HECO #591274, VOL 23, ISS 3 (June 21, 2011)

Ecological Psychology, 23:1–36, 2011

ISSN: 1040-7413 print/1532-6969 online
DOI: 10.1080/10407413.2011.591274

Routledge
Taylor & Francis Group

First-Order Languaging Dynamics and Second-Order Language: The Distributed Language View

Paul J. Thibault
Faculty of Humanities 5
University of Agder
Kristiansand, Norway
Faculty of Education
University of Hong Kong
Hong Kong SAR, PR China 10

This article articulates some aspects of an emerging perspective shift on language: Q1
the distributed view. Rejecting the classical view that the fundamental architecture of language can be localized as a number of interrelated levels of formal linguistic organization that function as the coded inputs and outputs to each other, the distributed language view argues that languaging behavior is a particular biocultural 15
manifestation and organization of the dynamic matter-energy-information flows that constitute the material reality of the world. This organization is irreducible to the formal abstracta that have characterized the focus on a de Saussure-type system of formal regularities in mainstream linguistics over the past century. Language, on the distributed view, is a radically heterogeneous phenomenon that 20
is spread across diverse spatiotemporal scales ranging from the neural to the cultural. It is not localizable on any one of them but involves complex interactions between phenomena on many different scales. This article focuses exclusively on talk rather than text. Talk and text are very diverse phenomena that ought not to be conflated with one another as in many recent discourse-analytical ac- 25
counts. A crucial distinction is thus presented and explained, namely, first-order languaging and second-order language. The former is grounded in the intrinsic

Correspondence should be addressed to Paul J. Thibault, Faculty of Education, University of Q2
Hong Kong, Hong Kong SAR, PR China. E-mail: pauljthibault@gmail.com

expressivity and interactivity of human bodies-in-interaction. On this view, first-order languaging behavior between persons exploits, synchronizes, and entrains the dynamical properties and functional capacities of pico-scale bodily events measured in timescales ranging from milliseconds to fractions of seconds. The real-time bodily dynamics of interacting agents are, in turn, constrained by second-order patterns emanating from the cultural dynamics of an entire population of interacting agents on longer, slower cultural-historical timescales. Second-order patterns are intrinsically normative. Lexicogrammar is one manifestation of such second-order constraints on first-order languaging dynamics. The article also engages in a dialogue with Gibson's ecological theory of perception–action. It argues that socially coordinated and culturally shaped languaging behavior affords extended forms of virtual perception–action that are guided by the seeking of values in the extended human ecology as agents affect and are affected by each other in their languaging behavior. An analysis of a videorecorded interaction serves to illustrate some aspects of the integration of scales involved in the whole-body sense making that is talk.

Recently, it has become possible to think about "language" in new ways. This means rethinking and even rejecting many of the core assumptions, descriptive formalisms, institutionalized abstracta, and modeling procedures of an entire family of 20th-century linguistic theories that were founded, with varying emphases, on some version of the idea that language is an abstract code, or de Saussure-type system (de Saussure, 1993 [*Cours* III, 2^{e} *partie*: 1910–1911]). It is now becoming possible to substantially rethink the abstract formalisms and associated assumptions on which this family of models is founded. Advances in the theory of dynamical systems grounded in our understandings of neurobiology; dynamic systems thinking; embodied, embedded cognitive science; and ecosocial systems have opened up the possibility of a different approach—an approach which recognizes that language is a cultural organization of process that is naturalistically grounded in human biology. These developments can help us to build new understandings of language that can stand as credible alternatives to the mechanistic and reactive models, founded on computational models of cognition, input-output coding systems of abstract symbolic forms, and text processing, that have predominated in the past few decades in the language sciences. The new approach stresses the centrality of coacting agents who extend their worlds and their own agency through embodied, embedded processes of languaging behavior rather than uses of an abstract language system (Cowley, 2007a, 2007b, 2008; Steffensen 2008; Steffensen, Thibault, & Cowley, 2010; Thibault, 2008). In this approach, the focus is, initially, on how dialogically coordinated first-order languaging is enacted and created in human activities that are spread across a diversity of timescales. The term "languaging" therefore reflects the focus on the dynamics of real-time behavioral events that are cocon-

structed by coacting agents rather than the more usual view that persons "use" a determinate language system or code (see First-Order Languaging section).

This new perspective is mindful of and takes part in the recent move away from social constructivism in the humanities and the social sciences to an approach that integrates ecological, social, cultural, biological, material, and bodily dimensions. These developments represent a renewed attempt to better understand the materially embodied, culturally/ecologically embedded, naturalistically grounded, affect-based, dialogically coordinated, and socially enacted nature of languaging as a form of whole-body behavior or whole-body sense making. This loose set of recent developments also goes hand in hand with significant shifts in orientation in the cognitive sciences and their ongoing project to naturalize cognition. This has taken the form of a shift from classical or first-wave (internal symbol manipulation) and second-wave (connectionist) cognitive science to third-wave theories and models of embodied, embedded distributed cognition and the place of languaging behavior in these.

In the classical view, humans are input–output systems that use language to "represent" and "communicate" information by means of mental representations housed in the brain of each individual. The externalized mirror image of this view is the notion that "text" is the encoding of a social semantic that again is abstracted away from material dynamics. The two versions, albeit with different goals and emphases, are entirely compatible with the computational theory of mind developed by classical cognitive science, namely, language is an encoding/decoding mechanism based on inputs and outputs from an abstract language system that is instantiated as spoken or written "text." Traditionally, language has been viewed as the processing of abstract symbolic forms on various levels of linguistic organization. Abstract forms replace real-time dynamics, body movement, feeling, and a history of first-person experience. The notion that brains operate on internal symbolic representations is no longer accepted by a growing number of researchers (Clark, 1997; Goldfield, 1993; Hutchins, 1995; Wheeler, 2005). Instead, human language is seen more and more as a suite of flexible and adaptive behaviors that are based upon a naturalistically grounded intersubjective sensitivity to the bodily dynamics (movement) of others and the sensorimotor coupling relations between persons and their worlds that result from this in the intersubjective matrix (Bråten, 2007; Cowley, 2007a; Trevarthen, 1998). It is through participation in this intersubjective matrix that persons learn to attune to the intentions and motives that inform bodily dynamics and the bodily feelings that modulate and evaluate the dynamics (Bråten, 2007; Bråten & Trevarthen, 2007; Trevarthen, 1987, 1992, 1998). Situatedness, bodily feelings, interpersonal engagement, sensorimotor coupling, and the capacity to move others and be moved by them are central. Cognitive and affective processes are distributed across brains; bodies; and aspects of the social, cultural, and physical worlds of persons (Hutchins, 1995). On this view, it is the coupling of dynamical processes

operating on diverse timescales (e.g., neural, bodily, situational, cultural) that shapes human behavior, learning, cognition, and language. These are seen as specific organizations of process of living human systems embedded in a human cultural ecology.

In the next section, we consider these questions in terms of the distinction between *first-order languaging* and *second-order language*. We consider how second-order cultural patterns (wordings) get integrated to, not encoded by, first-order languaging dynamics.

FIRST-ORDER LANGUAGING AND SECOND-ORDER LANGUAGE

The Problem: Form-Based Abstracta in the Language Sciences

Typically, linguists of diverse theoretical persuasions have postulated and focused on abstracta of various kinds and on different levels of linguistic organization, including, for example, phonology, lexicogrammar, and discourse. The focus in each case has been on abstract form, for example, phonological and grammatical forms and their regularities, seen as constitutive of diverse levels of "language." One consequence of this view was that regularities of synchronic forms replaced or displaced, with few exceptions, productive causal relations between events in languaging behavior. Abstract linguistic forms derive from a de Saussure-type *langue*, which de Saussure (1993 [*Cours* III, 2^{e} *partie*: 1910–1911]) constituted as the object of linguistic theory. In doing so, de Saussure, to be sure, made a shrewd academic political move. This amounted, in part, to staving off the boffins in white laboratory coats in the then nascent brain sciences and thereby insulating the study of language from any attempt to give it a naturalistic grounding (see Harris, 1987, pp. 39–49, for relevant arguments apropos). de Saussure's move effectively stymied fruitful dialogue between the biological and cultural sciences on the nature of language for many decades. The essentially negative consequences for the language sciences are registered to the present day. This led to many dead ends during much of the 20th century. de Saussure's theoretical construct *langue* was accordingly constituted as a collective social abstraction (Thibault, 1997). Synchronic, form-based theories and their descriptive formalisms consequently replaced the dynamical, time-locked organization and integration of processes on different timescales as the object of study and theoretical explanation. The resulting abstracta are artificially separated from the dynamics of first-order languaging behavior and accordingly analyzed as abstract linguistic (e.g., phonological and verbal) forms. On this view, meanings are mapped onto forms according to arbitrary social conventions

and codelike mechanisms (see Harris, 1981; Love, 2007, for relevant critique). According to traditional linguistics, language consists of abstract patterns such as phonology, lexicogrammar (morphosyntax), and semantics, seen as different levels of linguistic organization.

The cumulative effect of the emphasis on a de Saussure-type synchronic system, aside from the fact that it probably represents a very one-sided reading of de Saussure (Thibault, 1997, 2005a), has been to eliminate nonlinear causal capacities of material systems qua historical individuals on different ontological scales (e.g., persons, dyads, interpersonal networks). This has put the focus pretty squarely on a homogenized synchronic language system based on the constant regularities of formal abstracta and low-intensity equilibria. The productive capacity of causal connections and material dynamics is accordingly relinquished. The term "languaging," as distinct from the more usual "language," serves, therefore, to remind us, as Maturana (1970) understood, that what we habitually and unthinkingly call "language" is an open-ended meshwork of interlinked functioning components (De Landa, 2006) founded on material dynamics that know no single stable state based on abstract forms. If we assume that a language is a stable synchronic state, we end up with the idea that people "use" a stable language system and its concomitant that individual persons are "language users."

Abstract systems of this kind sideline persons, their internal complexity, and the complex material–causal interventions that persons must perform in their worlds. They also sideline the complex causal interlinkages between persons, artifacts, technologies, cognitive and bodily skills and capacities, and so on. These are seen as secondary to the postulation of formal systems, themselves the results of second-order practices that conflate theory and folk theory, as in the code view of language (Love, 2007). The emphasis on formal abstracta has had two main consequences in the language sciences: the deemphasizing of causal explanation and the elimination of material dynamics in favor of form-based regularities.

First-Order Languaging

The distributed language view that I present in this article is founded, in the first instance, on a very different theoretical object, namely, *first-order languaging*. The term *first-order languaging* refers to the organization of process on different scales that takes place when persons engage in talk together. Text and talk are radically different phenomena. However, many discourse-analytical approaches, owing to their reliance on verbal abstracta and associated transcription practices, tend to conflate the distinction precisely where it is most crucial to maintain it. This article is exclusively concerned with talk. Let me begin to define first-

order languaging by pointing out what many researchers have long recognized: languaging is, crucially, a dialogical phenomenon and is not therefore the possessions of individual brains and bodies. This is a fundamental point of departure. First-order languaging crucially involves synchronized interindividual bodily dynamics on very short, rapid timescales of the order of fractions of seconds to milliseconds. Following previous work in this area (see Cowley, 2006, 2007a [2006], 2007b; Cowley & Love, 2006; Steffensen et al., 2010; Thibault, 2008), I refer to these very rapid temporal frames of the dynamical properties of talk as pico-scale bodily events. Persons in talk enact, exploit, respond to, and attune to such events in order to engage with others and to coconstruct their worlds with them. Moreover, "language patterns" do not exist independent of bodily dynamics.

Unlike the purely metaphorical and misleading use of the notion of "construction" in social constructivism, the concept of first-order languaging takes it as axiomatic that persons assemble languaging events out of a heterogeneity of resources and component parts that are interlinked as wholes on the basis of objective historical processes on diverse ontological scales. The ontology is realist (not constructivist) not because it postulates reified essences corresponding to the identities of things but because it focuses on objective historical organizations of process on different ontological scales that cut across the nature–culture dichotomy in the way theorized by Deleuze and Guattari (1988; see also De Landa, 2006, p. 3).

Languaging was, originally, Maturana's (1970) term for exploring how talk is a complex behavior oriented to the creation and sustaining of consensual domains (see also Kravchenko, 2003). The biological basis of talk qua complex behavior means that human infants, in the first instance, depend on value biases that enable them to exploit dynamical movement patterns in the vocal and other bodily activity of adult interlocutors. In doing so, they learn how to align and integrate their behavior with the dynamics of complex socially coordinated languaging behavior. Infants synchronize with the rhythmical movement patterns of caregivers and in so doing they synchronize their own bodily movements to those of the caregiver. By synchronizing in this way, they learn to attune to the intentions and feelings that inform the other's movements (Trevarthen, 1998). At the same time, they learn, in time, to discriminate and to be attracted to higher level phonetic and lexicogrammatical patterns (Kuhl, 2007) in ways that connect both to aspects of situations and to cultural norms. Languaging behavior is grounded in the real-time dynamics of interacting bodies in and through which persons coordinate their actions, intentions, perceptions, and feelings with each other. Code models, on the other hand, depend on high-level descriptions of abstract linguistic forms on different levels (e.g., phonology, lexicogrammar, and discourse) that are not the basic stuff of everyday interaction between coacting agents in talk.

First-order languaging just is whole-body sense-making activity that enables persons to engage with each other in forms of coaction and to integrate themselves with and to take part in social activities that may be performed either solo or together with other agents. They are also able to attune to and hence to apperceive both aspects of situations and the informing intentions that flow from the agent's internal dynamics through time-locked bodily movement dynamics—intentions that modulate the time-locked trajectory of bodily movement until the completion of the desired goal state or intention (Juarrero, 1999, pp. 48–49; Thibault, 2004a, pp. 123–124). First-order languaging is not limited to vocalizing but includes a whole range of bodily resources that are assembled and coordinated in languaging events together with external (extrabodily) aspects of situations, environmental affordances, artifacts, technologies, and so on (see sections An Analysis and Turning Up the Microscope). Pico-scale body dynamics are not the totality of first-order languaging behavior. The pico-scale is the smallest timescale that is relevant to the present discussion, though it is a crucial one for the development of appropriate descriptive and explanatory models of real-time interactional events. For present purposes, first-order languaging events are viewed as the integration of the following three timescales:

- N + 1: The (sense of) flow of situated social events (measured in seconds, minutes . . .);
- N: Microscale sayings and doings (roughly, the scale of "moves" and "turns" in discourse-analytical and conversational-analytical traditions though the explanatory framework of the present perspective is different from these approaches (measured in tens of seconds to seconds . . .);
- N − 1: Pico-scale bodily dynamics (milliseconds to tens of seconds . . .).

Body dynamics of individuals exist under nonlinear and nonequilibrium conditions. Bodily dynamics are able spontaneously to give rise to forms, drawing on the inherent tendencies of the dynamics as well as the functional capacities of bodies to affect other bodies and be affected by them (to move and to be moved). The variable natural rhythms of brain and body prompt and guide the cosynchronization of brains and bodies in interaction. On the other hand, persons also have the functional capacity to materially and intentionally affect each other's neural and bodily dynamics. This capacity results from the "variable intensive affects" (Deleuze & Guattari, 1988, p. 408) of, for instance, bodily feelings and intentional modulations of bodily dynamics. The functional capacity to affect each other harnesses material bodily dynamics and individuates them: dialogically coordinated interactions between persons get bodies to behave in new, unexpected ways.

Persons individuate other persons (e.g., caregivers individuate infants) by causally engaging with their dynamics so as to explore their feelings, motives, intentions, beliefs, and so on. In engaging with each other, we collectively pool

our expertise and knowledge about each other in ways that we value and consider worth tracking over space and time. We develop interactional and cognitive routines that link persons to each other. These routines give rise to heterogeneous assemblages (meshworks) in which persons affect and are affected by each other. In such a meshwork of persons (e.g., an interpersonal network), persons are meshed together by causal links and constraints—sociocognitive dynamics— that in time become fine-tuned, honed, and standardized in order to better adapt the individual persons to each other and to the whole (e.g., the interpersonal network) of which they are interacting parts. In the assemblage as a whole, the interpersonal network, individual persons play their part in interactively stabilizing the whole. This is where second-order language plays a leading role.

Second-Order Language

First-order languaging is, therefore, distinguished from *second-order language*. The latter is what most people, including linguists, think of as language. On the distributed language view, lexicogrammatical patterns are attractors—future causes—that guide and constrain first-order languaging. They are stabilized cultural patterns on longer, slower cultural timescales. A given population's repertoire of vocal tract and other gestures in first-order languaging evidences distinctive tendencies toward increased differentiation of the rich, high-dimensional dynamical properties of vocal tract gestures. Variations in the dynamical pico-scale properties of vocalizations potentially afford more and more semantically salient differentiations. This tendency occurs at the level of populations, not individuals, in the sense that interactional needs and motives at the population level (e.g., the interpersonal network), not the level of individual persons, may exert pressures toward the increasing semantic differentiation of vocal tract gestures. By the same token, the enormous variety—the many degrees of freedom—of vocal tract gestures at the individual scale is subjected to increased standardization or coding to conform to normative patterns and expectations at the population level.

Phonosemantic operators arise in populations of agents on the basis of the recurrent interactions between agents. These processes give rise to a possibility space of virtual patterns that may be detected/enacted in first-order dynamics. The dynamics of first-order languaging and their evolution are attracted to a culturally distributed set of topological invariants (attractors) involving processes of historical differentiation and recategorization of their dynamics into second-order lexicogrammatical patterns. The vocal tract gestures of speakers are sculpted by lexicogrammatical attractors on the longer, slower cultural timescale such that the biomechanics of vocal tract gestural activity is itself norm replicating. Lexicogrammatical patterns are virtual (quasi-causal) topological constraints that structure the dynamics of first-order languaging along cultural lines. Speakers

are norm replicators who aim to produce the "same" articulatory and auditory patterns as those of the others in the same interpersonal network. With respect to the enormous potential variation in vocal tract gestures at the individual level, lexicogrammatical patterns algorithmically compress historical–cultural information into increasingly standardized ("grammaticalized") and culturally salient differentiations at the population scale in response to the distributed sociocognitive dynamics of the interpersonal network.

It is not possible to locate such a semantically salient differentiation (in vocal tract behavior) if we specified a region in gestural space that could only be identified with reference to individual speakers. Cultural selection shapes the first-order dynamics of *populations* of speakers on longer timescales that are simply not apparent to individual speakers on the very different spatiotemporal scale of the first-order speech events in which individuals engage with other individuals. This is so because cultural selection requires a large enough population of individuals and a large enough number of vocal tract gestural events over a long enough timescale in order to operate and to see its effects registered. Thus, on the population scale, lexicogrammatical patterns are quasi-causal topological constraints—attractor spaces—that may have a variety of physical realizations. Such constraints are virtual, not actual, and are mechanism independent (De Landa, 2006, pp. 29–31). They are not, therefore, causal because they do not refer to the actual mechanisms operating on any given scale. Quasi-causal topological constraints structure a phase space of possibilities. Constraints of this kind may be utilized as a formal resource to describe the formal constraints that structure the state space of vocal tract gestural possibilities for a particular population of languaging agents.

In the view I develop in this article, languaging is a distributed, nonlocal organization of process on a diversity of timescales. It is one manifestation of the property of nonlocality that is exhibited by all living systems. In their languaging behavior, agents integrate and choose between flows of information coming from both the past and the anticipated future. First-order languaging behavior individuates along a time-locked trajectory that is drawn into its future by attractors in the process evolving toward both entropy reduction and an increase in differentiation and organization. This is no less so on the scale of the situated utterance activity. Languaging agents, like all living systems, are anticipatory systems; they depend on and are pulled toward future states (Bickhard, 2004; Rosen, 1985; Thibault, 2005b, 2005c). This is not prediction, which is in actual fact founded on backward-looking and mechanistic computational models, based on probabilistic choices from a predetermined set. Languaging agents do not predict each other's responses in the sense of computing all the possibilities and coming up with an accurate prediction. This would lead to an intractable computational bottleneck given the very short timescales involved for the necessary neural processing of the vast amount of data involved. Anticipation is modal rather

than predictive. Agents' dialogically coordinated utterance activity sets up and anticipates future possibilities for the further development of the utterance and its meaning potential as the utterance activity is pulled into its future along its time-locked trajectory.

First-order languaging dynamics pertain to the dialogic coordination and bodily synchronization of persons in occasions of interaction on the scale of local here-now social events. Lexicogrammatical patterns are population-level organizations of process that pertain to slower, longer, cultural timescales. Languaging agents learn to exploit first-order languaging dynamics in order to synchronize with, to engage with, and to affect and be affected by the bodily states and perspectives of others. This is observable interactive behavior that is also integrated with different kinds of nonlinguistic tasks and activities such that observable behavioral patterns begin to differentiate according to situation type. Once this occurs and spreads through a population of interacting agents, statistical patterns based on frequency of co-occurrence are detectable in populations of utterances as regular or habitual patterns of conditioned probabilities owing to sociocultural factors (Halliday, 1991; Hasan, 2009). Over time, these patterns are further solidified as conventional patterns of word co-occurrence as they become institutionalized. The resulting "grammatical" patterns are enforced as normative and conventionalized usage patterns. They are second-order patterns that are culturally transmitted and socially enforced as community standards through schooling and other formal and informal learning situations. In this way, the enfolded or intertwined trajectories of persons on multiple timescales get linked to and are constrained by second-order cultural patterns, norms, and values. Second-order language patterns are the emergent historical–cultural products of the conventionalized solutions devised by collectivities of interacting agents to the numerous problems of coordination that they are called upon to solve in the many different kinds of social activities with which language is integrated. The normative character of the second-order patterns also means that layers of reflexivity are semantically compressed into lexicogrammatical forms. The forms themselves thus provide access to mutually enforceable cultural expectations and patterns of behavior that help to define and stabilize a community of interacting agents (see sections on Lexicogrammatical Differentiators and Body Dynamics).

In contrast, form-based theories of language have postulated a number of levels of internal linguistic organization that essentially conflate different ontological scales and accordingly treat language as a local phenomenon that is instantiated either through a recursive process of sentence generation that is constrained by a biologically inherited universal grammar or as text that is the output of selections from a determinate language system. The architecture or design features of language are thereby assumed to be localized either in the brain of the individual or in text/discourse as codelike input/output mechanisms that mediate between the different levels of the architecture. Language is thus

seen as (levels of) interconnected formal abstracta that are linked to each other and to the world in an essentially arbitrary way. The study of language as disembodied formal abstracta fails to reveal the intentionally and affectively modulated character of its dynamics along with the essentially nonarbitrary ways in which languaging behavior is a flexible adaptation to the ecosocial environments with which it is coupled (Thibault, 2004a, 2004b). Second-order language is no less real than the dynamical properties of first-order languaging, though it exists on a different ontological scale as a set of virtual patterns—a structured space or contrast set of cultural possibilities that defines and constrains the sociocognitive interactive capacities and tendencies of a population of agents. Mainstream linguistics, on the other hand, has split "language" into real and abstract formal patterns (e.g., the distinction between phonetics and phonology) and focused on abstracta in constructing the theoretical object "language." The distinction between "first-order" and "second-order" refers in this account to the genetic dependence of second-order patterns and relations on first-order ones (see Ladyman & Ross, 2007, p. 243). This says something about the phylogenetic, ontogenetic, and microgenetic relationships of second-order patterns with respect to first-order ones.

In the following sections, I further explore various aspects of the relationship between the two.

Lexicogrammatical Differentiators and Value-Weighted Attunement to the World

Languaging behavior enables agents to attune to their environments (Gibson, 1983, 1979/1986; Reed 1996; Verbrugge, 1977, 1980, 1985). In Gibson's theory of perception, media such as light, sound, chemical composition, and so on, are structured by the environment as higher order invariants that organisms have become attuned to over the course of evolution and development. Gibson claims that the information available to perceptual systems of different species is specific to environmental events and therefore it is *not* arbitrary. In Gibson's realist theory of perception, knowledge is attunement to environmental structure (Gibson, 1979/1986). Dialogically coordinated utterance activity between individuals in talk is an extension of this fundamental principle, not something that is completely different. Lexicogrammatical structures are higher scalar patterns that function to differentiate or partition the environment in culturally salient, value-weighted ways. Embodied, environmentally embedded utterance activity therefore is a culturally shaped affordance for the reciprocal attunement to structure and for the guiding of interactants' actions, perceptions, cognitions, and feelings in relation to (attunement to) environmental structure either actual or virtual (see later). In the case of vocal tract activity, attunement is attained by a synergy between dynamic constraints on vocal tract and related (e.g., facial and

gestural) activity and the dynamical properties of the surrounding medium (e.g., air and light). The capacity of organisms to attune to environmental structure and to develop more and more sensitivity to particular aspects of environmental structure is how knowledge about the environment is produced.

Verbrugge (1980) points out that language affords virtual perception of events—off-line perception of virtual events rather than on-line perception based on the pickup of stimulus information:

> . . . language leads us to experience events, to view them from fixed and moving points of observation, to move about in social and geographic environments. These imaginal experiences are similar in quality to experiences we have in nonimaginal contexts. This mode of experience will be called *virtual perception* (*action*), on rough analogy to the virtual experiences of optics, in each case one moves from the real to the virtual. (p. 93)

Virtual perception through language is only possible if we take it as axiomatic that utterances do not encode representations of things, events, and so on, in the world. Utterances are not encodings of anything at all. Instead, the functional differentiations they articulate have the power to evoke apperceptions of implicit underlying networks of representations that are presupposed and potentially activated by an utterance operator in a given situation (Bickhard, 2004). Instead of encoding/decoding, utterances work on the basis of the very different principle of differentiation; accordingly, they are operators on situations. They differentiate and partition the environment in ways that may or may not work in a given interaction (see Bickhard, 2004; Thibault, 2005b, 2005c). They do so by acting upon and modifying underlying networks of representations that are presupposed in a given situation and by its conventions. Bickhard (2004) points out that utterances afford the possibility of the apperceptive maintenance and change of the representations of the agent's environment, including, most crucially, environments that are extended across diverse time and place scales—past, present, future. Utterances and the presupposed underlying networks of representations that they act upon and modify (add, subtract, transform) are intrinsically modal and anticipatory; they anticipate what could happen in the interaction and what would be possible in the given environment. They are modalized, value-weighted ways of orienting to aspects of situations at the same time that they anticipate future interactive potential (Bickhard, 2004; Thibault, 2005b, 2005c).

Utterances do not directly act on the minds of the participants in some situation. They do so mediately by acting on and transforming the situation and its conventions. In the process, they therefore transform the common understandings of the interlocutors (X tells Y that A, therefore, both X and Y know that A was told) and its reflexivities (both X and Y know that both know that A was told, etc.; see Itkonen, 2008, pp. 288–290). It is therefore

crucial to investigate the often subtle indicators whereby interactants know or discover which representations are presupposed, which conventions are in operation, and, therefore, the ways in which particular utterances contribute to the maintenance and/or change of the underlying presupposed representations and their reflexivities. These points are entirely lost sight of if one assumes that propositions are encodings of representations of states of affairs. This is so because encodings necessarily are of already represented states of affairs. If language were a system of encodings, as commonly supposed, we would not be able to cross epistemic boundaries. We could not, therefore, create new knowledge and understandings of the world. On the contrary, language and the differentiations it affords for semantically partitioning the world are a means of dialogically engaging with and transforming agents' reciprocal understandings—understandings that are the ontological ground of the social reality on which language operates.

Verbrugge (1980) points out that

> . . . language can evoke and guide virtual experience. It activates and constrains an attunement to structure, a structure that may be very different from that specified by the immediate physical environment. (p. 94)

This possibility has its basis in the ways in which language dynamics bias and alter perception (see earlier). Kuhl (2007, p. 112) has shown that infants develop a sensitivity to the distributional patterns of vowels and consonants that is crucially affected by social learning. Social learning (e.g., of language-specific syllabic patterns) in the wild is particularly robust as distinct from laboratory learning and affects perception. Kuhl (2007, pp. 111–113) cites and discusses further research (Kuhl, Tsao, & Liu, 2003; Kuhl et al., 2008) to show (a) that infants are strongly attracted to, interested in, motivated and aroused by infant-directed speech from adults in natural settings and (b) that social learning provides enhanced information in the form of, for example, gaze following and pointing about the objects of reference indicated by utterances in particular situations. Rejecting the empiricist and nativist positions that have characterized the debate on child language development over the past 5 or 6 decades, Kuhl argues that infants learn "statistically."

In ontogenesis, persons sensitize to the distributional frequencies of the articulatory and auditory patterns in the ambient language in relation to the events they covary with and this alters or biases their perception of ambient speech sounds. Kuhl (2007) demonstrates that infants show evidence of a "perceptual magnet effect" for native language variants. In other words, the native prototype functions as an attractor that attracts perceptual variants thereby "reflecting prototype learning and categorization" (p. 112). The information available to the individual crucially depends on factors such as the individual's maturational

level and the extent to which he or she has learned to attend to the relations between various factors such as, for example, a vocalization, a gaze vector, a finger point, and a given environmental object or event (Verbrugge, 1977, p. 376).

As Verbrugge (1980) puts it,

> Sentences are embedded in richly structured social and physical events, and in fact are essential *constituents* of many events. As a result, specific word and sentence forms are experienced as recurrent covariates of specific types of events—that is, in the child's social environment they occur consistently and uniquely as constituents of these events. (p. 94)

Thus, the reexperiencing of a familiar vocalization in some other context has the power to induce an apperception of the previously experienced event with which the vocal pattern had covaried. Statistical learning is relevant here too. Infants 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. It is in this way that infants learn how language structure attunes them to virtual experience.

Dialogically coordinated languaging behavior is based upon real-time dynamics that are structured asymmetrically in ways that bias perception (see also Cowley, 2008, pp. 328–331). This is how differentiators get set up with the functional capacity to partition agents' worlds. Asymmetry means that there are always differences of perspective, beliefs, goals, motives, values, felt bodily responses, the dynamical patterns selected, and so on. In the light of such factors, utterances can be seen as catalysts (Verbrugge, 1985) that act on agents with complex internal organization (cf. Harré's 1983 "primary structure") in ways that cannot be predicted by models based on one-way efficient causality (Thibault, 2004a, pp. 203–204). The same utterance and its dynamical properties may catalyze different effects depending on the persons involved, their state of arousal, the history of their interacting, the situation, and so on. Utterances qua catalysts have the functional capacity to either activate or to inhibit flows of affect, cognition, action, and so on, within and between agents. Human agents' actions are informed by beliefs, reasons, and motives. In interpersonally coordinated encounters between agents, beliefs, reasons, and motives are always, to varying degrees, asymmetrical. This is one of the drivers of interaction. However, asymmetry needs to be managed and coordinated.

The management and coordination of interactive events and their inherent asymmetries depends on processes that connect individuals, cultural patterns, norms, artifacts, and aspects of situations. Persons thus become jointly enmeshed in a cultural ecology in which neural processes prompt individuals to integrate on very rapid timescales perceptual processes with dynamical bodily

events such as vocalizations. The brain does not produce inner representations of worldside things and events. Rather, neural processes prompt dynamical, socially coordinated languaging behaviors whose dynamical properties have the functional capacity to bias or skew perception, affect, and cognition in value-weighted ways that give rise to semantic differentiators (Ross, 2007). Semantic differentiators have the functional capacity to catalyze and to direct flows of cognition, perception, feeling, and affect in and between persons in ways that are constrained by social and cultural norms. The visible and audible aspects of language (e.g., vocalizations, gesturing, facial expressions), perceptual activity, and selected aspects of situations are all components of a distributed cultural–cognitive process that depends on a history of coordinating activities with others.

In the following section, I consider some aspects of first-order languaging as a form of whole-body sense making.

AN ANALYSIS OF FIRST-ORDER LANGUAGING DYNAMICS: WHOLE-BODY SENSE MAKING

In this section, I analyze a brief example to show first-order languaging dynamics in action. Figure 1 shows the main phases in the episode to be analyzed. In the example, two groups of 9-year-old children—a boy's team and a girl's team—are playing a hypermedia storytelling game. The Appendix provides a verbal transcript of the entire episode (not analyzed here).[1] At this particular point in the game, an instruction on the computer screen prompts the group of boys to begin the task of telling a story about aliens who have landed on planet Earth. Initially, one of the boys, as in our example, responds to an instruction to describe the aliens before the task of constructing the story about them is undertaken. My analysis focuses on the first part of the boy's response to the instruction. At this particular point, the speaker (Boy1) turns to one of his teammates (Boy2) and during the course of his utterance he playfully performs a number of acts on Boy2—pulling his ears, twisting his tie, pulling his ears up—in order to make him into an alien.

Typically, standard discourse-analytical models of language remain stuck in ways of thinking and modeling based on homogeneous hierarchies and their formalized representations. These are more likely to reflect the homogenizing concerns and ambitions of academia but not the heterogeneous flows of matter–energy and their modulations that characterize the many domains of social life in which people coordinate actions, feelings, perceptions, and understandings through their languaging in the fulfilment of their projects and the realization of

[1]I thank Anthony Baldry for making available to me the videorecording of the "Aliens" episode on which the present analysis is based.

FIGURE 1 Main phases of episode, shown as nine microscale events.

values. For example, a discourse-analytical approach is likely to focus on Boy1's utterance as a mainly verbal response (e.g., a discourse move or conversational turn) to the verbal prompt supplied by the adult immediately prior to Boy1. Such a view leaves out a great deal of what is critically important here. Boy1's whole-body sense making (e.g., phonetic gesture + verbal patterns (wording) + shift **595**

in posture + rising pitch, etc.) initiates an exploratory dialogue of Boy2 as an environmental affordance involving feedback provided by Boy2's response (see later). The pico-scale dynamics of the interaction between the two boys shows a structurally coupled system in which Boy1 actively seeks out and monitors in the microtime of the unfolding utterance the possibilities that Boy2's body and clothing afford for realizing the values and solutions prompted by the initial instruction on the computer screen.

The two boys engage in coaction: they lock into and exploit each other's body dynamics, first-person experience, bodily feelings about the situation, verbal (lexicogrammatical) patterns, and aspects of the situation in order to create something that could not have been created by either boy acting alone. They concert their actions with each other and with relevant affordances in the situation. They do so around a shared project—describing the aliens and participating in the computer game in competition with their opposite numbers in the girls' team—and its values. The two participants integrate bodily dynamics, verbal patterns, material artifacts, feelings, perceptions, and cognitions. A great deal depends on how participants integrate bodily events—aspects of vocalizations, gaze, gestures, shifts in posture, and so on—that occur on very short timescales of the order of milliseconds, or what can be referred to as the pico-scale of embodied coaction (Steffensen et al., 2010; Turning Up the Microscope section).

The episode presented in Figure 1 is analyzed as consisting of the following nine microscale events (First-Order and Second-Order section) as follows:

1. Boy1 is leaning forward to attend to the computer screen (00.00.22). Boy2 also attends to the computer screen.
2. In response to the adult supervisor's follow-up prompt, "See what you can do about this," Boy2 directs his gaze toward the adult (00.00.23).
3. Boy1 begins to step back from the screen while he utters, "Well they sort of . . . " At the same time, Boy2 raises his outstretched hands and moves his torso left and right in a slight swaying motion as his gaze shifts to track Boy1's movement.
4. On the onset of "look like," Boy1 begins to turn his head and torso toward Boy2 (00.00.24).
5. Boy1 makes direct eye contact with Boy2, stretches both hands, and takes hold of Boy2's ears, which he gently pulls outward while uttering "this." Boy2's torso moves slightly to the left while tracking Boy1's body movement, his hands outstretched toward Boy1 and moving gently up and down. In response to his ears being pulled, he utters a soft, low growl ("ow") when Boy1 says "this," which is significantly lengthened to coincide with the action of pulling Boy2's ears (00.00.25–26).
6. Boy2 leans backward and bends leftward, increasing the distance between the two while continuing to utter the low growl. His right hand is held

upward and outstretched toward Boy1. Boy2's movement is in response to Boy1, who, while still holding Boy2's right ear with his left hand, takes Boy2's tie (00.00.26).

7. With his right hand, Boy1 pulls Boy2's tie upright while still holding Boy2's right ear with his left hands, as before. While doing so Boy2 utters "which" on the onset of the pulling of the other's tie. Boy2 returns to an upright position, right hand outstretched as before (00.00.27).
8. Boy1 pulls Boy2's tie down while uttering "their ties . . . " His left hand moves away from Boy2's ear to his collar while Boy1's right hand moves lower and then pulls Boy2's tie toward Boy1's left at chest level while uttering "go like." Boy2, while remaining upright, slightly adjusts his movement toward his right in response to the movement of the tie (00.00.28).
9. Boy1's gaze focuses and is held on the new position of the tie to his left as he utters "that," which is significantly lengthened. Boy1 smiles during this phase. Boy2's right hand moves toward Boy1's right hand which, still holding the tie, moves toward the center of his chest (00.00.29–30).

In the episode, we see that vocal, 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 (Reed, 1996, p. 85; Thibault, 2008, p. 318). Each such movement or co-orchestration of movements is directed toward the realization of a diversity of values that are potentially at play in the situation (Hodges, 2007). These values include playing their part in the game, providing an effective response to the task stipulated on the computer screen, playing the game in a spirit of collaboration between the boys, competing with the girl's team, providing entertainment rather than being serious, pleasing the adults present (supervisor and camera operator).

TURNING UP THE MICROSCOPE ON THE PICO-SCALE DYNAMICS: AN ANALYSIS OF THREE PICO-PHASES

In this section, using PRAAT analysis, I turn up the microscope, so to speak, and focus in more detail on a short sequence from the episode analyzed in the section An Analysis. The sequence consists of three pico-scale phases in order to show more clearly the fundamental role of pico-scale vocal and other bodily dynamics and their coupling relations, as discussed in the First-Order and Second-Order section.

Pico-Phase 1

Overall, the movement of Boy2 constitutes a response to and a recalibration of his own body posture in relation to Boy1 (00.00.24). The vocalization /w□l/ (*well*) lasts 0.638 s and is coordinated with the onset of Boy1's postural shift. The coordination of these factors serves to prompt a shift in attention away from the computer screen that involves both boys. Directing the attention of both boys to the new focus of attention that now unfolds is an integral part of the activity. What is more salient than the content of the utterance at this point is the surge of intensity that goes hand in hand with the rising pitch and the lengthening of the final syllable in the Boy1's articulation of this vocal unit. Pico-phase 1 vocal and other dynamics are shown in Figure 2. **Fig2**

Pico-Phase 2

Pico-phase 2 lasts 1.548 s and integrates the shifting postural orientation of the two boys, shift of gaze, and Boy1's hand movements toward Boy2's ears. In this pico-phase, Boy1's utterance (ðe s□rt əv l□k lajk; *they sort of look like*) is characterized by falling pitch (see Figure 3) and five peaks of intensity, corresponding to syllabic pulses (see yellow curve in PRAAT analysis, Figure 3). These pulses are synchronized on the pico-scale with other aspects of the coordinated bodily dynamics. For example, /l□k lajk/ (*look like)* is synchronized with Boy1's shift in gaze and upper body orientation to Boy2 as he begins to extend his hands towards Boy2's ears constituting a pico-scale event lasting approximately 0.684039 s. **Fig3**

Pico-Phase 3

Pico-phase 3 lasts 0.768 s (see Figure 4). The vocal aspect is characterized by rising pitch, a distinct surge of intensity, and pronounced lengthening of the final syllable of /ð□s/ (*this*). The vocal unit /ð□s/ (*this*) is, in the first instance, a phonosemantic operator that depends on sensorimotor knowledge. The initial consonant /ð/ in, for example, the English demonstrative pronouns *this*/*that* and in the definite article *the* is an articulatory event that is associated with a submorphemic unit, {TH}, that functions as a deictic operator of anaphora (Bottineau, 2007) in demonstrative *this*. Deictic operators are mediated by patterns of sensorimotor dependence, as discussed by Noë (2004). In the submorphemic unit, {TH}, this sensor-motor knowledge constitutes in experience the representation of the concept [IDENTIFICATION OF PRESUMED OBJECT WHICH YOU NOW REMEMBER] (Bottineau, 2007, pp. 55–56). The experience of deictically locating objects in the physical world that the speaker is talking about and which are presumed, from the speaker's observational **Fig4**

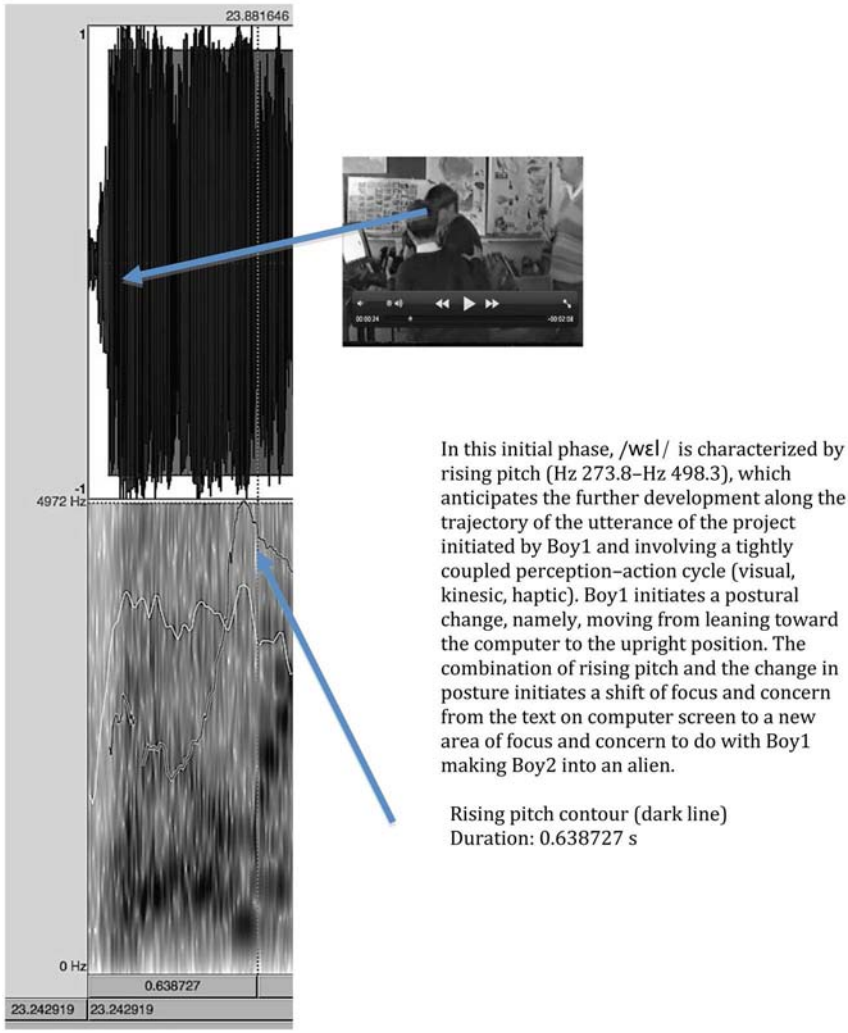

FIGURE 2 Pico-phase 1. (color figure available online)

perspective, to have already been indicated and therefore are available to the listener in working memory, depends on the way our sensorimotor grasp of such objects covaries with articulatory routines as speakers, listeners, and other observers co-orient to such indicated objects. This can be achieved by pointing, by gaze, and by other modes of bodily orientation. When we co-orient to 715

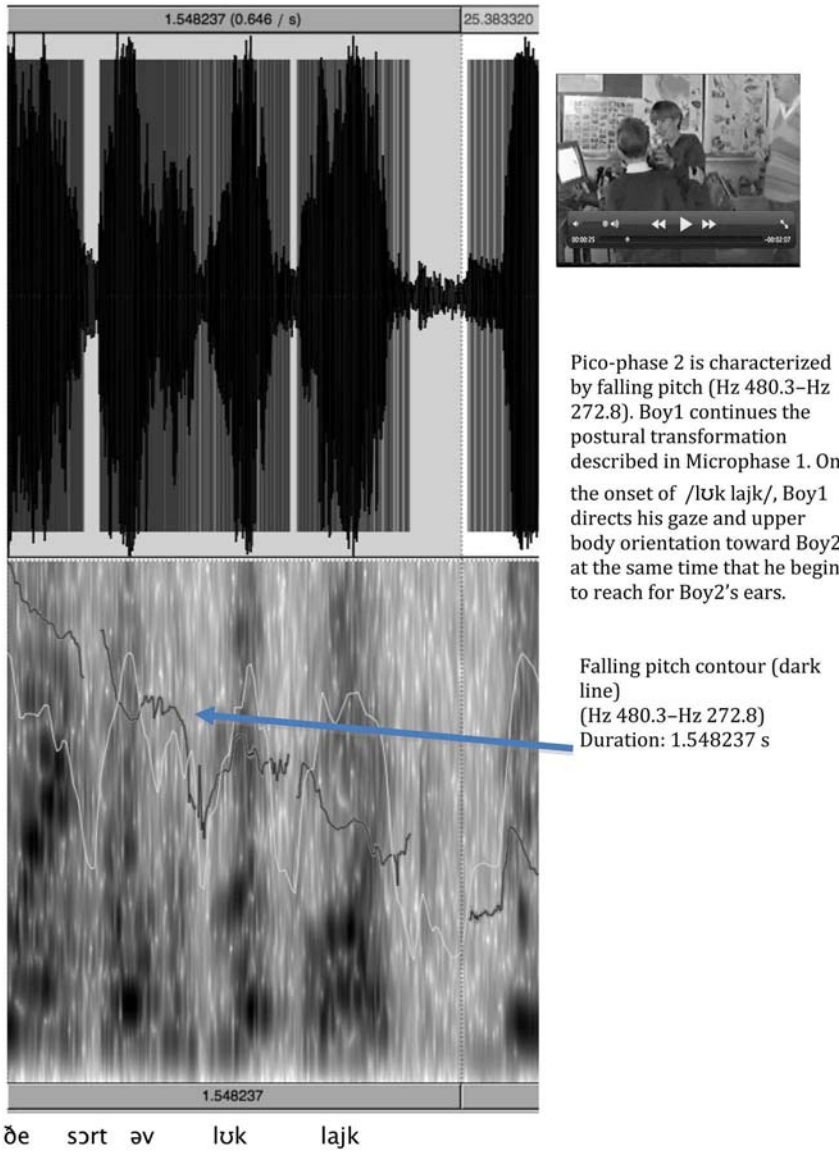

FIGURE 3 Pico-phase 2. (color figure available online)

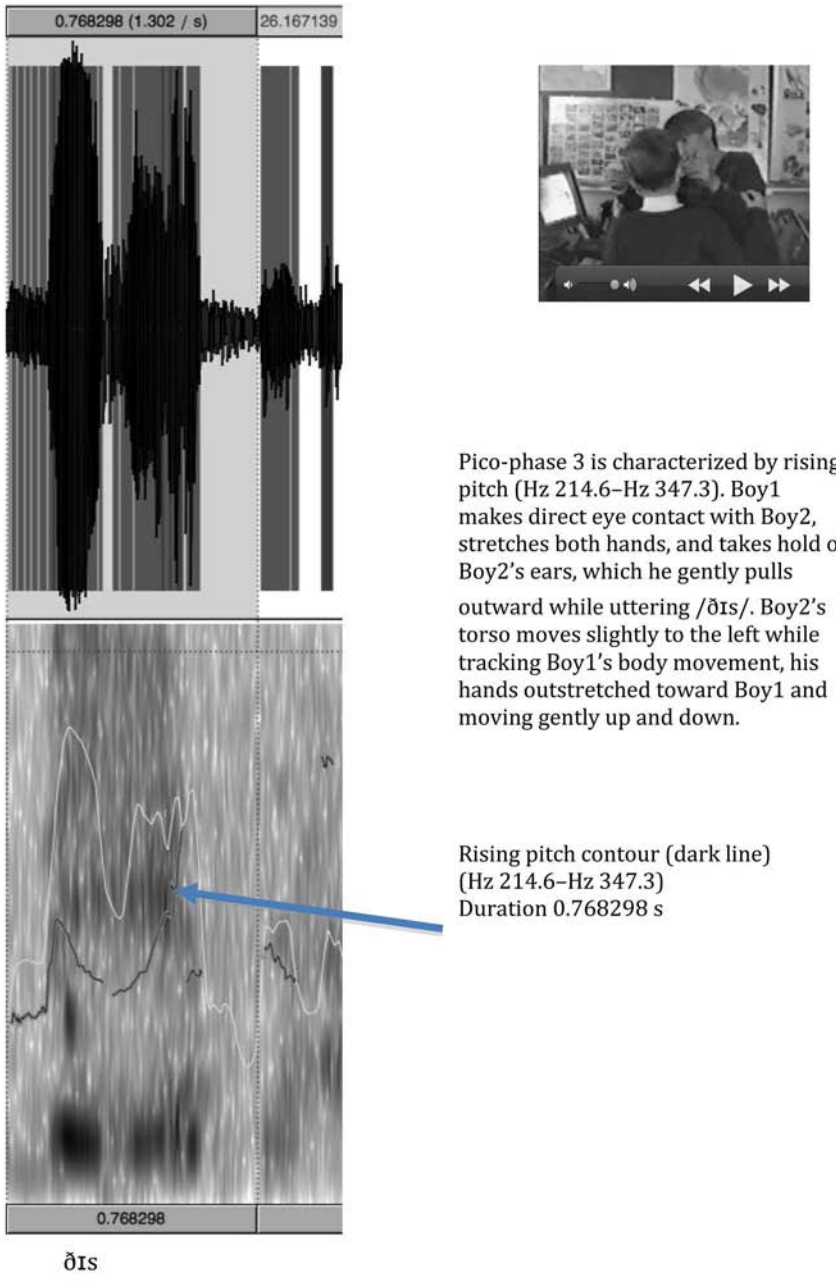

FIGURE 4 Pico-phase 3. (color figure available online)

someone's point, we have a grip on the ways in which the indicated object in the spatial purview, the direction of point, the orientation of the pointing agent, and so on, covary. This grip amounts to our grasping what it means for something to be pointed at and therefore brought to one's attention or made the current locus of cognitive processing. Demonstrative *this* is in principle no different. Following Bottineau (2007, pp. 55–56), demonstrative *this* is a phonosemantic operator consisting of the three submorphemic units {TH + I + S}, as shown in the parentheses. Our grip on it, so to speak, is our sensorimotor knowledge of what it means for the operator to invite the listener to attend to some object or event in the physical world as the current locus of cognitive processing and which is already presumed in the situation, that is, Boy2's ears, and so on, in the current example. The submorphemic unit {TH} can be glossed as [IDENTIFICATION OF PRESUMED OBJECT WHICH YOU NOW REMEMBER]. The entire semantic operator {TH + I + S} thus functions to assimilate {I} the semantic entity indicated by {TH} through the operation of anaphora to a further operation of present demonstration that is signaled by the final submorphemic operator{S} + NOUN/DEMONSTRATED ACTION. In the present example, {S} + Boy1's action on Boy2 = [WHAT YOU NOW PERCEIVE ABOUT THE OBJECT]). {S} indicates that the ensuing demonstration is, in the context, a novel one such that a virtual, extended perception of the aliens is enacted at the same time that this perception is assimilated to the anaphorically retrievable locus of current processing as seen by Boy1, that is, Boy2. Overall, the rising pitch, the surge of intensity, and the syllabic lengthening in /ðs/ (*this*), in concert with the other features of this microphase, enact a process of assimilation between the previously mentioned "they," which refers to the aliens mentioned in the text on the computer screen and the action that Boy1 performs on Boy2 in Pico-phase 3.

The three pico-phases analyzed earlier are phases in an intentionally and affectively modulated flow of co-orchestrated movement and other patterns toward an anticipated goal, namely, taking hold of and pulling Boy2's ears. Each of the three pico-phases is a component in a perception–action system that provides the system, comprising the two dialogically coordinated agents, with dynamical temporal information about the movement toward the anticipated goal. Listening to, watching, and moving in relation to such an unfolding event has the functional capacity to induce feeling states on the basis of our apperception of the rich dynamics of previously experienced movement patterns and the intentions they express. These rich dynamics give rise to intentionally modulated and affectively charged action trajectories that enact the essentially narrative forms of thought in terms of which movements and actions, and sequences of movements and actions, connect and interrelate persons, objects, goals, intentions, and so on (Reed, 1996, p. 174). The rich dynamics of movement dynamics means that the constant effort after meaning and value is always able to tap into more

information that has the potential to promote more and more cycles of thinking and imagining.

What then is the role of second-order verbal patterns? How are these related to first-order languaging? I focus on these questions in the following section.

LEXICOGRAMMAR AND THE CULTURAL MESHWORK

The verbal aspect of the utterance analyzed in the sections An Analysis and Turning Up the Microscope is an attributive clause (Halliday, 2004, pp. 210–226). It specifies a specific instance correspondence relation between the Carrier/Instantiator ("they") and the Attribute ("this + Boy1's pulling Boy2's ears"). In other words, the aliens referred to in the text on the computer screen are semantically construed as the instantiators of the property of standing in a relationship of visual correspondence ("look like") to what Boy1 does to Boy2's ears, tie, and so on. The lexicogrammatical pattern—the attributive clause—is a semantic differentiator, a cultural pattern that constrains the way in which the participants orient to and interpret the unfolding first-order languaging dynamics and those aspects of the situation that it connects to. The wording is not realized or encoded by linguistic expression. Rather, it is a virtual pattern emanating from cultural timescales that is integrated with the first-order dynamics in ways that facilitate coordination between persons and between persons and aspects of their worlds and interpretation. First-order languaging dynamics is a nonlinear, destratified flow of matter-energy and information. Lexicogrammatical patterns are attractors on longer, slower, cultural timescales that constitute basins of stability and becoming that may be embodied (not encoded) in first-order dynamics. In other words, the behavior of first-order languaging dynamics may be constrained by cultural attractors, seen as endogenously generated stable states or parameters.

In the example, (much of) the content is in fact being enacted or created by nonverbal means. The expressive bodily dynamics just i*s* the enactment of the "content"; it is not the expression of something else that is abstracted from the dynamics. In this case, the visual resemblance between the aliens mentioned in the computer text and what Boy1 does to Boy2 is to a large extent achieved by bodily (mimetic), not verbal, means. As the discussion in the Turning Up the Microscope section of Pico-phases 1 to 3 shows, each pico-phase exhibits a characteristic modulation of the unfolding event in relation to specific pico-scale dynamical features of the event. For many purposes, first-order languaging is all that is required in order to accomplish the task at hand and to enact representations without using higher level predicational language. Much of first-order language is explicable as the emergent outcome of intentionally modulated, dialogically coordinated agent-environment processes. The verbal aspect of the

activity plays its role in creating a link between two different foci of concern in the local representational topology in the situation, namely, the aliens previously mentioned in the computer text and what Boy1 does to Boy2.

In the first instance, we can say that the two boys participate in bodily based forms of intersubjective coengagement that are scaffolded by mimetic capacities (Hutto, 2008, p. 167; Zlatev, 2008). Mimetic acts such as pulling the hair, the tie, and so on, to imitate the aliens constitute and enact intersubjective engagement based on joint perception–action. A shared world is thus created in imagination. The two boys 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 (e.g., Boy2's hair) and items of clothing (e.g., Boy2's tie) are a part of our everyday, familiar embodied ways of acting and being in the world. These body parts and items of clothing function as intersubjective anchors because they afford nonarbitrary 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 (i.e., the aliens in the present example). Mimetic activities accordingly evoke absent objects and events on the basis of potentially shareable networks of sensorimotor associations that are held in working memory in the course of the interaction.

Infants learn in time to control their bodily dynamics, including vocalizing, in ways that determine perceptual input. Initially, the infant has little control over this input, sending random commands to the muscles. The infant 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., protoimperatives in infant protolanguage; Halliday, 1975). The infant learns to control vocal and other bodily behaviors in concert with others (and later solo) that establish a consensual domain of consistent motor–sensory relationships that are socially distributed in a given population of interacting agents. These motor–sensory relationships bias perception in value-weighted ways that lead in time to forms of higher order behavioral control. The use of the gesture for higher order control of this kind just is the meaning of the gesture. The achieving of control over vocal tract and other gestural activity means that the gesture can be intentionally directed and affectively modulated in order to get others to fulfill one's needs and wants. By the same token, gestures qua environmental events and the stimulus information they give rise to are shaped by interactional and cultural process in the activities with which they are meshed. They afford and are intentionally oriented toward opportunities for the seeking of cultural meaning and value (Reed, 1996, pp. 100–106). This is where second-order language enters the picture.

Take the example of the utterance *well they sort of look like this*. The utterance is grounded in sensorimotor learning and experience. The child, say, observes and learns that objects and events in the environment can be grouped together as resembling each other on the basis of perceived visual or other similarity. She learns that some objects and events can be distinguished from the rest of the environment in the local representational topology, that these can be perceived as affording relations of similarity, analogy, and so on, whereas others do not. She learns on this basis that the verb process "look like" can be used to construe a relationship of visual resemblance or comparison between different nodes in the local representational topology on the basis of perceived properties of objects and events that she observes. In time, she learns that the syllabic pattern /l□k lajk/ (*look like*) can be used independently of specific objects and events in on-line perception. The child learns that the syllabic pattern /l□k lajk/ (*look like*) is a public resource that affords visual comparisons with virtual entities that can be evoked on the basis of the apperception of previously experienced and remembered objects, events, perceptions, and so on that have covaried with the "same" pattern on similar occasions for different agents. A history of successful iterated associations between syllabic pattern and apperception in a population leads to a high degree of "interpersonal parity" (Fowler, 2010) such that the patterns of frequency of co-occurrence between syllabic pattern and apperception can be reliably detected in a population of utterances.

Lexicogrammatical patterns are conventional resources for compressing information about behavioral control strategies with which the patterns had covaried in previous encounters. Lexicogrammatical patterns semantically compress information about a whole set of behavioral control strategies and associated affordances that were learned in the first instance on the basis of on-line perceptual experience with which the words covaried. The typological-categorial or "digital" semantic character of lexicogrammar is well suited to organizing shared cultural worlds of virtual semantic entities on this basis (Ross, 2007). In off-line tasks, the lexicogrammatical pattern has the functional capacity to evoke and activate in imagination the entire set of previously learned strategies and affordances. This means that new situations such as the one the boys are faced with in the episode analyzed here are assimilated to what had previously been learned. The relevant functional capacities can be summarized as follows:

1. Lexicogrammatical patterns differentiate, in concert with others, a consensual domain of consistent sensorimotor relationships. They differentiate interaction possibilities that are presupposed in the selection of the given pattern. These possibilities involve implicit functional presuppositions concerning the presence in the local environment of the conditions required for the success (or otherwise) of the interaction. Language is normative in precisely this sense (Bickhard, 2004).

2. The patterns are stored in the agent as a means for setting up interactive potentialities even when not currently used. This is possible because of their functional capacity to compress information about the entire system of control strategies and affordances with which they covaried in prior experience.
3. When detached from on-line perception, such patterns have the functional capacity to operate on second- (and higher) order objects and events as a form of virtual experience. This leads, potentially, to an unbounded hierarchy of levels of knowing that can potentially be accessed.

Mimetic activities alone are both unreliable and subject to a high degree of indeterminacy of interpretation (Hutto, 2008, p. 269). Hutto is probably right to argue that pantomime is a "weak and highly ambiguous mode of communicating" (p. 269) whose limitations would push interactants toward "publicly established norms" of communication (p. 269). This explanation, however, provides no account of the mechanisms for the change. Lexicogrammatical patterns emerge and are culturally transmitted as normative solutions to the problem of coordinating understandings between agents. Lexicogrammatical patterns compress information that is historical and cultural. In dialogically coordinated interaction, this information gets integrated to the real-time bodily dynamics of agents.

Acculturated individuals are able to interpret lexicogrammatical patterns on the basis of culturally learned norms. They have learned to become "symmetric reasoners" (Gintis, 2009, p. 142) who are able to mutually access the same implicit representational topology. The participants who belong to a particular social group or interpersonal network have learned, or are expected to have learned, that in a given social situation-type the occurrence of an utterance of type-A means that contextually appropriate implicit representations are indexed and activated by the utterance. Participants are normatively predisposed as symmetric reasoners to orient to and select a given contrast set of representations rather than some other possible set as being appropriate to that situation on the basis of the information compressed in lexicogrammar. They are prompted to do so by experience and normative expectations grounded in a history of cultural learning.

In dialogically coordinated interaction, norms shape and direct expectations as to what one's interlocutor knows, believes, wants, and so on, and, therefore, what he or she is likely to do in a given situation. The compressed information in lexicogrammar that is potentially available to all the agents in the interaction is normative in this sense. Real-time interactivity (first-order languaging dynamics) is mutually choreographed such that agents mutually align their behaviors with cultural patterns. Bodily behaviors and displays are saturated with information emanating from cultural timescales. Individual motivations that control behavior are aligned with and reshaped by semantically compressed information patterns (lexicogrammar) that provide normative standards of interaction.

Lexicogrammatical patterns and categories (semantic differentiations) are, then, learned as normative cultural patterns that dialogically coordinate shared attention between participants by directing their interest and attention to a particular node in the local representational topology. The choice of a particular wording has the capacity to co-orient participants in the situation to a specific node in the local representational topology that is picked out as a local point of departure (i.e., the Theme) for the development of a message about the designated Theme (Halliday, 2004, Chapter 3). In resolving the reference of the pronoun "they" in the situation at hand, the attention of the boys is directed to a shared focus of attention about which something can be said. Thus, the aliens that were previously mentioned in the text on the computer screen are picked out and identified by the pronoun "they" as the current locus of shared attention and cognitive processing. The locally thematized referent is thereby specified as the locus of a message that can be predicated of it. Such messages can be jointly oriented such that the speaker directs the listener to attend to the particular aspect of the referent that is specified by the message developed about the theme in the local topology. With reference to our example, this process involves the following procedures:

1. Differentiate the current locus of attention by selecting a variable in the local network topology: the pronoun "they" anaphorically refers to the previously mentioned aliens in the instruction on the computer screen in this sense.
2. In so picking out and focusing on the aliens, the pronoun "they" is 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 prior discourse context (its reference is locally resolvable or interpretable).
3. The pronoun "they" is selected as the local point of departure—the Theme (Halliday 2004, Chapter 3)—for the development in the remainder of the clause (the Theme) of a goal-directed action structure.
4. 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 of presupposed representations that acts as a basis on attraction for the action trajectory.
5. The utterance induces apperceptions of representations of aliens by setting up in its transitivity structure a relationship of visual resemblance ("look like") between "they" and the demonstrative pronoun "this" and its coupling to the action Boy1 performs on Boy2 when he pulls his ears.

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, reconstituting them, and so on (see Bickhard, 2004). Apperception means that utterances are integrated with previous experience in memory in ways that get channeled in particular directions by the potential or likely contrast sets of representations that the utterance normatively 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 (potentially) 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 semantic differentiators (lexicogrammatical patterns) indicate more information than is available in the form of the utterance itself. They learn the normative conditions under which utterances are conventionally interpreted so as to activate and access the relevant systems of presupposed implicit representations.

BODY DYNAMICS, CONSENSUAL DOMAINS, AND EMERGENT UNDERSTANDING

How is understanding between the two boys displayed? Ideas of shared meanings or common codes will not take us far. Attunement is only part of the story. As the aforementioned analysis shows, Boy1 strikes up an orientation through his enactment of expressive body dynamics. His utterance activity elicits a response from Boy2. The dynamical patterns of Boy1's body dynamics signal to Boy2 (and to observers) purpose and orientation. As the analysis in the section Turning Up the Microscope reveals in greater detail, continuously changing body dynamics (vocalizing, movements, etc.) of agents have the capacity to affect, to move other agents. Boy2 responds and in doing so he signals his understanding by adapting his own dynamics to those of Boy1. In the segment analyzed, Boy2's vocal contribution is minimal. He adjusts his body posture and movements in microtime in ways that indicate a shift in orientation in response to Boy1. In other words, he is moved by and aligns with the orientation first struck up by Boy1. The result is co-orientation to an emerging consensual domain (Kravchenko, 2003; Maturana, 1970) as the dynamics and the perspectives of the two agents begin to mesh. Drawing on their experience of cultural norms, the two boys (and observers) adjust their perceptions of the first-order dynamics in ways that sensitize to verbal patterns (wordings). It is in this sense that they attune to verbal patterns as they draw upon and orient to cultural norms.

The verbal patterns so perceived have the capacity to resonate with cultural-semantic, thematic patterns (Lemke, 1983; Thibault, 1989) or "knowledge schemata" (Silverstein, 2004) and to provide access to them. Verbal patterns are not objectively present in the dynamics; they are not encoded "in" them, so to speak. Instead, understandings are affected by how we integrate a history of experience

of attending to such patterns, relevant normative expectations, body (vocal and other) dynamics, and second-order (virtual) patterns in order to semantically synthesize (Cowley, 2008, p. 330) a percept as a result of this integration. The perceiving of the interactive event has the power to evoke feelings and emotions on the basis of our experience of movement patterns (Schögler & Trevarthen, 2007, p. 291). The changing dynamics of the movement pattern, consisting of coupled shifts in posture and orientation and vocalizations (Turning Up the Microscope), engage participants and observers with temporally and interpersonally coherent and intentionally modulated mimetic displays that elicit "narrative" interpretation. As Schögler and Trevarthen (2007) point out, information in dynamical movement patterns "allows infants and adults to move in sympathy, communicating motives and intentions from brain to brain simply by moving together" (p. 291). In so communicating, we are moved to construe in the first-order dynamics verbal patterns that exploit and realize cultural meanings and values.

Verbal patterns are not pregiven formats or already constituted lexicogrammatical units that agents only have to retrieve from a stored system of options. Agents' understandings of interactive events are oriented to and shaped by mutual expectations and norms that provide cultural motivations and standards for evaluating patterns and integrating them to the dynamics in a given situation. Wordings are possibilities for (inter)action. They are "virtual multiplicities" (De Landa, 2002, p. 156) that are indeterminate yet progressively individuated through symmetry-breaking cascades in actualized occasions of linguistically mediated social interaction. An actualized social event is a historical individual. As such it "is but a specific solution" to a problem-space populated by virtual multiplicities (De Landa, 2002, p. 156) that inhabit cultural timescales but which interact with actualized events and are their constant shadow.

Construals of verbal patterns (wordings) are also affected by the results of previously constructed microsemiotic processes that are held in working memory. A concomitant tendency to minimize further constructive effort tends to favor the construction of lexicogrammatical schema that are topologically close to other schema that have been (recently) constructed and are currently held in working memory. As the episode under consideration shows, its dynamics are attracted to a particular region in the overall possibility space of English grammar: the patterns that are evidenced all cluster around the region in the lexicogrammar of English known, semantically speaking, as Attribution (Halliday, 2004, pp. 210–226). Consider in this regard the lexicogrammatical schema in the aliens example. Attribution as a general grammatico-semantic property can be seen as an overall topological semantic space to which the dynamics of the episode are attracted in the process of finding a solution to the problem that was posed by the instructional text on the computer screen, as shown in Table 1 and the Appendix. **Tab1**

Table 1 illustrates the three main subcategories of attributive clause that occur, though not necessarily according to their sequential ordering in the episode. The

TABLE 1
Attributive Clauses in Aliens Episode, Grouped as Three Subtypes

Carrier (Referential)	*Process: Attribution: Intensive: Phase: Perception ("Look," "Taste," "Like," etc.)*	*Attribute: Type-Category (Nominal) (Nonreferential: No Presupposition of Existence; Can Only Be Referred to in Relation to the Process)*	*Cotemporal Action*
They	(Sort of) look like	This +	Pulls ears and twists head
Their ties	Go like	That +	Pulls and twists tie
Carrier	*Process: Attribution: Possessive ("Have")*	*Attribute: Type-Category (Nominal): Body Part*	
They	Have	Three claws	
They	Have	A crinkled face +	Passes hand over face
They	've got	Green slime all over +	
They	've got	Very long tongues	
Carrier	*Process: Attribution: Intensive: Neutral ("Be")*	*Attribute: Type-Quality (Adjectival): Physical Characteristic: Size, Appearance*	
Their nose	Is	Long	
Their two front teeth	Are	Huge	
The rest of them	Are	Tiny	
All their hair	[Is]	Sticking up	

Appendix provides a verbal transcription of the episode, showing the sequential ordering of the utterances of the two boys. The point is that processes of **1045** semantic synthesis can induce functional topologies in the solution spaces of their microsemiotic construction in the real time of dialogically coordinated interaction.

CONCLUSION

The normative character of lexicogrammatical patterns (qua semantic differen- **1050** tiators) means that they specify information about classes of events that are

culturally distributed though never identical from one person to another in a given interpersonal network. Linguistic differentiators are, at the same time, both differentiators and integrators. In differentiating the environment in ways shaped by cultural norms, values, and associated semantic patterns, they also integrate first-order dynamics and the different experiences and perspectives of persons to culturally standardized verbal patterns in order to achieve some kind of harmony or equilibrium between the conflicting demands of differentiation and integration. A dialogically coordinated state of co-orientation between persons with respect to some event or class of event is always a local and provisional achievement in which two or more persons participate in a field of intersubjective resonance with each other's internal and external neural and bodily states and dynamics. In this way, persons in dialogically coordinated interaction can achieve a state of reciprocal intersubjective entanglement in each other's dynamics, leading to joint involvement in and attunement to some experience that has an intensified sense of intersubjective "nowness" when the subjective dynamics of two or more persons resonate as one.

Talk between persons is an intrinsically recursive and reflexive process built upon dynamic distributed processes of first-order languaging that are irreducible to second-order verbal constructs. The distributed language perspective argues that the latter can no longer be taken for granted as the essential architecture of language. Languaging is a distributed and heterogeneous biocultural resource that is spread over persons, environmental affordances, artifacts, cultural patterns, and values. It is grounded in the material interactivity and intrinsic expressivity of our bodies. This is the necessary starting point for rethinking "language" in the context of lived human experience.

ACKNOWLEDGMENTS

I acknowledge critically important discussions with and/or helpful advice from Stephen Cowley, Carol Fowler, Bert Hodges, Sune Steffensen, and Susan Stuart, all of whom have played a fundamental role in helping me to clarify my thinking about the topics addressed in this article. I express my heartfelt appreciation to all of these individuals for their inspiration, guidance, and support.

REFERENCES

Bickhard, M. H. (2004). The dynamic emergence of representation. In H. Clapin, P. Staines, & P. Slezak (Eds.), *Representation in mind: New approaches to mental representation* (pp. 71–90). Elsevier.

Bottineau, D. (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, 1*(2), 50–74.

Bråten, S. (2007). Altercentric infants and adults: On the origins and manifestations of participant perception of others' acts and utterances. In S. Bråten (Ed.), *On being moved: From mirror neurons to empathy* (pp. 111–135). Amsterdam, The Netherlands, and Philadelphia, PA: Benjamins.

Bråten, S., & Trevarthen, C. (2007). Prologue: From infant intersubjectivity and participant movements to simulation and conversation in cultural common sense. In S. Bråten (Ed.), *On being moved: From mirror neurons to empathy* (pp. 21–34). Amsterdam, The Netherlands, and Philadelphia, PA: Benjamins.

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

Cowley, S. J. (2006). Language and biosemiosis: A necessary unity? *Semiotica, 162*(1/4), 417–444.

Cowley, S. J. (2007a). Distributed language: Biomechanics, functions, and the origins of talk. *Emergence of Communication and Language, Part 1,* 105–127. Retrieved from http://www.tech.plym.ac.uk/socce/esg2006/cowley_distributed_language.pdf

Cowley, S. J. (2007b). How human infants deal with symbol grounding. *Interaction Studies, 8*(1), 81–104.

Cowley, S. J. (2008). The codes of language: Turtles all the way up? In M. Barbieri (Ed.), *The codes of life: The rules of macroevolution* (pp. 319–345). Berlin, Germany: Springer.

Cowley, S. J., & Love, N. (2006). Language and cognition, or, how to avoid the conduit metaphor. In A. Duszak & U. Okulska (Eds.), *Bridges and barriers in metalinguistic discourse* (pp. 135–153). Frankfurt am Main, Germany, and New York, NY: Peter Lang.

De Landa, M. (2002). *Intensive science and virtual philosophy*. London, UK, and New York, NY: Continuum.

De Landa, M. (2006). *A new philosophy of society: Assemblage theory and social complexity*. London, UK, and New York, NY: Continuum.

Deleuze, G., & Guattari, F. (1988). *A thousand plateaus: Capitalism and schizophrenia*. London, UK: The Athlone Press.

de Saussure, F. (1993). In E. Komatsu (Ed.), *Cours de linguistique générale: Premier et troisième cours d'après les notes de Reidlinger et Constantin*. Collection Recherches Université Gaskushuin n° 24. Tokyo, Japan: Université Gakushuin.

Fowler, C. A. (2010). Embodied, embedded language use. *Ecological Psychology, 22,* 286–303.

Gibson, J. J. (1983). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ: Erlbaum. (Original work published 1979)

Gintis, H. (2009). *The bounds of reason: Game theory and the unification of the behavioural sciences*. Princeton, NJ, and Oxford, UK: Princeton University Press.

Goldfield, E. C. (1993). Dynamic systems in development: Action systems. In L. B. Smith & E. Thelen (Eds.), *A dynamic systems approach to development: Applications* (pp. 51–70). Cambridge, MA: MIT Press.

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

Halliday, M. A. K. (1991). Towards probabilistic interpretations. In E. Ventola (Ed.), *Functional and systemic linguistics: Approaches and uses* (pp. 39–61). Berlin, Germany: Mouton de Gruyter.

Halliday, M. A. K. (2004). *Introduction to functional grammar* (3rd ed.). London, UK, and Melbourne, Australia: Arnold.

Harré, R. (1983). *Personal being: A theory for individual psychology*. Oxford, UK: Blackwell.

Harris, R. (1981). *The language myth*. London, UK: Duckworth.

Harris, R. (1987). *The language machine*. Ithaca, NY: Cornell University Press.

Hasan, R. (2009). Language in the processes of socialisation: Home and school. In J. J. Webster (Ed.), *Semantic variation: Meaning in society and socioliguistics. The collected works of Ruqaiya Hasan* (Vol. 2, pp. 119–179). London, UK, and Oakville: Equinox.

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

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA, and London, UK: MIT Press.

Hutto, D. D. (2008). First communions: Mimetic sharing without theory of mind. In J. Zlatev, T. P. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 245–276). Amsterdam, The Netherlands, and Philadelphia, PA: Benjamins.

Itkonen, E. (2008). The central role of normativity in language and linguistics. In J. Zlatev, T. P. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 279–305). Amsterdam, The Netherlands, and Philadelphia, PA: Benjamins.

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

Kravchenko, A. (2003). *Sign, meaning, knowledge: An essay in the cognitive philosophy of language*. Frankfurt am Main, Germany, and New York, NY: Peter Lang.

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

Kuhl, P. K., Conboy, B., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society, B, 363,* 979–1000.

Kuhl, P. K., Tsao, F., & Liu, H. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, USA, 100,* 9096–9101.

Ladyman, J., & Ross, D. (with Spurrett, D., & Collier, J.). (2007). *Every thing must go: Metaphysics naturalized*. Oxford, UK, and New York, NY: Oxford University Press.

Lemke, J. L. (1983). Thematic analysis: Systems, structures, and strategies. *Recherches Sémiotiques/Semiotic Inquiry (RSSI), 3*(2), 159–187.

Love, N. (2007). Are languages digital codes? *Language Sciences, 29,* 690–709.

Maturana, H. R. (1970). *Biology of cognition* (BCL Report 9.0). Urbana: University of Illinois.

Noë, A. (2004). *Action in perception*. Cambridge, MA, and London, UK: MIT Press.

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. New York, NY, and London, UK: Oxford University Press.

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

Ross, D. (2007). *H. sapiens* as ecologically special: What does language contribute? *Language Sciences, 29,* 710–731.

Schögler, B., & Trevarthen, C. (2007). To sing and dance together: From infants to jazz. In S. Bråten (Ed.), *On being moved: From mirror neurons to empathy* (pp. 281–302). Amsterdam, The Netherlands, and Philadelphia, PA: Benjamins.

Silverstein, M. (2004). "Cultural" concepts and the language-culture nexus. *Current Anthropology, 45*(5), 622–652.

Steffensen, S. V. (2008). Reassessing the systemic-functional subject: A hologrammatical interpretation. In N. Norgaard (Ed.), *Systemic functional linguistics in use. Odense Working Papers in Language and Communication, 29,* 782–806.

Steffensen, S. V., Thibault, P. J., & Cowley, S. J. (2010). Living in the social meshwork: The case of health interaction. In S. Cowley, J. C. Major, S. V. Steffensen, & A. Dinis (Eds.), *Signifying bodies: Biosemiosis, interaction and health* (pp. 207–244). Braga, Portugal: The Faculty of Philosophy of Braga Portuguese Catholic University.

Thibault, P. J. (1989). Semantic variation, social heteroglossia and intertextuality: Thematic and axiological meanings in spoken discourse. *Critical Studies, 1*(2), 181–209.

Thibault, P. J. (1997). *Re-reading Saussure: The dynamics of signs in social life*. London, UK, and New York, NY: Routledge.
Thibault, P. J. (2004a). *Agency and consciousness in discourse: Self-other dynamics as a complex system*. London, UK, and New York, NY: Continuum.
Thibault, P. J. (2004b). *Brain, mind, and the signifying body: An ecosocial semiotic theory*. London, UK, and New York, NY: Continuum.
Thibault, P. J. (2005a.) Saussure, Ferdinand de. In G. Ritzer, J. Stepnisky, & A. Wernick (Eds.), *The encyclopedia of social theory* (pp. 665–672). London, UK, and Newbury Park, CA: Sage.
Thibault, P. J. (2005b). 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, P. J. (2005c). 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, P. J. (2008). Face-to-face communication and body language. In K. Knapp & G. Antos (Series Eds.) & G. Antos & E. Ventola (Vol. Eds.), *Handbooks of applied linguistics (HAL) linguistics for problem-solving: Perspectives on communication competence, language and communication problems, and practical solutions. Vol. 2: Interpersonal communication* (pp. 285–330). Berlin, Germany: Mouton.
Trevarthen, C. (1987). Sharing makes sense: Intersubjectivity and the making of an infant's meaning. In R. Steele & T. Threadgold, (Eds.), *Language topics: Essays in honour of Michael Halliday* (Vol. 1, pp. 177–199). Amsterdam, The Netherlands, and Philadelphia, PA: Benjamins.
Trevarthen, C. (1992). An infant's motives for speaking and thinking in the culture. In A. H. Wold (Ed.), *The dialogical alternative: Towards a theory of language and mind* (pp. 99–137). Oslo, Norway: Scandinavian University Press.
Trevarthen, C. (1998). The concept and foundations of infant intersubjectivity. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 15–46). Cambridge, UK: Cambridge University Press.
Verbrugge, R. R. (1977). Resemblances in language and perception. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 365–389). Hillsdale, NJ: Erlbaum.
Verbrugge, R. R. (1980). Transformations in knowing: A realist view of metaphor. In R. P. Honeck & R. R. Hoffman (Eds.), *Cognitive and figurative language* (pp. 87–125). Hillsdale, NJ: Erlbaum.
Verbrugge, R. R. (1985). Language and event perception: Steps toward a synthesis. In W. H. Warren Jr. & R. E. Shaw (Eds.), *Persistence and change* (pp. 157–194). Hillsdale, NJ: Erlbaum.
Wheeler, M. (2005). *Reconstructing the cognitive world: The next step*. Cambridge, MA, and London, UK: MIT Press.
Zlatev, J. (2008). The co-evolution of intersubjectivity and bodily mimesis. In J. Zlatev, T. P. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 215–244). Amsterdam, The Netherlands, and Philadelphia, PA: Benjamins.

APPENDIX
Verbal Transcript of Aliens Episode

	Verbal Transcript	*Action*
1	B1: well they sort of	B1 turns head toward B2
2	look like	B1 touches B2's ears
3	*this (rising intonation indexing the action he performs on B2's head)*	James pulls Jonathan's ears
4	their ties go like that	B1 pulls B2's tie upward
5		
6		
7		
8		
9	and they have a	
10	crinkled	
11	I think they have	
12	a crinkled face	
13	and their noses are long	
14	Ah	
15		
16	and their two front teeth are huge	
17	and the rest of them are tiny	
18	and all their hair is sticking up	
19		
20	they look like this	
21		B1 pulls B2's hair
22	B2: they've got	B2 has been the butt of B1 but attempts to turn the tables by referring to B1's undisputed loquacity
23	very long tongues	
24	B2: and they've got	B2 puts his right hand on B1's chest
25	very long |tongues|	
26	B1: |tongues| Green slime	B1 looks down and then turns toward researcher
27		B1 looks at B2
28	B2: and they've got	B1 smiles
29	very long tongues	
30	R: let's try "next"	
31	Cliiiiiick	G1 has the mouse and clicks
32	B1: no	
33	B1: Funny Ashley: Ding	
34	B2: they talk like this	
35	B2:and they talk like this	B1 starts raising arm
36	Choral catlike noises produced by B1 and B2 imitating how the aliens talk	Two of the boys move arms imitating aliens' movements
37	B2: /rrr/	Chorus of "alien" noises by the boys
38	B1: /aH//Jau/	
	R: (softly) sssh shs	Continuation of same noises
39	B1: but they don't do that	
40	One girl starts reading from the screen "*Before on*"	

Note. Speakers = B1, B2, G1 (Girl1), R (adult researcher).