P. Racine, Chris Sinha, and Esa Itkonen (Eds.), *The Shared Mind: Perspectives on intersubjectivity*. Amsterdam/Philadelphia: John Benjamins, pp. 215-244.

Towards a Social Semiotic Approach of the Analysis of Emotion in Sound and Music

David Machin,

Cardiff School of Journalism, Media and Cultural Studies, Cardiff University
Cardiff, United Kingdom

Abstract

This paper explores the possibility of a social semiotic approach to popular music in order to document the voice qualities, melodies and instrumental sounds found in popular music using the case study of Michael Jackson's 'Billie Jean'. The paper explores how these semiotic features, as well as the lyrics, communicate and reveal something of the subjectivity, the emotional world, created by musicians. Here the analysis explores Jackson's use of restricted pitch ranges, breathiness, melodic patterns and gentle disjunctive articulation showing how these help to communicate the 'tension', and 'trouble' in the song. Taking a social semiotic approach the paper identifies the underlying available repertoire of meanings available to musicians and shows how musicians can skilfully draws upon these.

Introduction

Michael Jackson's *Thriller* album was originally reviewed using used words such as 'slick', 'harrowing', 'dark messages', 'tense and obsessive sound'. Connolly, writing in Rolling Stone (1983), commented that the record has 'a deeper, if less visceral, emotional urgency than any of his previous work'. The conclusion was that the words and sounds of the album revealed something about the changing identity of the troubled genius of Jackson pointing to his lack of fit to society and troubled family relationships. Yet such commentaries make observations through vaguer adjectives such as 'tension' and 'obsessive' rather than telling us exactly what it is in the music that communicates these meanings. There has been much less attention too in popular music studies, with some notable exceptions (Tagg, 1982; Middleton , 1996; Walser, 1993) to the way that sounds themselves communicate meanings. While we can of course carry out analyses of lyrics it is how these are delivered, the melodies and sound qualities that carry them, along with the sounds of instruments and drums, which also forms the way they convey moods, ideas and subjectivities. As regards Jackson's 'Billie Jean' the meaning of the words, the story about an obsessed female fan, lies also in the sounds. In this paper, taking a social semiotic approach, the aim is to reveal something of the subjectivity, the emotional world, created by Michael Jackson in his music through a systematic analysis of his use of a number of semiotic resources: restricted pitch ranges,

breathiness, and gentle disjunctive articulation, and instrumentation showing how these point to 'tension', 'obsession', 'troubled relationships' and 'emotional urgency'.

A social semiotic approach is interested in the way that communicators use semiotic resources to achieve particular goals, to communicate specific ideas, attitudes, values and identities. It is interested in exploring on the one hand the resources available to communicators, in other words the repertoire of sign/meaning potentials upon which they can draw in order to communicate, and on the other hand how these are used in specific cases, in particular combinations, to communicate particular meanings. So in the case of Michael Jackson 'Billie Jean', rather than being satisfied with adjectives such as 'troubled', the aim is to describe the features of his music that allow us hear such things. On what repertoire of meaning potentials in term of melodic form, sound quality, vocal and instrumental articulation does Jackson draw upon to communicate meanings like 'tense' and 'obsession'? This social semiotic approach draws in the first place on the work of the linguist Halliday (1978) and Kress (2010) and has been successfully applied to other modes of communication, to images (Kress and Van Leeuwen, 1996), multimedia (Baldry and Thibault, 2006), three dimensional objects (O'Toole, 1994) and more broadly to sound (Van Leeuwen, 1999) and to popular music (Machin, 2010).

To draw out the underlying meaning potentials available to musicians this analysis draws extensively on the work of the musicologists Cooke (1959) who believed that it was possible to describe and inventorise the 'rule-book' of classical music. In other words he thought it possible to reveal the underlying principles of melody and rhythm upon which composers draw in order to communicate things like sadness, outpouring of emotion, entrapment, etc. Tagg (1983) has described the way that in our culture we have come to make associations of particular musical patterns and sounds with emotions, attitudes, settings and events through their repetition in our lives and as they have become embedded in our shared cultural conventions. A composer can therefore rely on a certain combination of notes as being heard by listeners as 'romantic' or 'scary'.

Tagg (1997) discusses the emergence of sounds and music as communicative acts in early human societies in terms of the way they could be used to express the attitudes and ideas associated with certain activities such as initiation rites, marriage ceremonies, harvests and the hunt. So analysis could compare the rhythmic intensity of the kinds of sounds used as members of a group prepare themselves for a hunt as when they wish to send a child to sleep. Tagg states:

> Obviously, the pace required in conjunction with a hunt — intensity of heartbeat, speed of eye, of hands, arms, feet and breathing — will be far greater than that needed for singing a child to sleep (…) In the case of the hunt, quick, sudden movements enacted with the precision of split seconds are vital ingredients of the activity, but they would be detrimental when trying to send a child to sleep (1997:8)

Just as Tagg seeks to identify what more specific elements are present in kinds of musical experience: quickness, suddenness, versus gentle and lingering, so the same kind of analysis can be carried out for popular music songs. But first we need to identify what exactly are some of the basic semiotic resources in music, in a social semiotic sense what are the basic building blocks available?

The analysis also draws specifically on the observation on the meaning of sound quality by Van Leeuwen (1999) who attempted to inventorize a list of speech qualities that comprise the available resources for speakers to create meanings not only through word choice but additionally through voice quality. His observations are combined here with the linguistic theory of Brazil, Coulthard and Johns (1980), McConnell- Ginet (1977) and musical theory of Schaffer (1977), and Tagg (1982, 1984, 1994) in order to look also at the sound qualities of musical instruments. These analytical tools will be explained fully in each section.

The semiotic meaning of sounds

We have seen above that music is often described through adjectives. So we find words like 'tense' and 'obsessive' to describe the sounds of *Thriller*. The problem is that description and evaluation become mixed and the terms themselves are vague. They do not help us to identify and understand exactly what kinds of affordances the sounds are drawing upon. What exactly are the sound qualities that communicate 'tension' and 'obsession'? Drawing on Van Leeuwen (1999) and the cognitive psychologist Arnheim (1969) we can establish two kinds of origins for the meanings of sound qualities. These help to provide us with our first building blocks for identifying the affordances available to musicians which are then developed over the course of this paper.

Provenance

This is simply when a sound comes to have a particular meaning through cultural accumulation of associations. For example, to a Northern European listener pan pipes suggest 'nature' or simple, ancient cultures especially those from Latin America. The sitar is used to represent the whole of Indian culture or the esoteric and mysticism in general. Such associations may have no actual connection to time or place, for example, the bagpipes are associated with Scotland even though historians tell us they were only recently introduced early in the 20th Century. So a pop artist might use a violin to connote folk and tradition or a synthesiser to connote technology and modernity.

The fact that we experience certain musical notes and note sequences as communicating specific emotions too lies in a cultural accumulation of associations. Tagg (1983) was interested in the way that certain music came to be able to represent different kinds of landscape and character in the 19th century. These associations no longer appear anything but natural to listeners due to repetition. The music used in movies for romantic moments, drawing on this cultural history of sound, simply sounds 'romantic' to us when we hear it. However, while provenance explains why certain notes and note combinations have particular meanings it does not permit us to understand the meaning of *how* these notes are

played, the *qualities* of these notes. It is the second of our categories that helps us to think about this.

Experiential meaning potential

The meaning of sound quality may also derive from associations of things in the real world. Arnheim (1969) argued that communication is steeped in 'experiential associations' (p117). He explains that 'human beings are naturally aware of the structural resemblance uniting physical and non physical objects' (p118). So we might clap our hands together to suggest a conflict of interest between two people. There is no actual clapping or physical collision going on in the interaction but communication works by drawing on an experiential association of these to understand something of the way that people may not agree. In the same way the sound associated with crashing objects could be thought to suggest conflict as opposed to a gentle drifting sound that might mean some kind of mutual attraction.

Our physical environment produces noises all the time, all of which have meaning to us. These may be due to certain qualities of the element that makes the sound or even its meaning to us in our lives. Thunder makes a booming sound which may mean violent weather or lightening, things that frighten us or present danger. The sound of thunder also gives the impression of vastness where the whole sky appears to be filled with the noise. This means that such booming sounds therefore when made artificially can be used to communicate something ominous, something powerful or massive. The opposite can be the case for softer, or higher pitched sounds. Tagg (1993) notes that humans tend to use sounds and sound qualities that allow us to draw on associations in the physical world to understand our social and personal experiences. So a smooth sound will be more likely to be used to represent something like romance that a very rough raspy one. The point is, as we will see, that this kind of observation gives us a starting point for looking in greater detail at the sound qualities of voices and musical instruments.

It also appears to be the case that much of the sound qualities in music along with our experience of musical rhythm itself may be linked to our use of language (Levitin, 2006). Our ears and brains are finely tuned not just to listen for the meanings of words and grammar but also to the manner in which these are delivered, to voice quality and to the rhythms in speech. In linguistics much research has documented this phenomenon (McConnell-Ginet, 1977, 1988; Bell and van Leeuwen, 1990). More will be said about both these kinds of origins of meaning in sound as we look closely at 'Billie Jean'.

Social Semiotics of 'Billie Jean'

In the rest of the paper we look at the meaning of voice quality, pitch, pitch ranges, note value and the sound quality of instruments. In each case the meaning potentials will be first drawn out and explained using a number of examples from other well know popular songs and then applied to 'Billie Jean'. It is through this process that we can establish the underlying meaning potentials, the affordances, available to musicians in order to understand Jackson's own specific choices. Also we should bear in mind that in music meaning is created through all of the features working together. For analytical purposes we will deal with each in turn,

accumulating our understanding as we move through them. It is through this process that we will get access to the semiotic resources in 'Billie Jean' that communicate 'tension', 'obsession', 'troubled relationships' and 'emotional urgency'.

Voice Quality on 'Billie Jean'

Drawing on van Leeuwen's (1999) observations on voice quality that draws on a range of work in linguistics we can begin by describing some of the features of Jackson's vocal style on 'Billie Jean' although these will be built upon further throughout the paper. Van Leeuwen lists eight qualities of voice that account for basic sounds found in the sounds made by speakers and singers. Two of these are particularly relevant for 'Billie Jean'.

When we hear a vocalist such as Frank Sinatra, Sid Vicious, Bob Dylan, or Paul Simon it is not difficult for us to produce adjectives to describe the way they sound different. I asked a friend and they suggested that Sinatra has a more 'crooning', 'seductive' style. Vicious is 'aggressive' and 'in your face'. Dylan is 'wistful' and 'earthy' and Simon is 'gentle' and 'thoughtful'. The friend described Jackson's voice as 'excited', 'youthful' and 'innocent'. Van Leeuwen's inventory of voice qualities allow us to describe these differences using careful description as opposed to these vague adjectives that are often tied in with our perceptions of the artists themselves as much as description of sounds.

Tension

This simply describes the extent to which we speak or sing with an open or closed throat. When we become tense in everyday situations our throats tend to close up. This is what we hear when we say that someone's voice sounded tense. Of the singers mentioned above Sinatra and Simon tend to use open relaxed throats whereas Vicious and Dylan tend to use closed throats. Therefore a vocalist like Paul Simon can sound more thoughtful, gentle and relaxed. Dylan can sound troubled and Vicious can sound aggressive. On 'Billie Jean' we can hear that Jackson uses a lot of closed throat to communicate tension. These observations are also useful for thinking about degrees of tension and relaxation in the way that the instruments are played on 'Billie Jean'.

Breathiness

This is simply to do with the degree of intimacy suggested by a voice. To bring out the meaning potential here, we can think of the contexts in which we hear people's breath. This can be when they are out of breath and panting, because of some physical or emotional exertion or strain. It can also be in moments of intimacy and sensuality. When we hear a person's breath when they speak this may even be a moment of confidentiality as they whisper in our ear, or share their thoughts with us when they are experiencing emotional strain or euphoria. When we listen to the vocals of most rock singers we do not hear breathiness or the close texture of their voice. Intimacy is not what is important in such cases. We may however, hear the breath of Paul Simon on one of his more reflective recordings. On 'Billie Jean' Jackson often sings in short breathy gasps. This appears to communicate both intimacy and sensuality along with an element of euphoria. To some extent this may

bring something more of openness about 'Billie Jean' not necessarily communicated in the lyrics themselves. At the level of sound quality there is evidence that is not simply an account of an obsessive fan but something in which Jackson is emotionally implicated and moved.

As regards musical instruments 'breathiness' is also a useful way to think about the way we can hear the sound of a drum striking clearly and sensually or hear gentle sounds of fingers moving across a guitar fretboard or the sound of strings ringing out. This can be contrasted to cases where instruments sounds merge in with the mix.

We will be applying these observations on Jackson's vocals on 'Billie Jean' shortly and later to instrumentation. But first we turn our attention to the way he uses pitch, pitch ranges and particular kinds of note combinations.

The meaning of pitch

Melody is comprised of changes in pitch over different ranges of notes. We begin by looking at the basic meaning of pitch, moving on to how different ranges and kinds of pitch movement can have meaning potential.

Pitch is simply how high or low a sound is. A scream would be a high note, thunder a low note. The meaning of pitch is rich in metaphorical associations. Cooke (1959) has suggested that high pitch means effort, low the opposite; in other words contained, immobile and static (p. 102). We could think of this metaphorically as being like someone speaking in a low deep voice as compared to raising their voice in excitement. But higher pitch can also extend to mean agitation and lower pitch mean low drooping despair. Cooke shows that classical composers have used high pitch to suggest "up and away" due to its energy and low pitch to suggest "closer, down and relaxation" (1959, p. 103). In Western culture we have the association of up meaning 'feeling good' and down meaning 'feeling down'. We transfer this meaning to pitch even though there is no reason why a high pitch should be thought of as 'high' at all as height is not involved. Cooke adds that pitches beyond the range of the human voice can, in the cases of higher tones, give a sense of the ethereal, lightness, transcendence.

We also have associations of the way things in our everyday world produce different pitches of sounds. Heavy objects can make deep sounds when they move or fall. Smaller animals might make higher pitched squeaking sounds. Deep sounds give a sense of danger or something ominous as in thunder. This could be why deep sounds are often used to symbolise gravity or danger. These all provide meaning potentials in terms of the way an artist might use a low pitch, like Tom Waits, or a high pitch, like Michael Jackson. One can seem closer to drooping despair, heavy with gravity or menace, immobile, while the other can appear much less robust, agitated and high energy.

More will be said specifically about this when we move on to the meaning of movement in pitch. In music, after all, pitches rarely stay at one level. We will also later be able to comment on the pitch used in instrumentation.

The meaning of ascending and descending in pitch

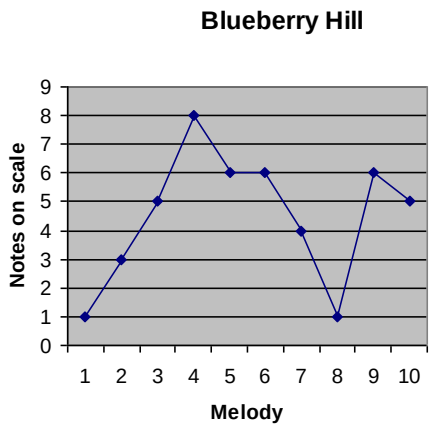
We have listed a number of meaning potentials for high and low pitches. But often when we hear sounds and music there is generally movement in pitch. A vocal line will usually rise and fall in pitch. The direction and extent of this movement can also have meaning potential.

Cooke (1959) suggests that in classical music ascending melodies are associated with outward expressions of emotions whereas descending melodies are associated with incoming emotion. This is due to the association of higher pitches with higher levels of energy and brightness and lower pitches with associations of low levels of energy. The movement from one to the other expresses a shift in either direction. A movement from a high pitch to a low pitch the meaning is of a falling of energy. The opposite, a gradual slide from low to high pitch, gives a sense of a picking up of spirits. National anthems use stepped increases in pitch to suggest the brightness and energy of the national spirit. Anthems will often also use some lower pitches to suggest the weight and gravity of the nation.

We can show how increase and decrease in pitch works in a familiar popular song to help us to establish the available meaning potentials draw upon by Jackson on 'Billie Jean'. Here is the melody from the start of "Blueberry Hill". First I show the direction of the movement of notes under the lyrics and then represent this as a graph. On the graph the scale on the right represents the notes going up and down.

I found my thri-ill on Blue-ber-ry Hill

↑ ↑ ↑ ↓ → ↓ ↓ ↑ ↓



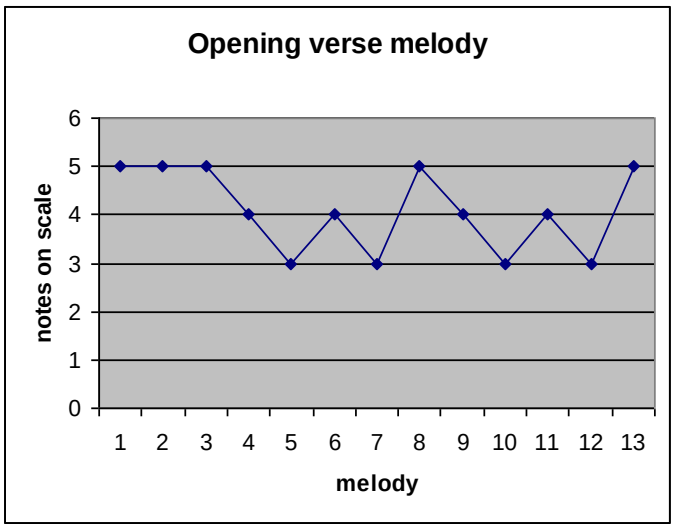
The arrows positioned under the lyrics indicate whether the melody of the song, the notes, are ascending in pitch, descending or staying at the same level. This can be seen clearly on the chart. In this case the first four words of the song ascend. If you hum the tune yourself you will be able to hear this. Therefore we can say that the singer here is expressing an outgoing emotion or an increase in optimism which is done with an open throat communicating no tension. The melody then levels off and descends back to its starting point, although there is an important ascending note before this on 'berry'. So the outward expression of emotion, of

joy, is followed by a move back to something more grounded and relaxed, perhaps more thoughtful. But the other ascending note on the second syllable of 'berry' marks another burst of emotion before it tidily resolves back to its starting point, giving a sense of closure. If the melody descended from the start then, according to Cooke, there would be incoming emotion, such as a received sense of joy or consolation or even a simple slide to bleaker thoughts and self-absorption. So in terms of pitch 'Blueberry Hill' is an outburst of positive emotion.

While more needs to be said about pitch before we have a comprehensive set of tools with which to consider 'Billie Jean' we can look at the pitch movement in the first line of the song which is typical of the whole song.

→ → ↓ ↓ ↑ ↓ ↑ ↓ ↓ ↑ ↓ ↑

She was more like a beau- ty queen from a mov- ie scene



We can see that Jackson tends to not use longer consistent increases or decreases in pitch but jerkier pitch movements. There are therefore no clear outbursts of emotional energy nor falling off of energy. This gives a sense of emotional instability, and since there is often tension in his throat as he sings these words along with breathy gasps we get a sense of a mixture of agitation and as we will see, also a kind of euphoria. The meaning of this can be drawn out more when we consider next the quality which is pitch range.

What Is the Range in Pitch Between the Highest and Lowest Note?

As well as whether pitch increases or decreases there is important meaning potential in the range of these changes. A large pitch range means letting more energy out whereas a small pitch range means holding more energy in. So we might think of a singer like Bob Dylan holding energy in, while a singer like Freddie Mercury lets it out as on songs like 'Find me Somebody to Love'. Linguists Brazil, Coulthard, & Johns (1980) argue that wider pitch range in speech is akin to excitement, surprise, and anger. Narrow pitch range associated with

boredom and misery. Pitch range in speech is also associated with emotional expressiveness. In Anglo-American societies, men have less pitch range than women (McConnell-Ginet, 1977, 1988). So small pitch ranges can be associated with holding in, or even modesty. In Bob Dylan's singing about tragedy or injustice, using a limited pitch can give a sense of resignation or contained pain, perhaps giving it a different kind of gravity than an opera singer using a large range to sing about the same subjects. Before we can think about how this applies to 'Billie Jean' we need first to look at how we can examine pitch range. We can most easily show this if we return to our example of 'Blueberry Hill'. Here we look at the way the melody uses notes from a scale of 8 notes in order to show pitch range.

Songs normally draw on a scale which is comprised of 8 notes. Note 1 gives the name of the scale and is called the root note and the 8 notes after this rise in pitch. Since there are 8 notes in the scale, note 8 is the same as the root but an octave higher. We can see this on the illustration of the piano key board below.

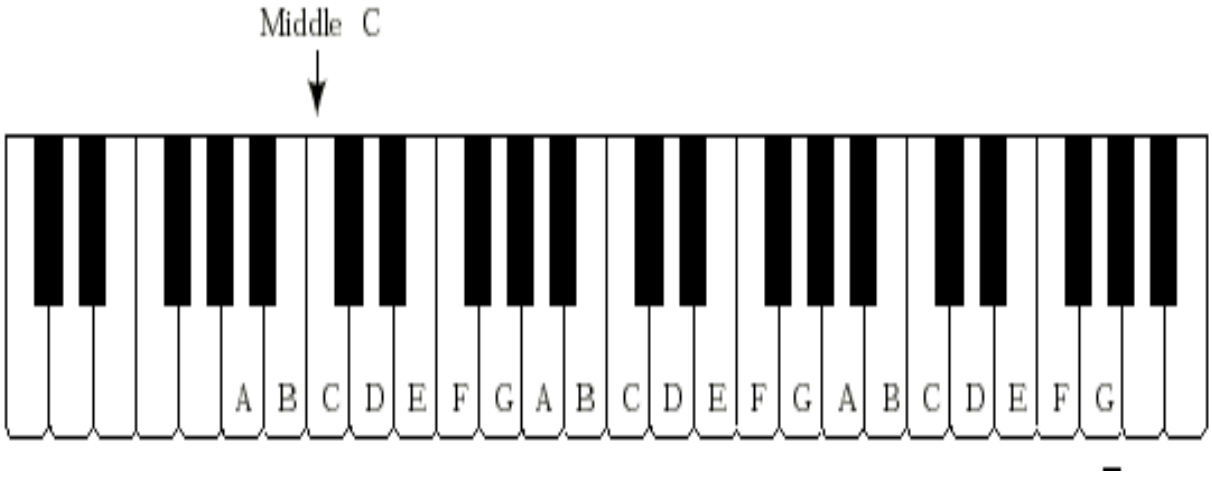


Fig. 1: 8 notes rising in pitch on piano keyboard

Middle C is note 1. If we move up 8 notes to the right we again come to C. If we move 8 notes down 8 notes to the left we find the same sequence of notes repeated.

Below the notes in 'Blueberry Hill' are represented with numbers to indicate where they are in the 8 notes of the scale. In this case the higher the number, the higher the pitch of the note, although if we see a 1 followed by a down ward arrow and then a 7 this would mean that the melody had descended to the 7th note below.

I found my thr- ill on Blue-ber---ry Hill

1 ↑ 3 ↑ 5 ↑ 8 ↓ 6 → 6 ↓ 4 ↓ 1 ↑ 6 ↓ 5

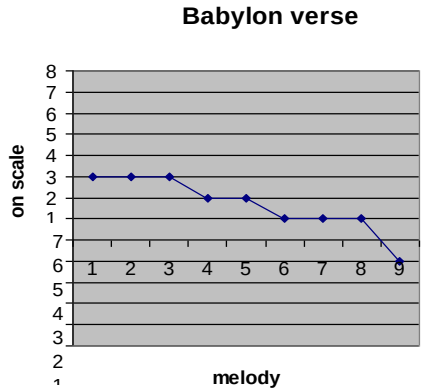
What we find in 'Blueberry Hill' is a large pitch raise in the first phrase 'I found my thrill'. Here there is a leap of 8 notes. This is therefore an extensive release of energy and brightness. Following Cooke (1959) we can say it is large outward expression of energy. This interpretation suits the lyrics, which seem to be highly positive. However, if the notes

had descended or been over a limited range, it would have resulted in a very different expression of those lyrics. So here the melodic form communicates a celebration of the event of finding the thrill; a descending pitch would perhaps lend a mourning, pining quality to the song a smaller pitch range something more emotionally restricted.

Below is an example of a song that uses a descending melody. In "Babylon" by David Gray we have an example where each line of the verse descends. Here is one of the lines:

Fri--day night and I'm go—ing no where

3 → 3 → 3↓ 2 ↓ 2 ↓ 1 → 1 → 1 ↓ 6



Here the melody begins on the 3rd note and then descends in steps as can be seen on the chart. This suggests the opposite of outgoing emotion. This can give a sense of not particularly attempting to reach out to communicate as well as a falling of mood. We can imagine that if the same lyrics were sung with an ascending melody it would seem as if he were pleased he was going nowhere.

Some melodies neither ascend nor descend but are very static. One example of this would be 'Anarchy in the UK' by the Sex Pistols. Much of this melody takes place with repetition of the first note. There is therefore very little outward giving of emotion or positive energy. Nor is there any falling away of energy. This means that there is something very contained, restricted or confident about the way it is sung. In fact the vocalist sings the song generally at a high pitch which conveys an emotional intensity. Yet in this intensity there is no emotional outpouring or pleasure as in "Blueberry Hill". There are only short sharp occasional outbursts to the fourth note towards the end of each line. .

For 'Billie Jean' we find much stasis in the melody with repeated notes and smaller pitch movements. There is a combination of emotional containment or entrapment and jerky melodic movements. We can see this in the case of the first verse:

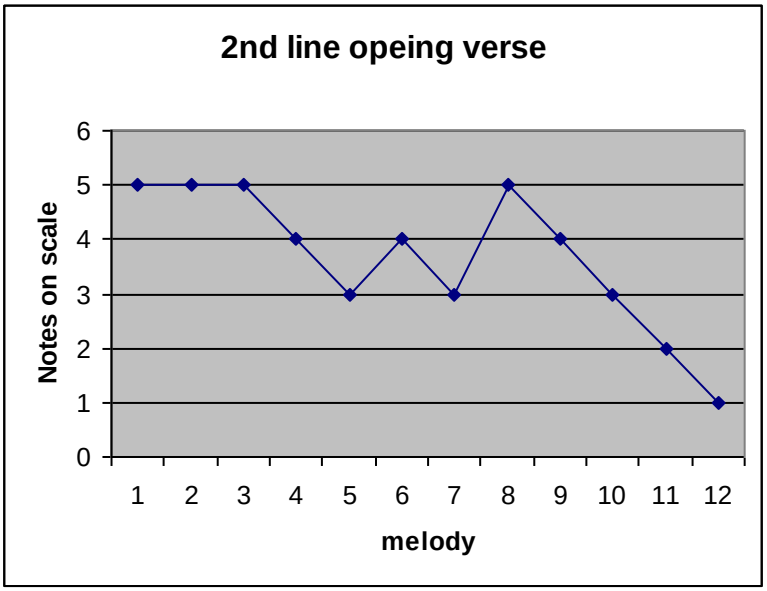
She was more like a beau- ty queen from a mov- ie scene

5 → 5 → 5 ↓ 4 ↓ 3m ↑ 4 ↓ 3m ↑ 5 ↓ 4 ↓ 3m ↑4 ↓ 3m 5 →

I said don't mind but what do you mean I am the one

5 → 5 → 5 → 5 ↓ 4 ↓ 3m↑ 4 ↓3m ↑5 ↓ 4 ↓ 3m ↓2 ↓ 1

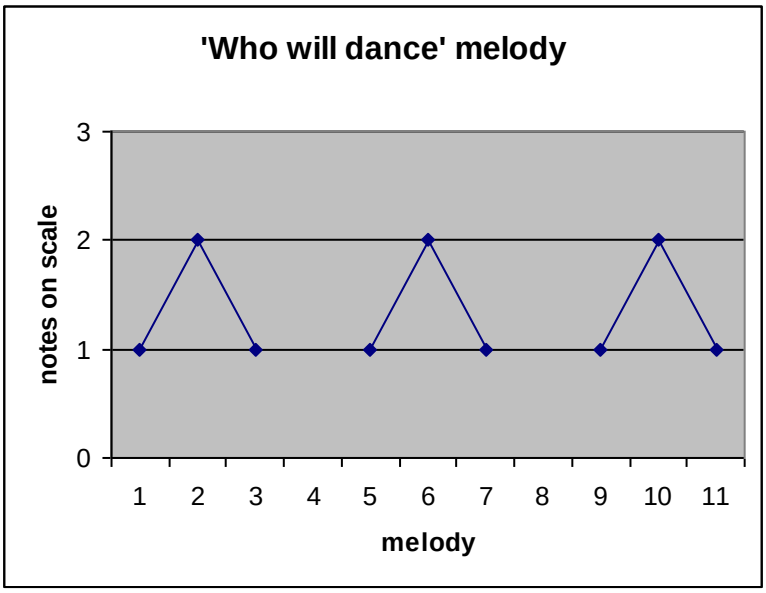
We have already seen the graph for the 1st line above. Here is the graph for the second line:



Much of this verse is characterised by jumpy melodies, with sections of stasis. It is only on the words 'mean, I am the one' where there is a clear descending pattern. Overall these jerky melodies sung at a high pitch give a sense of excitement and high energy, of lightness, and perhaps fragility, yet without clear direction. The jerkiness brings a sense of uneasiness of not staying still. This is also communicated through the sharp breaths and tension in the voice. We can see the same on the lines which follow the verse which start 'who will dance..'.

Who will dance on the floor in the round

1 ↑ 2 ↓ 1 → 1 ↑2 ↓1 → 1 ↑2 ↓ 1



These are sung as bursts of energy, but with no emotional expansion. This use of short bursts in what we can call 'disjunctive articulation' has important meaning potential. Some melodies will be sung in a way such that words are joined together in longer phrases as we find for 'Blueberry Hill' where there is less a sense of an exclamation as an emotional lingering. There are long extended rises and falls in pitch, often with extended gentle decreases in pitch over a word. 'Billie Jean' is different, and it is here that we get a further hint at the tension and obsession in the sounds along with the more fragile and almost anxious higher pitches.

Bell and Van Leeuwen (1993) have noted that these kinds of shorter phrases are associated linguistically with sincerity, certainly, weight and therefore with authority. We can imagine the effect if someone were to speak to us in sentences comprised of bursts words. This is common to the way that news readers speak, for example. In the case of news readers it can also connote the urgency and immediacy of news. We might therefore be less surprised to hear folk singers using such short bursts to communicate sincerity. The opposite case where singers produce longer lingering statements suggests rather slow burning internal emotion.

In the case of 'Billie Jean' we find that Jackson uses these shorter breathy bursts but combines them with high pitches and tension. We can imagine the effect if a news reader combined the shorter phrases with such higher pitches. Simply this would connote a level of sincerity and weight, but at a level of emotional involvement and fragility inappropriate for news reading.

The meaning of using different notes of the scale

This is the part that the non-musician may find slightly more off-putting. But it is very simple and technical terms are avoided. As pointed out above there are eight notes in a scale

and each, as Cooke (1959) has shown, has its own meaning potential. Jackson's note choices for 'Billie Jean' are one important way that 'trouble' and 'harrowing' are put into the music.

A pop song is normally based on one scale in which there are 8 notes. So in a scale of C there are eight notes following from note one, C. The eighth note, the octave, marks a return to C. In any song both the notes in the melody and the notes in the chords of the accompaniment are drawn from these notes. In between these eight notes are other notes that do not belong to the scale, but to other scales. These can be used in songs to add more drama and trouble. For none musicians it is worth pointing out that this division of notes in this way, into eight notes with specific identities, is completely arbitrary and can be explained by certain cultural and historical developments. Nevertheless due to repetition these now sound natural and inevitable to our ears (Tagg, 1983).

To make up a melody any of the eight notes can be used. But each of the eight notes has a different kind of sound which in turn has a different kind of effect for the listener. Because they have different effects, certain notes and certain note combinations are usually used. Tagg suggests that since we have been hearing these combinations and making certain associations for a long period in our culture, we can now easily recognise what is being communicated, what mood, what idea, what emotion.

In melodies certain notes of the eight in any scale are used a lot as they create a solid connection to the musical accompaniment, which will also draw on the same notes. These commonly used notes are mainly notes 1 and 5. Note 1 is the main defining note of the scale. In the scale of C therefore, note 1 is C. So this anchors the melody to the scale firmly and roundly.

Note 5 is similar in sound to note 1 and therefore is also good for anchoring the melody to the scale. Also important is note 3. These structures using notes 1, 3 and 5, have become the basis of western music.

Notes that anchor the melody to the scale and to the accompaniment allow the music to feel 'easy' or 'rounded'. In contrast jazz will use many notes that do not create this solid connection in order to create tension. Blues music often uses the occasional difficult, or "blue," note to create and release tension. I have heard music of boy bands which have at no point deviated from the 1st, 3rd, such is the importance of creating an 'easy' sound.

If we look at the melody of 'Blueberry Hill' above we can now think about exactly which notes are used in the melody. The first four notes are 1, 3, 5, 8. Since the 8th note is the same as note 1 we can say the first four notes are 1, 3, 5 1. Such a combination will anchor the melody, the song, to the music underneath very roundly. There is no complexity, trouble or doubt in what the vocalist expresses. In the case of 'Billie Jean' we can see on the opening verse that there is also the extensive use of the 5th and 3rd notes. Throughout the song these are important along with the 1st note. At this level we might say therefore that the song is highly grounded.

Interestingly the verse of 'Billie Jean' begins not on the 1st note as do many popular songs but on the 5th, only resolving to the 1st note at the end of the line on the word 'one'. This is interesting for two reasons. First, using the 5th in this way is a technique often used in jazz. An instrument will play a 5th in place of a 1st to give a sense of space, distance or very slight tension in the music. So Jackson begins telling the story with this kind of mood, of slight tension and distance. Second, it is notable that the 1st note, the grounding note, is first used in the melody only when Jackson sings the words 'I am the one'. This seems to make him only truly grounded at this point through being the 'one'. And it is here that he sings out the words with much less jerkiness, less breathiness and with an open throat. In Cooke's model this would suggest he may therefore have an ambiguous relationship with this situation. At first listening the lyrics appear to be a warning about fanatical fans. But melodically there is a suggestion that it may be through this process of being the 'one', the target of the obsession, which he truly finds himself, where he may have less gasping, nervous tension. As we have already seen the melody appears to indicate high levels not just of sincerity but also of instability and excitement. It seems that it is not just a warning story but also a view into Jackson's subjectivity.

The 3rd note is important for other reasons. It is the way that the 3rd and other notes are used that can bring much more complexity of emotion and feel to a melody. You might be aware that, in general, a minor key is sad and a major key is happy. Research suggests that major notes and chords are associated with positive feelings. A minor key is created by lowering some of the notes in the eight-note scale; the 3rd and the 7th notes. These are lowered to notes in between the notes of the scale. So they become a bit like a 2 and a half and a 6 and a half. If a melody has the standard 3rd and 7th notes then it is a major melody and is therefore happy and joyful. If it has the lowered 3rd and perhaps a lowered 7th, it is sad.

Looking at the example of 'Blueberry Hill' above we can see that the 3rd note used for 'found' is a standard 3rd, also known as a major note. The song goes straight to this defining happy note from the 1st note. Therefore it is happy. Since it is an ascending melody we can comfortably say that we have an outward expression of joy and brightness. In contrast had this note been a minor 3rd the word 'found' would have sounded very sad and regretful. If you listen to the opening line of 'Ain't No Sunshine', you can here this kind of sadness. This song opens, like 'Blueberry Hill' with an ascending melody but uses the minor rather than major 3rd. The pain of this song therefore lies in the melody as well as the lyrics. Of course there is nothing 'painful' about the minor 3rd in itself. It is just an association that has become established through semiotic provenance.

The 4th is a note that Cooke suggests is associated with building or moving forwards. It can also give a sense therefore of space and possibility. In the case of the Sex Pistol's 'Anarchy in the UK' the verse melody is comprised of the 1st note which repeats for most of the verse and the 4th note. So while the melody is dominated by stasis, which we suggested in this case could mean confidence or at least contented lack of interest, there is a repeated 4th to suggest some building. The meaning would have been very different had the first note always moved towards a minor 3rd or minor 7th.

The 2nd note has been associated with transition by the suggestion of movement, or the promise of something to follow, or lengthened 2nd notes can suggest limbo or entrapment. This is because of its position between the strongly related 1st and 3rd. 'Aint No Sunshine' uses a 2nd note for the word 'when' for the words 'when she's gone'. The lingering on this note suggests a sense of limbo.

The sixth note is very much like the major 3rd and can therefore be used to ground the melody. On 'Blueberry Hill' we find much use of the 1st, 3rd and 5th which give the melody a grounded solid feel and also the 6th which has the same effect. The only note that does not have this grounding effect in this melody is the 4th used for the word 'Blue' which suggests building.

If we look at the verse and chorus melodies on 'Billie Jean' we find an extensive use of sad and painful minor notes along with 4th notes. We can begin with the opening lines which were illustrated above.

The opening vocal line is a confined melody mainly based around the limited pitch range of the 5, 4 and 3m. There is a combination of stasis through repetition of the same notes several times and of shorter fluctuations in melody which, in contrasts to songs like 'Bluebery Hill' does not suggest steady outpouring or receiving of emotion. This fluctuation is characterised both in the melody and in the high pitched disjunctive articulation of how they are delivered by Jackson in tense, breathy bursts rather than more legato, smoother articulation as we find on 'Blueberry Hill. What is communicated throughout the 'Billie Jean' melody is emotional instability with some stasis, communicated with a level of sincerity, tension and high excitement.

The opening verse is anchored through the 5th note but since this is used from the start this brings a slightly floating effect. The 1st note is used only at the end of the line for the word 'one'. So it is only here where it is truly grounded and resolved. It first passes through the 3m note on the word 'am' indicating some painful about this fact. We find an extensive use of the building 4th note in the melody. In combination with the extensive use of 3m notes this creates a building sense of sadness and pain. And this raise to the painful note along with the high pitch and tense throat gives a sense of euphoria. While this is painful Jackson appears to be highly stimulated by this fact.

The linking section 'Who will dance, on the floor, in the round is comprised of small bursts of breathy, sensual energy using the 1st and 2nd note. This again creates a jerky, nervous effect. The 2nd note is appropriately used for the asking of a question.

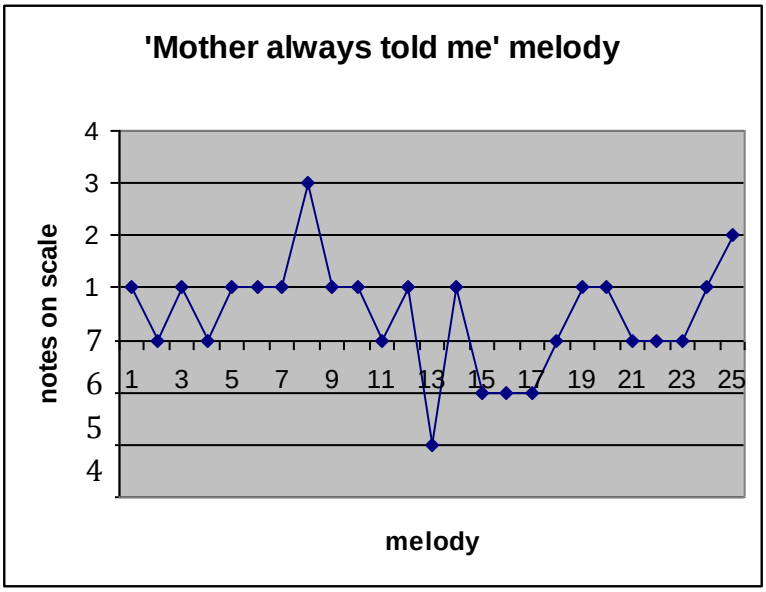
For the next section which begins 'People always told me' we find an increase in pitch of a whole 8 notes. This suggests increased energy.

7m ↑ 1 ↓ 7m ↑ ↓ 7m ↑1 → 1 →1 ↑ 3m ↓ 1 → 1 → 1 ↓ 7m ↑ 1 ↓

And mo-ther al-ways told me be care-ful of who you love.

5 ↑ 1 ↓ 6 →6 → 6 → 6 ↑ 7m 1 →1→ 1 ↓ 7 → 7 → 7 ↑ 1 ↑ 2

And be careful of what you do 'cos the lie be comes the truth



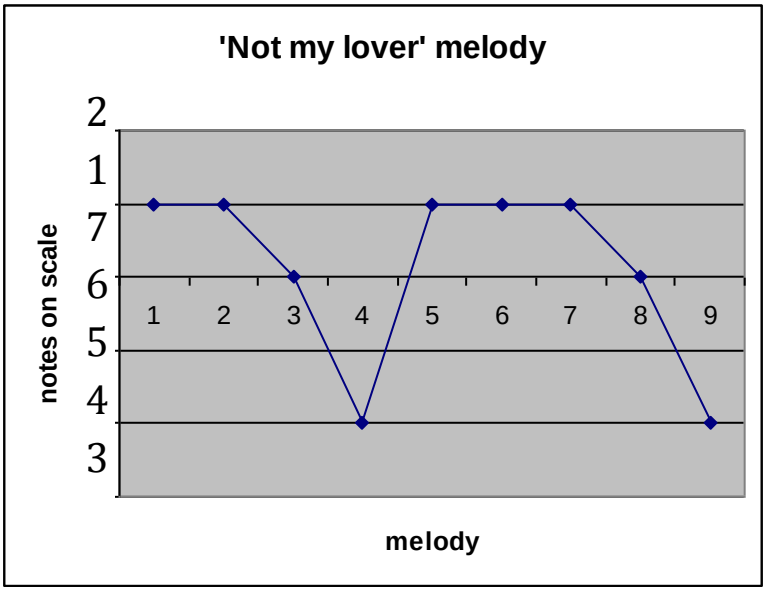
The first section is comprised mainly of stasis with movement between the 1st note and the painful 7m. There is one burst of energy leaping to the sad 3m note on the word 'Careful' and then a return to the painful stasis. The second section has much more emotional range. It starts by falling down to a grounding 5th note which brings a sense of gravity for 'And be careful' and then stays on the relatively positive 6th note, suggesting this is good advice. It then returns to more stasis over the 1st and the 7th notes again but this time interestingly not using the 7th minor but the natural 7th note which jars a little alongside the minor key. So the words 'lie becomes the truth' sound dischordant and uncomfortable. The line ends on the 2nd note which gives a sense of something to follow, as if this was the prophecy that was fulfilled.

Overall here we have a surge in pitch and energy levels yet a continued use of stasis and emotional confinement and more jittery, nervous melodic pitch changes. There is a dominance of painful minor notes and the uncomfortable 7th used to express the nature of lies becoming truth. In the lyrics too it is unclear of what has actually taken place and who has seduced who.

For the chorus we find less erratic melodic movement.

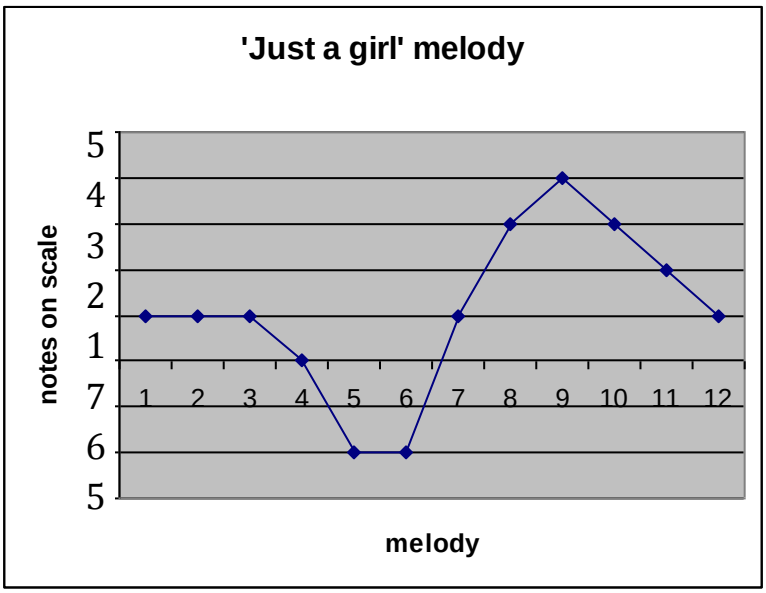
1 → 1 ↓7m↓5 ↑ 1→1→1↓ 7m ↓5

Bi-llie Je- an is not my lov- er



1→ 1 → 1↓ 7m↓ 5→ 5 ↑ 1 ↑ 3m↑ 4↓ 3m↓ 2↓ 1

She's just a gi- rl who claims that I am the one



The chorus line starts with a falling away of pitch sinking straight to the painful 7m note and then to the 5th which repeats. This suggests a falling away of energy, a feeling of regret. Here the emotions are much more identifiable that for the verse and there is smooth legato articulation rather than the nervous bursts of phrasing for the verse. But what it interesting is

the large surge in pitch for the lines 'claims that I am the one'. This uses a 3m which makes it sad and also the 4th note which suggests something building and possible ominous, but also communicates a sensed of building excitement again with the high pitch. But this then resolves back to the 1st for 'I am the one'

We then have the following lines that repeat throughout the song:

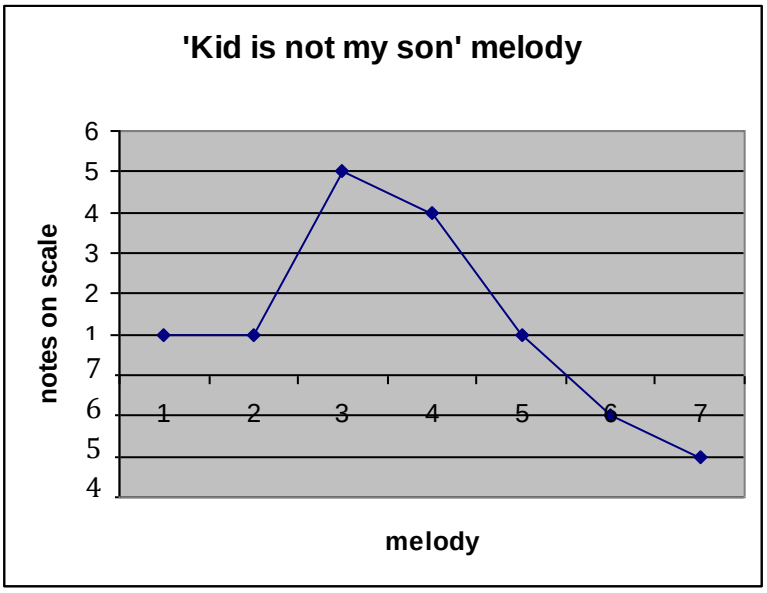
1 ↑ 3m↑ 4 ↓ 3m↓ 2 ↓ 1

She said I am the one

This line goes straight form the 1st note to the 3m which for Cooke would suggest heavy outward expression of pain. It then goes to the building 4th and descends back through to the 1st note. Since the resolution to the 1st note goes through the 3m from the building 4th we can see the menace in this accepted state of affairs suggested by the resolution. Here it is the 4th 3m 2 1 combination that bring the beautiful sound.

1 → 1 ↑ 5 ↓4 ↓ 1 ↓ 6↓ 5

But the kid is not my son



I find this a particularly beautiful section. Here we have a much more grounded melody, the most 'easy' and 'unproblematic' of the whole song with the use of the 1st and 5th notes and the 6th note which sounds very much like the happy major 3rd note. And we find the resolution here dipping lower in pitch than the 1st note as if to give weight to what he says or possibly simply a sense of resolution. There is also the use of the 4th which here suggests moving forwards. And importantly on the chorus we have less breathiness and more openness to the throat in the vocals. The chorus therefore shifts away from the higher levels of nervousness expressed in other sections of the song.

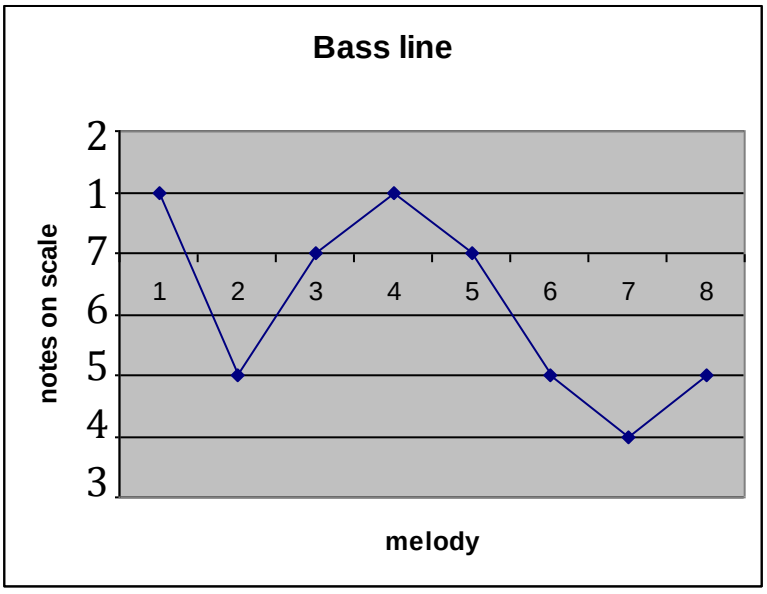
Pitch and sound quality on instrumentation

Overall in terms of vocal melody and voice quality we find mainly nervous agitation and euphoria with limited pitch ranges suggesting emotional confinement but at the same time lack of stability within this. What is important to now consider is the way that this vocal line fits with the instrumental sounds. In this section we make a number of observations on the bass line and the synthesizer stabs using the observations on melody and sound quality developed over the preceding sections.

Bass Line

The bass line, to use adjectives in the first place, is a riff that cycles and gives a sense of relentlessness. Here is the melody which is repeated, cycling, throughout most of the song. It appears after a short introduction where we hear a synthetic dance drum beat:

1↓ 5 ↑7m↑ 1↓ 7m↓ 5 ↓4 ↑5

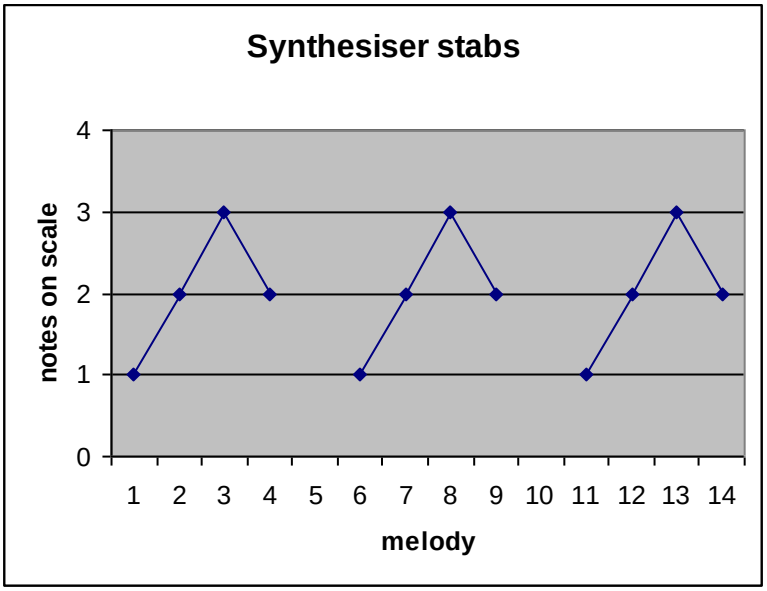


This is a grounded melody using lots of 1 and 5 notes. That the melody takes place in pitches lower than the starting 1st note brings a sense of gravity through the metaphorical meaning of lower pitches. For mood it uses no 3rds but uses the painful 7m. There is also the use of the 4th note at the lowest point in pitch to create a sense of building menace. The pitch range is not extensive and takes place over 4 notes. It is not therefore particularly emotionally expansive. The actual sound quality of the bass notes are relatively softly played and do not ring out. This relentlessly repeating cyclic bass riff with the lower 4th note appears to give a creeping but relentless movement to the song, always moving forward and never staying still on one note and which remains just in the background pushing the song forwards. It is over this forwards movement that the vocal melody places its relative stasis and emotional instability at high levels of excitement and energy. It is as if Jackson is carried along by this momentum, gasping and agitated.

Both Tagg (2001) and Schaffer (1977) have discussed the relationship between melodies and backing instruments and argued that certainly in popular music lead vocals or instruments either merge with or rise above the sounds and rhythms of the backing as a representation of the individual's place in society. In rock music guitars and drums are played at high volume and then the singer screams out above them when with amplificiation technology this is in fact unnecessary. The authors suggest this reflects a sense of the individual's place in the anonymous post industrial society where we need to shout to be heard, although even so our voices are only just heard. In this sense we might say that this bass line of 'Billie Jean' represents the social context in which Jackson finds himself – one that does not drown him out as is suggested by the backing of rock music, but one that moves him along relentlessly and with slightly darker undertones, but nevertheless with a stylish energy. .

Synthesiser stabs

The synthesiser stabs are played once the momentum of the bass line is established. These stabs are basically three chords that repeat returning to chord one. Chords are basically notes taken from the scale, usually the 1, 3, 5 and 7. These can be played in any order of pitch. The way the chords are played on 'Billie Jean' allows for a melodic sound to come from the sequence where the highest notes sound. This makes for a melody of notes 1, 2, 3m, 2 . It is very compact and limited in pitch range suggesting a kind of emotional confinement. As with the bass line there is a sense of repetition here, of entrapment or of relentlessness.



Melody notes on synthesiser stabs

What is of great importance is the way that these notes are played. The synthesiser chords are not played in a way that allows the notes to ring out which may have suggested an emotional release. They are played with a degree of tension that suggests hesitation. The metaphorical meaning potential here relates to tension versus relaxation in voice quality as discussed above. These stabs over a very confined pitch range are placed against the relentless bass line. The

use of the 3m note provides them with a degree of pain. So there is a painful hesitation or warning. In the model provided by Schaffer (1977) we might argue that this represents another layer in the social world occupied by Jackson. There is the relentlessness and menace of the bass line and the hesitant slightly tense, painful emotionally confined synthesiser stabs. Does this mean that Jackson finds himself in a world almost out of his control, where he is a passenger? Here the synthesiser stabs might communicate the soft yet creeping nature of the threats that face him or even himself attempting to negotiate them. It is over these that he places the melody and sound qualities of the vocals used to tell the story. It is over this relentless bass line and hesitant confined synthesiser that he places his nervous, energetic, sincere and sensuous story.

In sum the lyrics of 'Billie Jean' tell of an obsessed fan who appears to claim that Jackson is the father of her child which he says calmly is not the case. He is warned by people and his mother to take care of such situations, with some rather sad moments where his mother tells him to be careful of who he loves and lies can become truth, yet it appears he has been drawn in to some extent and it is unclear to what extent he is an agent in the process. The instruments provide a relentless grounded yet slightly menacing bass pulse along with a hesitant tense synthesiser over which he gasps and shifts around agitated and breathy in small jerky pitch shifts with lots of painful notes and also building notes. The story is told at higher pitches through tense throat communicating high energy, agitation, excitement and anxiety and at other times a breathy intimacy and sensuality or even arousal. Only in the chorus do we find more calm articulation and clear emotional direction. Moments of high energy and grounding appear to take placed when he tells us that he is 'the one'. Through these combinations we certainly gain a sense of 'tension', of 'obsession' of 'trouble' and of 'emotional urgency'.

Conclusion

Unlike visual and textual stimuli sound is something which enters our bodies. We can literally resonate with a thumping bass line or riff of an electric guitar. Sound can pass into and through our bodies and become part of us. We don't simply observe it at a distance as we can a picture on a screen or words on a page. In this way sound has a more immersive quality than other kinds of perception, than other acts of human communication. Sound and music can feel like they are a part of us, inside us. In this sense sound and music form a very special kind of set of semiotic resources that have this power to literally enter us and move us. In this paper I have attempted to take a step in exploring the way that Michael Jackson uses sound alongside lyrics on 'Billie Jean'. If we wish to understand the way that musicians communicate subjectivities then we need especially to understand how they communicate through sound as while through lyrics they can certainly communicate ideas these enter our bodies on a different level allowing us to resonate with them. If we wish to understand why we are so moved by certain musicians we should think more closely about the exact semiotic choices they make. We like to understand he linguistic and stylistic patters in literature which can mean analysing grammar and word choice. In the same way a semiotic analysis of music can help us to appreciate how musicians use the semiotic resources in sound in the same kind of special, creative way.

References

Arnheim, R. (1969) *Visual Thinking*, CA: University of California Press.

Baldry, A. and Thibault, P.J. (2006) *Multimodal Transcription and Text Analysis*, London, Equinox

Bell, P. and Van Leeuwen, T. (1993) *The Media Interview*, Kensington, University New South Wales Press

Brazil, D., Coulthard, M. and Johns, C. (1980) *Discourse Intonation and Language Teaching*, London, Longman Higher Education.

Cooke, D. (1959) *Language of Music*, Clarendon paperbacks

Cook, N. (1990) *Music, Imagination and Culture*, Oxford, Clarendon Press

Frith, S. (1996). *Performing Rites: On the Value of Popular Music*, Harvard University. Press, Cambridge, Mass

Halliday, M.A.K. (1978) *Language as Social Semiotic: The Social Interpretation of Language and Meaning*, Baltimore, University Park Press

Kress, G. and van Leeuwen, T. (1996) *Reading Images: The Grammar of Visual Design*, London, Routledge

Kress, G. (2010) *Multimodality*, London, Routledge

Levitin, D. (2006) *Your Brain on Music*, London, Atlantic Books

Machin, D. (2010) *Analysing Popular Music*, London Routledge.

McConnell- Ginet, S. (1977) 'On Stress and Linguistic Rhythm', *Linguistic Inquiry* 8, 249-336

Middleton, R (1983) `Play it again, Sam': some notes on the productivity of repetition in popular music'. *Popular Music*, 3: 215-234 (Cambridge).

O'Toole, M. (1994) *The Language of Displayed Art*, Rutherford, Fairleigh Dickinson Press

Schafer, R.M. (1977) *The Tuning of the World*, Toronto, McClelland and Stewart,

Tagg, P. (1982) Nature as a Musical Mood Category, Nordens working paper series. http://www.tagg.org/articles/xpdfs/nature.pdf

Tagg, P. (1984) 'Understanding Musical Time Sense' in *Tvarspel - Festskrift for Jan Ling (50 år)*, Göteborg, Skriften fran Musikvetenskapliga Institutionen http://www.tagg.org/articles/xpdfs/timesens.pdf

Tagg, P. (1994) From Refrain to Rave: The decline of the figure and the rise of the ground, *Popular Music* 13/2, 209-222

Walser,R. (1993) *Running with the Devil: Power, Gender, and Madness in Heavy Metal Music* Hanover, N.H.: Wesleyan University Press

FAST FORWARD?
Reflections on fashion and modernity[4]

Alexandra Verschueren,

University of Antwerp & Antwerp Royal Academy of Fine Arts
Antwerp, Belgium

Abstract

This article starts from the striking obsession of the fashion world with innovation and change. Then a systematic attempt is made to relate this feature to the notions of modernity, modernism, and post-modernity. An exploration of these notions in relation to each other leads to a description of the position of fashion in the economy, in a world imbued with social status and politics, in the sphere of individuality and individual needs, and in the context of art and culture. A complex picture emerges in which the artistic and practical problems faced by a designer can be summarized in the following question: what can be made how and where in a sufficiently creative and innovative way, appealing to people's ambivalent needs for identity, conformity, and individual self-expression, and producible in an affordable and ethical manner?

Keywords: *Fashion; Design; Modernity: Modernism; Postmodernity*

An obsession with the future

The January 2011 issue of *Harper's Bazaar* was at first sight unremarkable, meeting any fashion magazine reader's basic expectations with cover lines such as *"HOW TO* DRESS *NOW*" and "MODERN *NEW* MOOD". Skipping quickly to the *Style* section (pp. 39-60), however, we see that editor Clare Coulson gave it a very specific label, *The new MODERNITY special*, which is repeated in a form resembling a quality sticker on most pages. The section is introduced as follows:

[4] The research for this article was carried out under the academic supervision of Kurt Vanhoutte and with the financial support of the Flemish Fund for Scientific Research. Needless to say that I am greatly indebted for both.

> At the start of a brand-new year, *Bazaar* presents a future-fashion special. From advanced fabrics and digital wardrobe services to virtual catwalk shows and the latest apps and websites – plus the hottest new trends and designers – we take a glimpse at stylish things to come.

The superlatives of novelty are hard to miss: brand-new, advanced, latest, hottest, new. And as if that were not enough, there is the promise of getting a look at what is not yet there: future fashion, things to come. This promise is given substance with reference to 'Calvin Klein's *futuristic minimalism*', a feature on 'The *coming* storm' (new British designers "at the forefront of future trends"), and an article – 'Fashion *forward'* – on the 'brave new world' of new technologies for marketing and buying which everyone will have to embrace in order not to be "stuck in the Iron Age". This is all summed up as "Driven by a combination of smart technology and cool common sense, fashion is undergoing a seachange".

Around the same time, the Barbican Art Gallery in London was exhibiting a retrospective of thirty years of Japanese fashion design with the title 'Future Beauty' (15 October 2010 to 6 February 2011). Thus in the world of fashion even things past can be felt to be sufficiently modern to be related to the future. This may be simply a matter of shifting oneself to a past perspective from which one can then look ahead. But such a conceptual maneuver may even emphasize more than the market-driven and market-driving language of *Harper's Bazaar* the need for fashion to present itself as really new and reaching into the future.

Figure 1: *Harper´s* modernity 'sticker' and the Barbican´s exhibition flier

Such anecdotal facts certainly reveal fashion's somewhat uneasy, because compulsive, relationship with time or, more specifically, with the future and 'modernity'. The uneasiness follows from the fact that the future can never be reached or, as soon as it is reached, it becomes past. As Wilson (2003: vii) puts it, "the 'now' of fashion is nostalgia in the making". Much has been written about fashion and modernity.[5] But what still seems to be lacking is a compact but explanatory way of situating fashion in the landscape of notions that includes modern, modernity, modernism, but also late modern, postmodern, postmodernism and postmodernity. A modest attempt is made here to fill that gap.

Modern, modernity, modernism

It is commonplace to say that the concepts just listed do not have the same meaning for everyone. The first task is, therefore, to trace the main differences that may be relevant for situating a specific cultural phenomenon such as fashion. But we need a warning to start with. It is also commonplace to say that for speakers of English the relation between fashion and modernity is less obvious than it already was for Baudelaire in the middle of the 19th century because he spoke French and could therefore easily associate *mode*, French for 'fashion', with *modernité*, 'modernity'. That link, though a handy rhetorical tool to start a discussion of the subject, is merely associative and by no means semantic. *Mode* derives from Latin *modus*, which simply means 'manner' or 'measure'. It entered French in the 15th century to refer to a way of living and thinking typical for a certain time or place (a sense which is now obsolete except in expressions such as *à la mode de* …). Not until the 17th century did it acquire the meaning of fashion in the sense of collective and passing habits of dressing, and in the 18th century it began to refer to fashion as a commercial and industrial activity.[6] By contrast, the etymological source of *moderne* (and therefore *modernité*) is Latin *modernus*, which was itself derived from the adverb *modo*, meaning 'recently'. As we shall see later, the link between *modern* and *modernity* is not a straightforward one, let alone the link between *modernité* and *mode*.

According to Habermas (1983: 3), Hans Robert Jauss (1982) places the first use of *modernus* in the 5th century "to distinguish the present, which had become officially Christian, from the Roman and pagan past". But the most common use made by historians of the term ***modern*** is in the distinction between ancient, medieval, and modern times. The term *modern*

[5] The best-known monographs on the subject are probably Wilson (2003) and Lehmann (2000), but there are also collective volumes devoted to the topic, such as Breward & Evans (eds) (2005), as well as chapters in other publications (e.g. Faurschou's contribution to Kroker & Kroker, eds, 1988, or Breward's chapter on the 19th century in his 1995 history of fashion, which gets the subtitle 'Fashion and modernity').

[6] In a succinct form, this history of the French word *mode* can be found in basic disctionaries such as *Le Petit Robert*.

then covers the period from the 15th century to the present, and the main characteristics, according to Styles (2005), would be:

- changes in relation to subjectivity (the rise of individualism)
- changes in technology (a scientific and industrial revolution)
- changes in geopolitical relations (an expansion of European global control from Columbus and Vasco da Gama onwards)

One famous historian, Arnold Toynbee, attaches precise dates to the Modern Age, 1475-1875. According to Malpas (2005: 33-34), he saw this period as the rise of humanism, with human beings as the source of knowledge and – through their free will – the source of action, and as culminating in the rationality and freedom ideals of the Enlightenment. For Toynbee (1954), then, this progress declines towards the end of the 19th century when a period of 'postmodern' nationalist conflict and war starts.

5th c >>	15th c	16th c	17th c	18th c	19th c	20th c	present
modern(us) to distinguish the old from the new (originally the Roman pagan past vs. a European Christian present)							
	modern times, as distinguished from ancient and medieval						
	1475-1875, *modern* age according to Toynbee						

Table 1: The term *modern* in relation to history

The second term to be explained, ***modernity***, is most often associated with political, social, economic, and cultural developments that started in the 18th century with the Enlightenment, the industrial revolution and the rise of industrial capitalism. A property that is often emphasized is a transition, under the influence of scientific, rational, and emancipatory Enlightenment thinking and of technological progress, from traditional and only slowly changing societies to rapidly self-transforming modern ones. Yet, the properties that Styles (2005) identifies for cultural, technological, and geopolitical developments since at least the 15th century, play such a central role in discussions of modernity, even in its most recent forms, that there may be good reasons to emphasize a strong form of underlying continuity. In fact, many discussions of modernity center around divergent views of continuity and discontinuity. While in line with Horkheimer & Adorno (1979), one may see 20th-century fascisms and totalitarian communism as a direct continuation of an Enlightenment logic that produced rational bureaucratic structures of domination, it is equally easy to side with Toynbee's view of the rise of European nationalisms as the true decline of Enlightenment ideals (thus effectively ending modernity towards the end of the 19th century), or with Habermas (1983) and his normative claim that modernity is still today an unfinished

project because there is a real need for the further pursuit of Enlightenment ideals in the form of rational debate as a basis for decision-making in public and political life, or of further progress of knowledge, not in the service of power but of moral and social improvement. Continuity, or the lack thereof, becomes even more problematic when the the terms ***late modernity*** or ***postmodernity*** are introduced. Especially the term *postmodernity* suggest a real break with preceding modernity. The literature on this topic is extremely complex.[7] Simplifying more than would be acceptable for other purposes, the basic properties usually singled out for a characterization of postmodernity include:

- at the level of ideas:
    - a questioning of values which Enlightenment thinking held to be universal, including first and foremost rationality itself
    - the individual subject is no longer seen as a member of a universal human category; there is not so much a clear identity as mutable performances of identity
    - emphasis on cultural diversity in a postcolonial world
- at the level of observable and accelerated change:
    - fast expansion of scientific knowledge
    - accelerated technological change (especially in communication technology, which is partly responsible for the changed experience of subjectivity)
    - rapid shift to an information-based economy (with a strong role for media and mediatization)
    - ever faster expansion and mobility (i.e. 'globalization') in commercial markets

Malpas (2005: 38) quotes Lyotard's view of postmodernity as a clear break with the grand Enlightenment narratives of progress, emancipation, and freedom, or as a complete destruction of the project of modernity by economic forces: "[…] the criteria of universalism and emancipation have been replaced by a single criterion: profit". In other words, knowledge no longer serves progress, but economic profit. It has become a commodity, the basis of power; it must serve measurable efficiency. If, however, constant change and self-transformation could be seen as the basic property of modernity, one could as easily say that the postmodern second half of the 20th century is simply a continuation of the same pattern. And if the focus is to be placed on the impact of a profit-driven economy, there may be good reason to side with Wood (1997) and her claim that the notion of growing capitalism provides a better explanation for developments from the 18th century until the present day than the distinction between modernity and postmodernity:

> The old Fordism used the assembly line as a substitute for higher cost skilled craftsmen and to tighten the control of the labour process by capital, with the obvious objective of extracting more value from labour. Now, the new technologies are used to

[7] My sources, rather than the writings of Jean Baudrillard, Zygmunt Bauman, Fredric Jameson, Charles Jencks, Jean-François Lyotard, and many other major figures, are the more accessible introductions by Delanty (2000) and Malpas (2005).

> the same ends: to make products easy and cheap to assemple (how else, for instance, would outsourcing be possible?), to control the labour process, to eliminate or combine various skills in both manufacturing and service sectors, to replace higher- with lower-wage workers, to 'downsize' the workforce altogether – again to extract more value from labour. What is *new*, then, about this so-called new economy is not that the new technologies represent a unique kind of epochal shift. On the contrary, they simply allow the logic of the old mass production economy to be diversified and *extended*. Now, the old logic can reach into whoile new sectors, and it can affect types of workers more or less untouched before. (Wood 1997: 550)

Thus in her rejection of *postmodernity* as a useful explanatory term, Woods basically presents the same analysis as Lyotard: the economy has taken over; or, in her words, capitalism has simply been universalized.

It should be clear that a really simple picture cannot be given. There are many paradoxical tendencies, some of which would point towards discontinuities in development, while others would indicate continuity. There are, for instance, still grand narratives (as pointed out by Malpas 2005): the idea of scientific progress certainly is one (which does not even differ so much from what started in the early stages of modernity); the war on terror may also be one (breaking the initial postmodern idea of the postcolonial coexistence of a diversity of cultures – an idea which is also broken locally by anti-immigrant political movements in western societies). And while communication technologies and travel patterns seem to annihilate borders, new borders are created and old ones reasserted all the time (cf. Augé 2009, who also points at the paradox of a world in which one could theoretically do everything one wants without moving, but where people seem to move more than ever).

<table>
<tr><td>15th c >></td><td colspan="2">18th c</td><td>early 19th c</td><td>late 19th c</td><td>early 20th c</td><td>late 20th c</td><td>present</td></tr>
<tr><td></td><td colspan="4">*modernity* (view that emphasizes origins in
- Enlightenment [ideals of rationality and personal freedom]
- industrial revolution and capitalism [mass production, speed, mobility, quick profits]
and accompanied by the development of rational bureaucratic structures of organization and control, with [cf. Horkheimer & Adorno] excesses in 20th-century fascism and totalitarian communism)</td><td colspan="3">*postmodernity* (seen as differing from modernity in terms of ideas and along parameters of accelerated change)</td></tr>
<tr><td colspan="3">*modernity* (for Toynbee, who views the rise of nationalisms as the decline of Enlightenment ideals)</td><td colspan="5">*postmodernity?*</td></tr>
</table>

modernity (emphasis on continuity in cultural, technological and geopolitical terms; cf. Styles)	
	modernity (emphasis on continuity of the capitalist model; cf. Wood)
	modernity (emphasis on the need for a further pursuit of Enlightenment ideals; cf. Habermas)

Table 2: Views of *modernity* and *postmodernity*

Another pair of terms to be briefly discussed are ***modernism*** and ***postmodernism***. These are relatively unambiguous in relation to styles in architecture. *Modernism* in architecture can be identified with the International Style that was developed between the two Worlds Wars and that was used extensively in the rebuilding of Europe immediately after the Second World War. It is clearly linked to the ideas of modernity, as it emphasized universality and rationality. The idea was to build, on the basis of straightforward principles that could be used anywhere in the world, uniform, ageeable and practical environments that would provide for all the needs of citizens (in sharp contrast to the slum-like conditions of many urban areas before). In doing so, the new materials which industry could provide, were used extensively. Postmodern architecture, by contrast, while continuing the use of advanced materials, reacted to the universalist uniformity of the modernists by reintegrating regional cultural elements, but without copying, and while keeping an ironic distance.

In the arts, the distinction is less simple. Modernist art started in the last decades of the 19th century as a challenge to established forms and styles. The term applies to a range of different movements (including e.g. Surrealism. Dadaism, Futurism) which were all somehow *avant-garde*. The common aim was to push the limits of art. What is, by contrast, described as postmodernist art, is basically a continuation of such experiments, but without the same sense of seriousness about the role of art in society. Another property of postmodernist art is its 'democratization', in the sense that materials and techniques were introduced that had never before counted as artictic, and that sections of the community were involved that had been excluded from art before. The latter effect was achieved by breaking down the wall between so-called high art and popular culture. In Delanty's words:

> Postmodernism in the arts and in questions of aesthetic style is a continuation of early twentieth-century radical modernism in its intensification of the symbolic domain, which is no longer confined to the aesthetic but includes the wider category of the social, the everyday life, and reaches beyond to include history and myth. (Delanty 2000: 133)

Last but not least – and here a link with the more critical views of postmodernity must emerge – postmodernist art tends to be highly commercialized. In Wood's (1997) words, we would be inclined to say that art did not escape from the expansion of capitalism, not only across the globe, but across fields of culture.

late 19th c	early 20th c	late 20th c >>
	modernism in architecture (the International Style)	*postmodernism* in architecture (incorporation of 'local' cultural elements)
modernism in the arts (various avant-garde movements)		*postmodernism* in the arts

Table 3: *Modernism* and *postmodernism* in architecture and the arts

Situating fashion in modernity

How can we situate fashion, with its own obsession with what is modern, in relation to modernity, bearing in mind the complex conceptual landscape of which I have tried to present some central points? But, first of all, why should we care? Clothes are an extremely important aspect of culture. We should remember that, as Faurschou (1988: 78) says, "the adornment of the body has rarely been a question of strict material or functional necessity". Garments have always played a defining role in the marking of differences between people: men and women, urban and rural people, adherents of different religions (even protestant vs. catholic in European history), generations, occupations, neighborhoods. This social and cultural significance of clothing also has a political side: one can impose a dress code as a restriction on one's freedom, one may dress in a way to signal submission or rebellion, or to underscore sameness or difference. These functions are probably as old as human civilization, but it is important to remember that they do not cease to play a role in the periods we want to deal with when discussing fashion specifically in relation to modernity.

The next question is what we mean precisely by *fashion*. We must specify what the term refers to in order to define its position in relation to modernity. A convenient starting point is Breward's definition:

> Fashion is taken to mean clothing designed primarily for its expressive and decorative qualities, related closely to the current short-term dictates of the market, rather than for work or ceremonial functions. (1995: 5)

As Breward shows, fashion in that sense is a phenomenon that is no doubt *pre-modern* in origin even if we place the beginnings of the modern age as early as the later decades of the 15th century (see Table 1). Restricting himself to England, he finds historical evidence for fashion in the middle of the 14th century, linked to mercantile capitalism in medieval European cities. In that period, the conscious introduction of differentiation (e.g. according to gender roles), choice, and change can already be traced, and these processes are related to market tendencies.

This is no doubt the way in which fashion continues to develop into the modern age. What is striking in this description is an immediate emphasis on economic factors, the 'market'. Breward mentions 'expressive and decorative qualities', but does not ask the question of what clothing could be an expression of. In what follows I will also focus on the economic embedding of fashion first. But then the wider social, cultural, artistic, and even political significance will also be touched upon. In the context of a discussion of fashion and modernity, those aspects should occupy an important place. Otherwise, we would have to claim that fashion is linked to only one of the pillars of modernity (industrialization and capitalism) and not to broader consequences of Enlightenment thinking (see Table 2).

Fashion in a (post)modern economy

Differentiation, choice, and change in fashion, which was already observable at least from the 14th century onwards[8] in the context of mercantile capitalism, underwent a serious transformation in the course of the 18th century. *Industrial manufacturing* techniques changed the production process, and a stronger *commercial orientation* produced faster changes. About the 19th century we read:

> Violent changes in the broader scenario of production and consumption were also reflected in the style and the form of fashionable dress itself, which in terms of variety and complexity underwent something of a visual transformation. (Breward 1995: 147)

[8] Note that such historical statements are based on research carried out on developments in European fashion. Maybe completely different statements would have to be made in relation to developments elsewhere, as for instance in Asia.

Indeed, certain clothing shapes could not have been produced so easily without the role of industrial and technological progress. The production of steel, for instance, made it possible to create the flexible cage frames of the wire crinoline. Though their popularity was short-lived, the wire was produced on a large scale, and there was a clear link with efforts at protecting newly invented products and designs; patents were taken out for crinolines and so-called 'skirt improvers' in the period 1847-1870. Another example is the invention of aniline dyes, which made it possible to produce fabrics in colors such as purple and mauve which had been very hard to obtain before.

Possibilities of industrial mass *production* went hand in hand with *trading* methods oriented to finding cheaply produced goods in large quantities, sold quickly with small profits. Serious business empires were built that way, for instance by James Morrison in the first half of the 19th century (see Dakers 2005). At the level of trade, there were also two other innovations, the *department store* and *fashion magazines*:

> The rise of the department store and the expansion of women´s fashion magazines, both designed to serve all classes, undoubtedly transformed and 'modernised' the culture and consumption of dress in the second half of the century. (Breward 1995: 166)

By 1870, *Le Bon Marché* was well-established in Paris and the model was soon adopted elsewhere in Europe, especially in London. The purpose of department stores was to put on display the widest range of new branded fashionable goods at fixed prices. They made it possible to browse goods that had to be displayed in such a way as to make them compete for attention. Having different products available thus became a necessity. This development went together with the rise of a middle class of (mostly male) white-collar workers whose wives had some money to spare and for whom shopping was a welcome break from the chores of the home and for whom consumption became part of their bourgeois identity. Also men consumed fashion, but with the exception of the obvious excesses of dandyism (with Beau Brummell as an early prototype) they were more 'hidden' consumers (see Breward 1999); there were, for instance, no male fashion magazines until the 1920's.

Yet, however strongly fashion may be associated with the rise of industrial capitalism, both in terms of production and of trading, its position in relation to those economic developments is not unambiguous. There are at least three specific developments that modify the picture.

First, in spite of large-scale production and trading, clothes-making required a serious amount of detailed and mostly manual labor which, in turn, made it possible to preserve a high degree of differentiation. Quoting Styles,

> Dress, despite all the links that scholars have made between it and a variety of modernities, never fitted easily into those narratives of economic history whose culmination was Fordist mass production. Economic historians wrote extensively about how clothing materials were manufactured, but with few exceptions their [p. 34] histories concluded at the textile factory gate. The production of clothing itself was too fragmented, too small-scale, too primitive to incorporate into their grand narratives of industrial progress. But with the undermining of those grand narratives and the reassertion of the small-scale, the flexible and the diverse, the distinctive, fashion-driven features of the way clothes are made appear less a historical backwater than one of the main currents in the history of modern manufacturing. (Styles 2005: 34-35)

Whether or not we go along with Styles' suggestion that similar processes can be observed in other areas of manufacturing, this picture is certainly correct for the fashion industry. We must remember that even though steel wire for the crinoline was really factory-produced, once the wires were there and had been cut at specific lengths, the pieces went to workshops where the cage frames were manually constructed.

A second remark pertains to the way in which heavy industry itself facilitated small-scale home production. From the early 1850's, sewing machines were manufactured. These were soon adopted for home alterations of pieces of clothing. And by 1865 free dress patterns for home production were distributed with the fashion magazines. In other words, small-scale production was itself stimulated by the availability of machines that large-scale industry provided.

Third, the 19th century witnesses the emergence of the 'designer' as a profession. Fashion design as a specific professional role aimed at the creation of differences and changes in style, is generally assumed to date from the middle of the 19th century. The British Charles Frederick Worth is often presented as the first real 'fashion designer'. In 1858 he established the first real 'haute couture' house in Paris. What distinguished him from the 'tailors' before him was that he did not just make the clothes that customers wanted him to make. Instead, he decided what he would make for whom and how. In other words, he really 'designed' their clothes. This activity, by definition, does not fit the picture of industrial mass production. Still it suits ideas of modernity: tradition is consciously deviated from in a search for new forms adapted to the individual

All these properties of the positioning of fashion in an economy of modernity continue into the 20th century and until the present day, with an ever stronger incorporation of fashion in a general world-wide *consumer culture*. There is a vast expansion of the market for fashion magazines, supplemented with websites and blogs. There is the increased role of the mass media, starting with the Hollywood film after 1920, but generalized through television and the internet: sales of a fashion line boom when a celebrity is seen to wear a certain brand. Chain stores have multiplied, and so have mail order companies (these days mostly internet-based).

And fashion is produced and marketed by giant multinational corporations, ranging from cheaper single-brand companies (H&M, Zara, etc.) to business conglomerates at the high end of the fashion scale (such as LVHM). This does not mean that smaller businesses have disappeared:

> At the same time, paradoxically, a reorganisation of business practices, or marketing and advertising, prioritised particular strands of society as fashion leaders. A cult of 'the designer' revolving around ideals of couture and 'high fashion' or strong subcultural identities ensured the survival of older hierarchies based on notions of quality, style and individuality […]. (Breward 1995: 183)

The paradox here, however, is not simply an opposition between mass production and smaller (and more expensive) labels. Mass production and distribution with small profit margins is obviously in line with the industrial trends since the 18th century. Affordable products are made, which seems democratic, but such quantities are produced that really big capital is required, and profit margins are handled that are often made possible only by economic exploitation on the production side. But breaking away from this pattern by single designers or small design teams makes clothes more expensive, and the survival of designer brands requires consumers with more buying power. Thus fashion seems trapped in a truly (post)modern capitalist economic logic at both ends. This picture looks even worse when the effect of advertising and image building is taken into account: some conglomerates with serious capital reserves incorporate expensive brands that have achieved such a status in the hierarchy of desirability that high profit margins can be maintained without harming the expansion of sales. When looking at the economic side it is hard, therefore, not to side with Wood's emphasis on the expansive influence and continuity of a capitalist model in the modern and postmodern world, rather than to assume a significant break between modernity and postmodernity (see Table 2).

Fashion, social status, and politics

As said before, garments have always played a defining role in the marking of social differences between people. This universal phenomenon manifests itself in specific ways since the 18th century. At the earliest stages of industrial manufacturing, clothes were beginning to signal the difference between members of the rising bourgeois classes on the one hand and the old aristocracy on the other. In the beginning, bourgeois dress reflected mostly rather conservative values: durability (based on the high quality of the manufactured materials) and modesty (though not neglecting elegance). With the fading role of aristocratic

elites, this original distinction became much less outspoken. Many types of new distinctions emerged over time, others became more outspoken, and still others disappeared, often in conjunction with modern social developments.

One of the more striking developments may have been that in the 19th century the difference between *male and female dress* became more visible. Male dress was inspired by earlier aristocratic sportswear (with tight-fitting pants reminiscent of riding outfits). Female dress was beginning to place much more emphasis on female elegance, also in public, where it was no longer obligatory to hide behind a cloak and veil. Such gender differences, while persisting, have also been played around with in many ways: the introduction of trousers for women from the 19th century onwards, the emergence of androgynous fashions for men as early as the 1920's, and so on.

In addition to fluctuating ways of signaling gender roles, *social status differences* manifested themselves in dress in varying ways. In the modern cityscape is has always been possible to distinguish, on the basis of dress as much as behavior, bourgeois passers-by from street vendors, from shopkeepers, from office workers, from factory workers, and so on. And even though today well-made fashionable clothes are widely accessible for many people in western societies, it does not take much training to distinguish those who can afford (or choose to invest in) luxury brands, even if the products are not conspicuous and worn with modesty.

Clothes may of course also be chosen and worn to make explicit statements of *distinction* or of *belonging*. The phenomenon of 'uniforms' ranges from the unchangeable aspects of clothing which are self-prescribed by certain ethnic or religious groups (whether Amish, Sikhs, or Hasidic Jews) to the garments of a variety of occupations (e.g. nurses, whose uniforms have themselves undergone fashion changes over the years). Any type of uniform signals one´s belonging to a group. But belonging itself is also a matter of being distinct from others. And distinctness can be explicitly emphasized with what Wilson (2003: 179-207) calls 'oppositional dress'. In this context, Wilson also shows the fundamental ambivalence of fashion as a social, and in many ways political, phenomenon: forms of oppositional or rebellious dress may become highly 'fashionable' and even set the norms for later standard developments.

A first example of this we can already find in late 18th and early 19th century dandyism. When political upheavals made aristocratic birth and wealth less relevant, dandies consciously sought social distinction in style and pose. They paid extreme attention to what they were wearing, being careful to not to look conspicuously adorned. Their style, which is often referred to as a style of 'understatement' or even 'anti-fashion' in dress, also included the introduction of adaptations of earlier aristocratic sportswear. Though their habits of spending hours polishing their boots or deciding how to knot their ties did not spread widely, the resulting inconspicuous and laid-back style set the tone for much of the later development of men's fashion – turning men, in Breward's terms, into somewhat hidden consumers of fashion rather than a species not interested in fashion at all.

Another example would be the 'bohemian' style which developed in the early 20th century in London's Chelsea and New York's Greenwich Village. The inhabitants of these areas, artists, writers, journalists, consciously challenged reigning bourgeois norms, for instance by leaving off hats at a time when all men and women were wearing them, by not using cosmetics, by letting women wear either men's clothes or flowing robes, etc. But soon this radicalism turned into a new form of consumerism, because the rebellious dressing habits became fashionable themselves. The same can be said of the black colors that were worn by beatniks as a statement of dissent, or of the hippie style that emerged as an aspect of student counterculture. Hippies' use of second-hand clothes, for instance, was not without its own paradoxes:

> Yet although this was undertaken in a spirit of anti-consumerism, it did involve the expenditure of much time if not money, and reintroduced the snobbery of uniqueness, since there was, necessarily, only *one* of the 'frock' you had found – just as much as if you'd bought a Dior original. (Wilson 2003: 193)

About the introduction of unisex dress, itself related to the sexual revolution of the 1960's, Suzy Menkes is said to have remarked as early as 1984

> […] that this form of 'cross-dressing', which is opening up the way to 'gender-bending' unisex departments in exclusive fashion stores, is simply a new fad and that – significantly – the market it is aimed at is the market of affluent heterosexual *couples* for whom androgynous dress symbolizes not an attack on gender but merely a reaffirmation of middle-class togetherness. (Wilson 2003: 203)

Looked at in this way, there seems no escape from patterns of social dominance. Some analysts have made the link with political power relations. Particularly popular are illusions to the relationship between Parisian fashion houses and the Nazi occupiers during the Second World War. The recent commotion surrounding John Galliano's "I love Hitler" scandal and his subsequent dimissal from Dior made Peter Popham write an article entitled "Fashion and fascism – a love story" for *The Independent on Sunday* (6 March 2011, p. 48). He claims that the fashion industry's vulnerability when there are allusions to fascism (and hence the quick dismissal of Galliano) simply result from the fact that it does not want to be reminded of a history of past collaboration. Some even go further: there was a comment on the incident saying that the ideals of beauty that are stimulated by the fashion industry are inherently comparable to Nazi ideology aimed at the creation of a pure and perfect breed of human being …

Without denying the paradoxical status of fashion in relation to economic and social structures, and hence inevitably to politics as well, we must remember that

> […] because fashion, like capitalism itself, is so contradictory, it at least has the potential to challenge those ideologies in which it is itself enmeshed – as can all popular cultural forms, so long, that is, as we have some coherent political position […] from which to criticize. (Wilson 2003: 205)

In this sense, fashion can even play a fundamental role in the further development of modernity in Habermas' interpretation of it as an unfinished Enlightenment project (see Table 2).

Fashion and individuality

As we have seen, modernity and postmodernity are closely associated with developing notions of subjectivity. Even if a postmodern logic no longer sees the individual subject as a member of a universal human category with shared rationality, and even if people are no longer conceived as having a clear identity but rather as displaying mutable performances of identity, an emphasis remains throughout the (post)modern period on individual responsibility, choice, and positioning. Not surprisingly, the relationship of fashion with individuality is as ambivalent as its relation with the economy and social structures. Group identity and conformity, as well as individual identity and self-expression, all play a role. According to Hill (2005), in the world of clothes, *sameness* dominates, the apparent opposite of individuality. Observing similarities in dress in a modern city, he observes:

> Perhaps though what people were wearing was indicative of the type of individuals they had become – adrift in a world of consumption, without direction, context and meaning. (Hill 2005: 75)

He adds:

> This offers an intriguing perspective on the state of the contemporary individual – dressing in a way that signifies very little, but at the same time anxious about what they are wearing. (Hill 2005: 76)

However, the opposite may be stressed as well:

> Twentieth-century fashion, rather than producing a sense of undifferentiated sameness, has actually supported fashion changes as frenetic and diverse, as open to the possibilities of individuality, as any other period. (Breward 1995: 184)

Some analysts formulate this in extremely negative terms:

> In contrast to the productivist ethic of industrial modernism, late capitalism is the society of consumption, the society of the mass market and multinational capital, the age of media, information, and electronic reproduction. It is no longer an economy seeking to fulfill the needs of a modernizing society but a society driven to create a perpetual *desire for* need, for novelty, for endless difference and instant satisfaction. (Faurschou 1988: 82)

What is referred to here is the logic of planned obsolescence: these days, industries have a tendency to produce things in such a way that they are outdated very soon so that they need to be replaced by newer products. Faurschou (1988: 92), thinking about the fast seasonal changes in fashion, goes as far as to say that "[…] fashion has become the commodity 'par excellence'". This verdict ignores the entire range of electronic gadgets, where the need for trend followers to always possess the newest item seems a lot stronger. In that field, moreover, the choice of 'newest items' is more restricted, and an item that gets outdated loses its value (and often even its usability) completely. Fashion is notably different: even someone really trying to keep up with fashion trends must constantly choose between styles and items that best suit his or her personal tastes, so that individuality remains an important factor, and one can always go back to pieces of clothing from earlier seasons when choosing something to wear on a specific occasion. Obsolescence, in other words, cannot be truly planned in the world of fashion.

Similarly, thinking about fashion in exclusively negative terms as reducing the body to a perfect 'object' without meaning, completely stripped of human individuality and personality, so that it could as well be replaced by a lifeless mannequin[9], seems extremely one-sided.[10]

[9] Such views are attributed to a theorist of postmodernity such as Jean Baudrillard.

[10] For a balanced account of fashion as an expression of individual identity (in relation to gender roles, social status, sexuality), see Davis (1992).

The artistic and cultural positioning of fashion

There is no need to deny that fashion occupies an ambivalent position in relation to economic processes, social structures, and even ideas of individuality, and that this positioning is related to aspects of modernity that often have negative as much as, if not more than, positive sides. But its being embedded in a capitalist economy (as simply as aspect of money culture and conspicuous consumption – cf. Veblen 1953), and the strong potential it carries for marking social class distinctions (as analyzed by Bourdieu 1979), certainly are not the only things that need to be emphasized when trying to make sense of the relation between fashion and modernity. In other words, fashion is not just an overheated piece of international economic machinery that is permanently in ´fast forward´ mode. It is true that

> […] fashion is a space where industry articulates issues of identity and signification for the purposes of competitive advantage to such a degree that culture and economy become mutually constitutive to the extent of being analytically inseparable. (Briggs 2005: 81)

But it is important to also emphasize the *cultural side* of the equation, and in particular the *artistic status* of fashion.

Particularly helpful is Lipovetsky's (1987) study of fashion in which he places, as Breward (1995) does, the origins of fashion at the end of the 14th century, about a century before what historians call the modern age (see Table 1). He adds that, roughly speaking, there are two periods to be distinguished in the history of fashion: from the 14th century until about 1850, and the period from 1850 onwards to the middle of the 20th century. This distinction is adopted by Stern (2004) to delimit the period of his own study of clothing as art: 1850-1930. Note that this period starts roughly with Baudelaire's reference to fashion prints to illustrate his view of beauty in modernity as historical and changeable rather than universal and absolute, and coincides with the period of *modernism in art* (see Table 3). One defining property of the period, for Lipovetsky, is the fact that, with the establishment of *haute couture* houses (Worth being the first in the 1850's), fashion becomes a form of art. Recognition as art leads to an integration of fashion in the modernist avant-garde art scenes:

> […] after Baudelaire, the debate over fashion only became more intense. It involved controversies over fundamental issues in the art theory and aesthetics of the second half of the nineteenth century, including the abolition of the traditional hierarchy between 'major' and 'minor' arts, a questioning of the difference in 'status' of artists

> and craftsmen, and the artist's wish to go beyond the traditional boundaries of art. For many artists at the end of the nineteenth century and the first half of the twentieth, dress design was something far too important to be left to couturiers alone. The historical avant-gardes would appropriate dress design as a privileged field in which the artist could overstep the limits of 'pure' art and act directly on daily life.
>
> These artists' dress proposals are very diverse in terms of style, but they all proceed from a common will to reject 'official' fashion, refusing its mercantile logic and striving to replace it by a utopian 'antifashion'. (Stern 2004: 3)

In this way, fashion came to share the artistic goals of avant-garde art movements. Stern describes, by way of illustration:

- French romanticims and its movement from individual eccentricity to artistic dress;
- rational, artistic, and aesthetic dress in England
- Henry van de Velde's fashion designs based on principles comparable to his modernist architecture (and to their integration in an environment that could be seen as a *Gesamtkunstwerk*)
- designs by Gustav Klimt and the members of the Wiener Werkstätte
- Giacomo Balla's futurist dress designs
- the utility-oriented dress style of the Russian avant-garde
- Sonia Delaunay's experiments with 'simultaneous dresses' (incorporating a diversity of materials, shapes, and colors)

None of these developments were accidental. They were carefully thought about, as appears from the writings by many of the protagonists, collected in Stern's *Against Fashion*. In other words, just like developments in painting, sculpture, and architecture, developments in fashion were meant to be meaningful, expressing ideas about beauty and attitudes to society, social trends and structures. An explicit discussion of many of the issues can for instance be found in a work by another modernist architect, Le Corbusier, *L'art décoratif d'aujourd'hui* (1925), which Wigley (1995) uses as his starting point for a lengthy account of the relations between modernist architecture and fashion, which both also represent a field of tension between art and technique in attempts to create interesting new forms with relatively simple variations and without directly aiming at ornament.

The link between fashion and art goes beyond the work of artists who ventured into fashion. There are numerous examples of fashion designs inspired by works of art. But more importantly, designers and visual artists share both goals and methods. As to *goals*, they all try to bring an innovative visual message which cannot necessarily be formulated in words but which is nevertheless pertinent to people's lives. As Wilson puts it:

> Art is always seeking new ways to illuminate our dilemmas; dress, however tainted a medium – from its association with the body and with daily life and behaviour – nevertheless does this too. (Wilson 2003: 247)

As to *methods*, designers and artists alike get their inspiration from all areas of experience, including older art (which is why progress is sometimes made by moving cyclically rather than straight ahead). As Brassett (2005) suggests, a useful notion to understand fashion, as well as art, may therefore be the notion of 'emergence': in creative processes the whole becomes more that the sum of the parts, and there is a form of self-organization at work; changes may happens without direct outside input, and qualitative changes of direction may unintentionally grow out of a struggle with sources of inspiration, ideas, techniques, materials, and all their limitations.

When discussing recent developments, Wilson (2003) observes that *haute couture* may be losing the dominant leadership position it once had. Fashion shows usually run at a loss, presenting clothes that are hardly sold but that serve as expensive advertisement for more affordable ready to wear and derived products such as cosmetics and accessories. In her opinion, this development brings fashion even closer to the art world. A symptom may be the frequency with which fashion is featured in museums nowadays. New may be the fact that not only the traditional visual arts are involved these days, but a strong performance element is introduced as well.

A complex picture

When Obama was running for the US Presidency, there was a New York homeless man trying to appeal to passers-by with a cardboard sign saying "I'm just like Obama – I want change". There might be a temptation to use this man's play on the polysemy of 'change' as a metaphor for the fashion world: innovation, yes, but with money as the driving force. The importance of the economic side cannot be downplayed. It is certainly the dominant factor from the point of view of the fashion trade. But the phenomenon of fashion is not that simple, especially from the designer's perspective. A designer's basic, and highly complex, question is: **what** can be made **how** and **where** in a sufficiently **creative and innovative** way, appealing to people's ambivalent **needs** for identity, conformity, and individual self-expression, and producible in an **affordable** and **ethical** manner? This formulation of what is both *an artistic and a practical problem* addresses all parameters of modernity (and postmodernity, if one wants to use that term).

References

Augé, Marc (2009) *Pour une anthropologie de la mobilité*. Paris : Éditions Payot & Rivages.

Bindas, Kenneth J. (1990) 'Review of Elizabeth Wilson, *Adorned in Dreams'*. *The History Teacher* 24(1): 127-128.

Bourdieu, Pierre (1979) *La distinction*. Paris : Éditions de Minuit.

Brassett, Jamie (2005) 'entropy (fashion) and emergence (fashioning)'. In Breward & Evans (eds), 197-209.

Breward, Christopher (1995) *The Culture of Fashion: A New History of Fashionable Dress*. Manchester: Manchester Univerity Press.

— (1999) *The Hidden Consumer: Masculinities, Fashion and City Life 1860-1914.* Manchester: Manchester University Press.

— & Caroline Evans (eds) (2005) *Fashion and Modernity.* Oxford: Berg.

Briggs, Adam (2005) 'Response'. In Breward & Evans (eds), 79-81.

Dakers, Caroline (2005) 'James Morrison (1789-1857), 'Napoleon of shopkeepers', millionaire haberdasher, modern entrepreneur'. In Breward & Evans (eds), 17-32.

Davis, Fred (1992) *Fashion, Culture, and Identity*. Chicago: The University of Chicago Press.

Delanty, Gerard (2000) *Modernity and Postmodernity: Knowledge, Power and the Self.* London: Sage.

Faurschou, Gail (1988) 'Fashion and the cultural logic of postmodernity'. In Kroker & Kroker (eds), 78-93.

Foster, Hal (ed.) (1983) *Postmodern Culture*. Port Townsend, WA: Bay Press.

Habermas, Jürgen (1983 [1981]) 'Modernity – An incomplete project'. In Foster (ed.), 3-15.

Hauser, Kitty (2005) 'The fingerprint of the second skin'. In Breward & Evans (eds), 153-170.

Hill, Andrew (2005) 'People dress so badly nowadays: Fashion and late modernity.' In Breward & Evans, 67-77.

Horkheimer, Max & Theodor Adorno (1979 [1944]) *Dialectic of Enlightenment*. London: Verso.

Jauss, Hans Robert (1982) *Toward an Aesthetic of Reception*. Minneapolis: University of Minnesota Press.

Kroker, Arthur & Marilouise Kroker (eds) (1988) *Body Invaders: Sexuality and the Postmodern Condition*. London: Macmillan.

Le Corbusier (1925) *L'art décoratif d'aujourd'hui*. Paris: Éditions G. Grès et Cie.

Lehmann, Ulrich (2000) *Tigersprung: Fashion in Modernity*. Cambridge, MA: MIT Press.

Lipovetsky, Gilles (1987) *L'empire de l'éphémère: Le mode et son destin dans les sociétés modernes.* Paris : Gallimard.

Malpas, Simon (2005) *The Postmodern*. London: Routledge.

Purdy, Daniel (2004) 'Review of Ulrich Lehmann, *Tigersprung*'. *The Journal of Modern History* 76(4): 943-945.

Radford, Robert (2001) 'Review of Ulrich Lehmann, *Tigersprung*'. *The Burlington Magazine* 143(1179): 381.

Ray, Larry (1997) 'Post-Communism: Postmodernity or modernity revisited?' *The British Journal of Sociology* 48(4): 543-560.

Stern, Radu (2004) *Against Fashion: Clothing as Art, 1850-1930*. Cambridge, Mass.: The MIT Press.

Styles, John (2005) 'Response'. In Breward & Evans (eds), 33-37.

Toynbee, Arnold (1954) *A Study of History*, vol. IX. Oxford: Oxford University Press.

Veblen, Thorstein (1953 [1899]) *The Theory of the Leisure Class: An Economic Study of Institutions*. New York: Mentor Books.

Wigley, Mark (1995) *White Walls, Designer Dresses: The Fashioning of Modern Architecture.* Cambridge, Mass.: The MIT Press.

Wilson, Elizabeth (2003 [1985]) *Adorned in Dreams: Fashion and Modernity*. London: I.B. Tauris & Co Ltd.

— (2005) 'Fashion and modernity'. In Breward & Evans (eds), 9-14.

Wood, Ellen Meiksins (1997) 'Modernity, postmodeernity or capitalism'. *Review of International Political Economy* 4(3): 539-560.