

Enaction

Toward a New Paradigm for Cognitive Science

Edited by John Stewart, Olivier Gapenne, and Ezequiel A. Di Paolo

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

Contents

Introduction vii

John Stewart, Olivier Gapenne, and Ezequiel A. Di Paolo

1 Foundational Issues in Enaction as a Paradigm for Cognitive Science: From the Origin of Life to Consciousness and Writing 1

John Stewart

2 Horizons for the Enactive Mind: Values, Social Interaction, and Play 33

Ezequiel A. Di Paolo, Marieke Rohde, and Hanne De Jaegher

3 Life and Exteriority: The Problem of Metabolism 89

Renaud Barbaras

4 Development through Sensorimotor Coordination 123

Adam Sheya and Linda B. Smith

5 Enaction, Sense-Making, and Emotion 145

Giovanna Colombetti

6 Thinking in Movement: Further Analyses and Validations 165

Maxine Sheets-Johnstone

7 Kinesthesia and the Construction of Perceptual Objects 183

Olivier Gapenne

8 Directive Minds: How Dynamics Shapes Cognition 219

Andreas K. Engel

9 Neurodynamics and Phenomenology in Mutual Enlightenment: The Example of the Epileptic Aura 245

Michel Le Van Quyen

10 Language and Enaction 267

Didier Bottineau

First MIT Press paperback edition, 2014
© 2010 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

For information about special quantity discounts, please email special_sales@mitpress.mit.edu

This book was set in Stone Sans and Stone Serif by Toppan Best-set Premedia Limited. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Enaction : toward a new paradigm for cognitive science / edited by John Stewart, Olivier Gapenne, and Ezequiel A. Di Paolo.

p. cm.

"A Bradford book."

"Based on an International CNRS Summer School organized by the Association pour la Recherche Cognitive (ARC), held from 29 May to 03 June 2006, Ile d'Oléron, France"—Text.

Includes bibliographical references and index.

ISBN 978-0-262-01460-1 (hc : alk. paper)—978-0-262-52601-2 (pb.) 1. Cognition—Philosophy. 2. Philosophy and cognitive science. 3. Cognitive science.

I. Stewart, John Robert, 1941– II. Gapenne, Olivier. III. Di Paolo, Ezequiel A.

IV. Association pour la recherche cognitive (France) V. Ecole d'été du CNRS sur les sciences cognitives (2006 : Ile d'Oléron, France)

3F311E495 2011

153—dc22

2010006930

10 9 8 7 6 5 4 3 2

10 Language and Enaction

Didier Bottineau

Linguistic behavior is orienting behavior; it orients the orientee within his cognitive domain to interactions that are independent of the nature of the orienting interactions themselves.

—Maturana and Varela 1980, 30

How can I tell what I think until I hear myself speak?

—C. S. Peirce

10.1 Toward an Enactive Outlook on Language

Describing language in the light of the enactive paradigm is a most challenging issue: language is to be reconsidered in terms of sensorimotor interactions with an environment in which both the individual and the environment are modified, in which not one, but several individuals are involved—an experience that is, all in one, that of the speaker and hearer at the instant of uttering or thinking; that of the child developing into an adult through social intercourse; that of the tribe turning to a full-fledged civilization; and that of the linguist interfering with his object of scrutiny by linguistic means. As a selection has to be made, this chapter will primarily focus on the immediate experience of languaging, and secondarily broach more general subjects like acquisition and evolution.

The proposals of the enactive paradigm (Maturana and Varela 1987) bring about a whole range of novel views that upset firmly established dichotomies: the subject and the object, the innate and the acquired, the interior and the exterior, the physical and the mental (Stewart 1996). This model destabilizes what Foucault (1966) called the epistemological basis of our knowledge, and challenges traditional scientific protocols.

In a recent book, Keller contends, in a purely Whorfian way, that the lexicon in use plays a major role in shaping a culture's thinking and

provides ready-made conceptual frames out of which any discourse is to be assembled, restricting intellectual creativity to recombining preestablished stereotypes (2006). Keller's view can be interpreted as an application to language of Varela's view of experience: once the subject/object dichotomy is revoked, the traditional distinction between individual experience of an external object and scientific "objective" scrutiny of an "objective" object is made redundant or at least questioned. The cliché has it that linguistic science is the only discipline to study its own object, language, using it as an instrument in the form of lexically marked concepts and discursively expressed descriptions and theories. But *talking* about language *using* language will inevitably *alter* the language, to the extent of making many words redundant, among which the words *language* and *word* themselves. Either linguists have *faith* in the existence of abstract entities such as *words, sentences, morphemes, nouns, verbs, subjects, objects, phonemes, structures, syntax, meaning, their evolution, the transmission of information, the acquisition of a language by an infant, representation, expression, communication*, and so on, and make a consistent use of them; or they reject them, with two possibilities: coining ever more abstract concepts, or looking back on experience to restore a phenomenological contract between the empirical basis and the theory. In a subversive fashion, the enactive questioning will consist in suggesting that all of those concepts correspond to no pre-given reality but are enacted by the linguist experiencing the linguistic environment he is studying and, above all, such as his own perception has reorganized it: is the NP/VP pair actually experienced, does it have any consistency other than on paper? A supposedly descriptive grammar is a self-ignorant formalism, and a deliberate formalism should be clear about the nature of the experience it enacts. For the same reason, it seems hazardous to postulate the existence of a conceptual architecture that gets revealed by linguistic forms in the process of communicative externalization, unless the individual and collective history and experience of the communicative process play a central role in the formation of the model¹. Linguistic models are based on an astonishingly poor samples of data with regard to the actual complexity of the multimodal processes experienced by speakers, and the enactive outlook aims at including this diversity in the modeling process.

The linguist is not the only one to enact metalinguistic experiences. Children are known to anticipate morphological systems by analogy (Slobin 1996): *to bring, brang, brung*. Some French learners of the English language produce such utterances as *he can played football*: their past experience in romance languages invites them to enact a universally inflected

infinitive, making the perception of the uninflected verb stem unacceptable in their system of anticipations, which causes them to override the evidence by superimposing a "rule." And in the absence of any available inflection in the English system, they will simply retrieve the homophony linking the infinitive and the past participle in French: *jouer/joué*; hence the English infinitive will sound like *played*. This cannot be explained in simple terms of analogy or overgeneralization. In so doing, the erring beginner is actually enacting an experience of the infinitive that is motivated by previously formed anticipations leading to decision and action upon passive perception; coming up with a personal pattern is so much easier than sorting out the perceptual chaos.

As for the speaking community, the way it enacts its own language is monitored by the standards set by the academic authorities and vulgarized by the schooling system: a language is an ensemble of morphological features (a lexicon, an inflectional system, morphosyntactic patterns) governed by explicit rules; one may learn and obey them and/or create new ones characterizing rebellious communities rejecting the authority of parents and institutions symbolized by the norms. For example, French *verlan* (from *l'envers* "reverse" reverted into *vers-l'en* > *verlan*) consists in coining new words by (1) inverting the order of syllables, (2) deleting the nucleus and coda of the final syllable (apocope), and (3) opening the vowel in the new final syllable: *femme* > **mefa* > *meuf*. Teenagers are outraged when their parents violate their own symbolic authoritative status by using such words or even worse, coining new ones as I did (for "priest," *curé* > **récu* > *reuc*). As soon as a *verlan* word is firmly established and perceived as a standard, it may be reinverted to renew the rebellious shift: *Arabe* > **Beara* > *Beur* > *Reubeu* (with epenthetic *e*). In some cases, the young reinvent words that had long existed in previous generations and fallen out of use, which they will never acknowledge.

In the introduction to his recent *Grammaire du Gallo*, whose title refers to a romance regional language spoken in the eastern half of French Brittany, Patrik Dreano (2006) declares that his initial goal was merely to collect instances of how he and the natives of his region actually spoke, on the basis of recordings of spontaneous conversations and storytelling by native subjects of varying generations, dialects, and origins. However as the corpus developed, he became aware of an emerging general consistency and diversity in the data, which made him turn the initial project into an extensively documented comprehensive grammar including the lexicon, phonology, morphology, syntax, and expressivity; scavenging through the data resulted in enacting the language, that is, turning the

dim awareness of the *parler* into the proud identity of an actual *langue* by producing a grammar, a seminal representation that inevitably sets itself as a reference, and therefore a standard. From a sociolinguistic viewpoint, a *langue* may be considered as a heterogeneous, multidialectal *parler* that has turned into a unified, consistent system through the selective and organizational decisions required by an academic description, which leads to institutional prescription and the emergence of a consensus mustering all social classes and domains of activity; a *langue* emerges as an autopoietic dynamic system (Maturana 1980) in the process of detecting and ruling its own collectively constructed experience. It is not an individual entity or self-organized system (Oudeyer 2006) endowed with a transmissible genetic identity, but it does form a structured dynamic phenomenon involving a boundary or “membrane”—the general coherence of lexical and morphosyntactic patterns, the phonological and prosodic system as homogenized human behaviors—that delimitates what is recognizable as belonging to the system (sounds, words, correctly formed sentences) or not. Dreano’s introduction is remarkable in that it autobiographically summarizes the enactive metalinguistic experience that he went through in a personal and historical shortcut, along with the growing awareness that by publishing the book, he was actually spreading the word throughout the potential community; and indeed Gallo rural native readers are surprised when, leafing through the book, they realize that their own *parler* does comprise as many consistent and expressive features as French, all the more so, as they would normally not have regarded Gallo as a language.²

10.2 The Act of Speech, the Interlocutors, and the Linguist

Modeling language in the light of the enactive paradigm entails some methodological choices that highly constrain the heuristic approach. First, one must eliminate all theoretical prejudice regarding the general categories: the traditional objects of linguistics (*langage*, *langues*, *parole*, grammar, etc.) cannot be taken as scientific objects of scrutiny as long as they have not been redefined in terms of dynamic sensorimotor experience (McNeill 2005): language, like any form of living cognition, is embodied (in the sense of Erard 1998, Wilson 2002, and Ziemke 2003). Strikingly enough, the very notion of *experiencing language* is commonly found among literary writers or critics, (especially of poetry and drama), pathologists, and psychoanalysts, but much less so among linguists (Gardiner [1932] 1951). Second, this requires an empirical selection of how language is actually

experienced by speakers in real life: the act of speech in all its forms, *languageing*. This heuristic funneling will not restrict the domain to the empirical data, but all abstract general concepts will be rooted in the dynamics of living. And third, the linguist must hold fast to her own position: languageing is an interactional process. When he thinks he observes his object, he is effectively interacting with the linguistic environment, that is, (1) playing the role of the interpreter of the linguistic data and (2) *languageing* a discourse about it using his own lexicon and grammar. So this section is about how the immediate experience of languageing—the act of speech—may be enacted by the linguist.

At least three basic types of languageings may be experienced (an inventory that is not to be found in the introductory section of manuals and treatises): *introverted languageing*, experienced in the form of verbalized mental discourse; *extraverted vocal languageing*, or speaking; and *extraverted manual languageing*, or writing and signing. Other marginal communicative forms of manual languageing, which will not be described here, involve various sensorimotor experiences such as whistling,³ beating drums, and sculpting smoke clouds, all of which have possibilities and constraints of their own.⁴ Foot languageing does not seem to ever have emerged,⁵ but one may imagine it as potentially experienceable in the form of tap dancing. The preceding typology will be used because it is convenient, but it is inaccurate for two reasons:

1. Languageing, like any experience, is multimodal, and cannot be reduced to any one of the media involved. In vocal languageing, the voice is experienced as the prevailing feature, the one playing the major role in the elaboration of meaning. But this does not preclude the role of head, body, and hand gestures, and spatial positioning and eye contact, known as the kinetic nonverbal modalities of conversation.⁶
2. Those categories must be mapped against the interactional pattern involved. Intraverted languageing is executed by a single languager. Extraverted vocal and sign languageing is simultaneously experienced and concerted by at least two languagers mustering in the same perceptual environment in space and time, which can be artificially expanded by technological means (Bruner’s *amplifiers*; see Bruner 1983), both spatially (telephone, etc.) and temporally in the case of recordings. And finally, graphic manual languageing involves a stabilized alteration of the material environment, a piece of writing, that can bridge a spatial and temporal gap between the writing and the reading, no matter whether the latter is carried out by the same person or by a different one.

The preceding elements are of course trivial and are commonly overlooked. Yet they show how heterogeneous the languaging experience is, how difficult it is to construct a unified theory of language out of this diversity. Most models happen to be constructed out of an extremely narrow, if not adulterated, selection of empirical features. When speech is said to *express one's thoughts*, the stress is laid on the communicative role of voicing or writing, but nothing is said about why intraverted speech requires languaging too and how it reflexively *impresses* individual thinking.

10.2.1 Acoustic Languaging

Vocal utterance is a process in which a "speaker" embarks upon affecting the atmospheric environment by somatic means: the air's molecules are made to vibrate following oscillatory patterns that can be controlled by articulators including the vocal chords, the tongue, the different parts of the oral tract, and, occasionally, the nasal cavity. This eventful action is perceived by all human beings present: (1) by the hearer(s), whether they are the addressees or not (Coursil 2000), who will match the auditory inputs with all relevant other experiences items acquired through perception (the speaker's general appearance and attitude, the situation, the psychological contexts, the cultural references and knowledge, whether shared or not); (2) by the speaker him- or herself, who happens to detect his or her own activity both directly and indirectly: by the tactile proprioception of his or her muscular activity, and by the auditory perception of his or her own acoustic output. In short, if the speaker causes the hearer to construe meaning, the speaker will also cause him- or herself to follow the same procedure, so that voicing amounts to a physical process of semantic mental co-alignment. "Communication" should not be modeled as a one-way arrow as in Bühler's ([1934] 1990) and Jakobson's empirically selective models, but rather as a retroacting radial propagation that will constantly affect the consciences in presence and be reprofiled in real time according to transitory effects and actions: a binary structural loop in which two living bodies' cognitive experiences are alternatively controlled and synchronized through somatic interference with the shared medium.

In other words, linguistic cognition involves cortical, muscular, and environmental dynamic events shared by individual beings in a continuous experiential shell forming a *social body*. In this process, vocal interplays synchronizing mental dynamics amount to forming *complex transitional synapses between conscious selves at the intercortical level using acoustic signals as transmitters across the atmospheric medium* (Bottineau and Roulland 2007), with a dramatically enhanced level of complexity: whereas the "physics"

of the chemical synapses is simple (one given chemical transmitter with a stable structure), that of the vocal synapse is infinitely versatile (lexical, morphosyntactic, prosodic variation); whereas the semantics of the former is simple (a binary 0/1 alternative between firing or no firing), that of the latter is infinitely open.

This synaptic coordination of mental dynamics by linguistic verbal processes causes both individual and collective consciousness (experience, knowledge, culture, concerted actions, *civilization*) to emerge and gives cognitive-biological substance to the notion of social body, whose spatial boundaries and temporal extension and continuity will vary with the profile of the communicative pattern (number of participants, technological amplifications, synchronization of the alternations or recordings). A large-scale social body (in time and space) such as a linguistic, ethnic, religious, cultural, tribal social body may be enacted as a continuous entity precipitated—in the chemical sense—by discontinuous speech acts experienced by each individual member joining in the incorporative, emerging process.

10.2.2 Optical Languaging

If one is to define a piece of writing rigorously, including the way in which it is experienced by the writer and the reader (including the chapter-writing linguist), a text is a light-reflecting (or emitting) surface whose visible profile has been altered by manmade gestures (from carving to dictating through writing and typing) in a way that can be perceived visually and converted into a *reading*, that is, a simulation of what actual voicing might have been had the human participant shared the same medium in space and time.⁷ Whereas talking affects the medium the time that the voicing lasts, writing affects the environment more durably by introducing a permanent alteration and makes it possible to extend, amplify, and multiply synaptic connections between consciences over spaces and times spanning extensions far beyond the reach of human voice and even individual life: if talking is basically local, writing is universal, although both can be amplified (Bruner 1983) in space (telephone, television, radio, the Internet) and time (oral and written recordings and multiple copies).

Writing systems correlate varying aspects of recorded experience with the act of reading: letter and syllabic systems associate the visual bottom-up input with a vocal prediction of the acoustic counterpart, as in the reformed syllabic alphabet of present-day Korean, a vocal and cultural writing that causes the reader to enact human vocal output in the first place, and, from there, the semantic counterpart in general experience;

ideogram systems (Chinese keys and first-generation Egyptian hieroglyphics) associate the optical unit with the experience of the "object" and not that of any human vocal output: the writing system is targeted at the nonvocal part of physical experience. Some hybrid systems combine the two approaches like Japanese ("kanji" ideograms for lexical units and compounds vs. hiragana syllabic characters for grammatical morphemes, with kanji allowing for contrasted historical readings anchored in the Chinese and Japanese vocal traditions known as *kun* and *on* interpretations); lexical reference to natural experience is paired with the optical marking of the natural sensing, the sensory side of the enactive interface, and grammatical combination is associated with the graphic marking of vocal, cultural output (the motor, controlled side of the enactive interface). For English- or French-reading learners of written Chinese and Japanese, the challenge is to explore those undocumented, exotic enactive interfaces. Some writing systems based on letters are *linear*, like the roman alphabet, and tend to blur the borders between syllables; other nonlinear systems focus on the syllable, either phonetically (hiragana) or semantically (Chinese); some systems assemble syllable around vowels by peripheral nonlinear extensions (Mongolian) and some linear letter-based systems underline syllables by masking their central component, the vowel (Arabic). So to model a writing system is both to decipher the nature of the vocal or nonvocal experiences the writing gesture is connected with and to understand how the very format of the gesture (in terms of linearity and fragmentation) is reciprocally intertwined with its meaning.

As a result, reading is never exclusively linear, even in linear systems, and involves the mapping of visual perception with a system of predictions (involving lexical selections, grammatical order, etc.) acquired through experience and training. Moreover, the reader experiences meaning to emanate not from a physically present biological agent (the speaker), but from a stabilized material object, the piece of writing, left by some remote and dimly envisaged agent, the writer. Meaning acquires material autonomy, the word becomes a kind a fetish, nonhuman sources may be envisaged (with far-reaching consequences for religions, especially if based on partially cryptic writing systems concealing the vowels and calling for interpretation), and humans are willing to fight over objective notions and truths crystallized by the written word, forgetting that the original word is no more than a mind-guiding piece of voicing. Stabilized pieces of writing lead to the formation of versatile social bodies with spatial boundaries that materially fluctuate in time with personal connections with the support, but that are enacted by the community as stabilized assemblies

known as *readership*. Barring oversimplification, writing is essentially *public* and speaking *private*, and both can be extended by technological artifacts and social rituals (public meetings).

Interestingly enough, the English lexicon reminds the reader of the sensorimotor experience of writing in the very morphological structure of the corresponding words. In the verb *write*, of Germanic origin, the onset /ɪ/ is spelled *wr-*, a consonant cluster that has long been identified as an *ideograph* connected with the notion of a circular movement of torsion, as in *wrist*, *wring*, *wrath*, *wriggle*, *wrestle*, *wrought*, *wreck*, *wreak*, and maybe *work* and *word*.⁸ And in the onset of the stressed stem of *scripture*, the consonant cluster *scr-*, of romance origin, is connected with a phonosymbolical imitation of the scratching sound of the writing as in *screech*, *scratch*, *scrub*, *scrawl*, *Scrooge*, *scrum*, *scribble*, etc. (some of them of Germanic origin); those words are morphologically categorized into a selected and relevant sensorimotor class of experience, and writing is mainly categorized as a type of movement and marginally as a movement-caused sound. The two enactive traditions were introduced by two of the ethnic and cultural contributors of lexicon in English (the Romans, the Saxons) and hybridized, randomly or coherently, by their mixing offspring. Semitic languages are known to display feature matrices of consonant clusters motivated by the common multimodal experience associated to the voicing and the notion (Bohas 2006). Whether a neural theory substantiates these correspondances is another debate (Feldman and Narayanan 2004; Gallese and Lakoff 2005).

10.2.3 Reflexive Linguaging

Inner discursive, linguistic thinking is a serious challenge. It is essentially *intimate* languaging (as opposed to *private* and *public*), in that it is not meant to transitively affect any identified target other than one's own self, reflexively. It consists in interpreting one's own verbal production and actively become aware of some construed, enacted representation—in the dramatic sense, not in the diplomatic one—that could not be achieved through any other means than languaging. The difficulty is that the function of transitive speech, externalizing one's preestablished thoughts, is in blatant contradiction with that of reflexive speech, making up one's ideas by verbal means; this paradox is best summed up by the following quotation: "Language is the formative organ of thought. Intellectual activity, entirely mental, entirely internal, and to some extent passing without trace, becomes through sound, externalized in speech and perceptible to the senses. Thought and language are therefore one and inseparable from

each other" (von Humboldt ([1836] 1988, 54). Do Austin's (1962) and Searle's (1969) classifications of speech acts apply to introverted speech? It would seem that they do not, as they are based on a communicative conception of discourse requiring a clear-cut distinction between an addresser and an addressee, even in the case of soliloquy.

As regards action, the experience of languaging can be located between two extreme situations. One is when the thinker speaks with himself, that is, fully voices the whole utterance, causing himself to recapture it through proprioception and audition and interpret it. The other is when the speaker totally refrains from any single detectable motor action during the process of mental wording and contents himself with an intimate experience of what the utterance would sound like if it were actually voiced, which includes word and sentence stress, tone units, melodic patterns, and so on. And in between, the commonest of all, is the case when a speaker unwittingly "motors" some barely detectable movements of the lips, jaws, tongue, and other members ordinarily involved in private, vocal (multimodal) languaging. Admitting this simple experiential insight into the empirical basis of languaging has far-reaching consequences.

First, it implies that in all cases, intimate thinking is a sensorimotor experience in its own right, no matter how and how much the private sensorimotor interferences with the perceivable environment are inhibited. Recent neural imaging has revealed unexpected connections between perceptual and motor areas even in cases when vocal production is not involved, and recent studies in infants have shown these connections to be active long before the articulatory system is operational: regardless of whether one does or even can speak, the processing of acoustic signals in linguistic terms is correlated with the controlling of the muscles that will produce them. Should this be interpreted as a hint of the inborn character of the coupling? An alternative hypothesis is that in its experience of adult languaging the infant jointly sees and hears adults speaking and develops relevant sensorimotor matchings acquired through perception, and that these pave the way for future production without requiring immediate practice at the time of their emergence. Vygotsky (1962) hypothesized (against Piaget 1972) that the child's egocentric speech, far from vanishing with development, would functionally and structurally split between specialized internal discourse and external speech. This illustrates the difference between a direct feedback effect and an enactive structural loop, in which a relevant system of anticipated decision and action is synthesized out of a selective and reorganized set of empirical data that may not involve the execution of the action in question.

Second, it entails that in the languaging experience in all its forms, even the innermost intimate pondering, the elaboration of meaning can never be envisaged out of the realm of bodily action: "linguistic cognition" coincides with the sensorimotor experience of the voicing and hearing of the speech act, and strictly confining this structural coupling within the motor-tight intracortical neural compartment turns out to be virtually impossible—let the reader just try to look as if he or she were not reading these lines while still doing it and he or she will feel the impossibility of not "somatizing" somehow the innermost neural dynamics of languaging. Languaging is embodied. This calls for more explanation.

At any time in the history of human thinking about language, spanning from Presocratics and Cratyle to present-day models and including other traditions (Arabian, Indian, Chinese), there has always been what may be loosely called an *enactive bias*:⁹ yes, the lexicon and the morpho-syntax do somehow derive from pragmatic experience and reflect sensations phonosymbolically or organize one's mental categories according to experienced entities (the Sapir-Whorf hypothesis: Sapir 1951; Whorf 1956), actions (Langacker 1987) and encounters or confrontations involving force dynamics (Talmy 2000), all of which are recorded in lexical forms, cognitive patterns underlying semantic representations (Pottier 1992), abstract grammatical processes (Guillaume 1964), elocutive or interlocutive configurations (Givon 1994; Coursil 2000; Douay 2000), lexical distributions (Whorf 1956), and culturally formed phrases and metaphors (Lakoff and Johnson 1999). The idea here is not to deny that somehow language (as general know-how) and languaging (as a singular doing) reflect bodily experience of the world by the subject, or to suggest that all this is not commonplace. It is to suggest that embodiment *stricto sensu* is not the binary, mental, symbolical copy of experience, and that metaphors do little to illustrate embodied languaging—they underline how the mind acknowledges the body's personal life in a kind of dialog: metaphors underline how general knowledge is formed through action (cf. the role of practical or sensorimotor intelligence in children's cognitive development according to Piaget 1972), but they do not show how language in general, even if not metaphorical, is a "cognogenetic experience." By definition, languaging is embodied in that experiencing the sensorimotor coupling of voicing and hearing is instrumental in the construction of thinking at every level of the experience—intimate, private or public, and this is what the enactive paradigm has to bring in beyond the traditional embodiment described by the symbolic cognitivist paradigm.

10.3 Meaning

Languaging (the act of speech) is understood here as an intimate, private, or public sensorimotor process, *la parole*, enabling all participants to construct some form of mental event or scene. This section is about the kind of meaning achieved through languaging, an enactive experience that could not be achieved by any other means. Before turning to the respective roles of lexicon, morphology, syntax, and prosody, a first approach to meaning is required.

10.3.1 Languaging as an Alternative Medium

In private languaging, an addresser causes an addressee to come up with an original piece of experience. Not that the "idea" should be new or original—*Buses are always late, aren't they?*—most of our daily production is cliché. A speaker will make a hearer become aware of something (real or not), that is, enact a piece of sensorimotor experience through languaging, either because he or she has seen from the hearer's attitude that the "real fact" has not been taken into account and does not bring about the expected intellectual, emotional, or pragmatic reaction (Sperber and Wilson 1995), or because the speaker computes that introducing this piece of experience might serve his or her purposes, directly or indirectly (seducing someone else), and so on. Speaking does not *refer* to the world; it *causes an experience* that happens to coincide or not with the narrow situation or the larger reality such as it is enacted, and has to be mapped against the environmental medium, including the psychological environment.

As long as the hearing or reading of the sentence lasts, the interpreter's awareness, will, and action, in one word, dynamic experience, is entirely concentrated on plotting out (neither *computing* nor *representing*) the sensorimotor enaction gradually emerging from the lexical inputs as they are matched with one another, following a constructional procedure specified by the morphosyntactic input. The reader of these lines has probably become unaware of the table he or she is sitting at, if any, unless it is too low, causing an experience of discomfort that may override the reading. In one word, speaking is an alternative cultural medium that can override the natural medium and be utilized to control enactive experiencing instead of letting the physical world "decide" on what should be lived by humans. By "polluting the atmosphere" with acoustic waves, the emitter will make all receiving souls present focus on experiencing the forefronted acoustic interference at the expense of the backgrounded unperturbed medium, including the table.

Example 1 In a famous comic strip for children, *Rupert*, the layout falls into two parts: above are four pictures representing an adventure, like the teddy bear catching a kite, accompanied by a small caption in the simple present tense, *he catches the kite*; and below is an autonomous paragraph narrating the story. The caption is not in the continuous present *he is catching the kite*, because it does not refer to or describe the picture standing for actual experience; it causes the young reader to experience the catching of the kite by interfering with the text as a parallel and alternative experience to interfering with the natural environment symbolized by the picture. The caption is the verbal alternative in the presence of the picture, and the narrative below is the verbal alternative in its absence, as the layout compels the reader to lose sight of the figurative drawings. The child is thus gradually taught to disconnect the verbal process from the nonverbal one and learn how to enact worlds through alternating media. Other strategies will teach the distinction between fictional worlds and downright lie.

Example 2 *Black humor is the dark side of the farce*. In this statement, several polysemic units (*Black*), references (*Dark Side: Star Wars*, *Pink Floyd*), substitutions (*force/farce*) cause the interpreter to enact a semantic chain reaction, a kind of mental fireworks (one *cracks* a joke) in which diverging trains of thought are explored simultaneously, echoing and feeding one another in proliferating structural loops, implying irreconcilable evaluations of the speaker's attitude ranging from genius to delirium through playful but serious subversion, until all hell breaks loose. If there were no sensorimotor "fuse" to interrupt the runaway plottings, any good joke might trigger sensorimotor crises comparable to epilepsy. At some point, the victim bursts out laughing: an exhausting, air- and energy-consuming spastic periodic emotion generating tremendous tactile and auditory signals that will momentarily suspend all interpretation, all conversation, all verbal thinking—just try to think anything while you laugh, just try *even to remember to try*. Laughing is a periodic spasm overriding runaway sensorimotor experiences including the plotting of a joke and physical tickling. Alternative vocal spasms, characterized by other frequencies and the absence of periodicity, namely the cries of physical and emotional pleasure or pain, play analogous roles in attempting to master similarly dangerous, exceedingly intense nonverbal events. The very existence of a sensorimotor antidote for runaway verbal plotting underlines the coexistence of rivaling media: a laugh, a body-made respiratory and vocal spasm, will disconnect the verbal medium by mere physical exhaustion and bring about a gradual

restoration of the natural medium before social verbal interplay can be restored.

10.3.2 Defining the Import of the Language Medium

The verbal event affects the producer (intimate languaging) and/or the "consumer(s)" (private and public languagings). In the latter case, the producer may be causing the "consumer" to enact an intended experience following the relevant ritual lexical and morphosyntactic coupling. And as the sentence makes the consumer become aware of a world that has not (or could not) be captured through experience, this world is intrinsically an original one and need not be structurally different from the experience that might be enacted through direct interference with the natural, non-verbal medium. If one centers languaging on communication, one need not hypothesize that linguistic meaning is any different from natural meaning in direct experience.

Very different is the case of intimate languaging. If I see and hear my dog barking, I will recognize both the being and the action and have no difficulty in identifying the class—this is a dog, my dog, whose name I know, and what he is doing is to bark, a loud cry they do when happy, angry, afraid, and so on. All this can be executed without any linguistic help, and I do not need to name the dog and the barking to identify them as such. However, something in this experience remains unsatisfactory—the barking annoys or frightens me, I do not detect its origin, I cannot anticipate its end, and I do not see what I have to do, whether I should start investigating into its cause—an intruder in the garden?—or incite the dog to put an end to it. In a nutshell, *enaction is stalled and action is paralyzed, and comes to a standstill*: I have failed to enact the natural medium. In reaction and with the positive purpose of restarting the experiencing dynamics, I launch the alternative enactive process, languaging: *Why on earth is the idiot barking?* In so doing, I actively take hold of the enactive problem of the cause (*why*) and the displeasure (*on earth*) involving the identities that I have recognized and enacted in this perspective (*the idiot*: because he is annoying me) that which have to be mentioned not because wording them is required for identification, but because they are involved in the enactive stalling and may be verbally enacted in the perspective of the problem they have raised: a structural loop. This might seem to relate with the pulsional basis of phonation (Fónagy 1983) and the Thwarting Theory of the origin of language (Salzen 2006), except that this model is not exclusively centered on the producer and gives the interpreter's reception a major role (Coursil 2000). In talking,

I make myself focus on and become fully aware of what really matters, the anxiety, excluding all other pending matters, and I start acting upon myself by verbal means in order to proceed toward the determination of the relevant decision to be made and course of action to be undertaken; if consciousness is devoted to action in the world rather than computation (Bruner 1990; Shanon 1993), languaging is used to launch or relaunch action in the face of an enactive stalling or obstacle in natural experience, which includes the social encounter with other selves. Private and public languaging will trigger or concert actions including other selves, while intimate languaging will install personal experience. Action is here taken in the sense of general sensorimotor living and cognitive experience; it includes physical doings as well as intellectual learning, sensations, emotions, judgments, and the like.

10.3.3 How Languaging Catalyzes Action

A word like *dog* is a reiterated vocal action over the manifold experiences of the animal (both direct or indirect) in an individual's life and in social context from the infancy into adulthood. Using the word an additional time will serve as a behavioral trigger connecting the immediate experience with the long-running encyclopedic knowledge about it, acquired through all previous "verbal encounters" and forming a general network of hierarchized features known as notion or prototype (depending on the semantic theory considered: Culioli [1981] 1990); the word dynamically re-presents (re-stages) an historical excerpts of individual experience organized into a category (involving perceptual properties, human actions, cultural values, emotions, etc.) that can be mapped against the immediate situation causing a problem: I have seen dogs bark in a whole range of situations, so maybe the solution for this one is to be found among those. To speak is to command vocally the connection between the immediate and recorded experiences by reproducing the controllable sensorimotor experience, the word, used as a token or common denominator (rather than *symbol*): the lexicon has to be reconstructed from a phenomenological point of view, and is connotational rather than denotational (Maturana 1978; Kravchenko 2004).

The deliberate experience of voicing and hearing of the word, either intimately or privately, is the key to unzipping the encyclopaedic file required for installing enaction through further verbal and nonverbal action. The sensorimotor action of voicing is used as a corporal levy catalyzing that of recollecting an extensive network that could not activate itself alone: literally, *languaging amplifies intelligence* and enables a shift

from a network of local spatial-temporal correlations to an enormous life-escape trespassing the boundaries of immediate experience. *In so doing, languaging is instrumental in constituting reflexive consciousness*: it consists in a general course of actions enabling to match immediate, *actual experience* against the encyclopedic knowledge of sensations and actions retained from previously recorded *virtual experience*. Language makes it possible to convert what is experienced in the here and now into an occurrence of something more general. It is not simply that "this dog" is compared with "other dogs" but that "my experience of this dog" is set against "my previous experiences of this or other dogs." Thus, merely voicing the token *dog* will mechanically induce the distance-taking effect (*Verfremdungseffekt* in the Brechtian sense) and cause the voicer to put immediate and distant livings in perspective—"ceiving" discrete instances of one's living self both as a background (virtual experience) and salience (actual experience); "*déjà vu*" entails "I remember now that I was there."¹⁰ The reflexive self is the one that emerges through languaging out of this contrast between the reflecting and reflected consciousnesses. The continuous entity representing the constant linking of two occurrences of analogous experience is separated from the varying behaviors and events (situations) in which it is perceived to be engaged: comparing the stretching cat and the meowing cat will sort out the unvarying parts of the cat and the varying attitudes. As Bickerton (1990) pointed out, no language in the world has simple words mixing entities and attitudes. By definition, languaging, a vocal comparator of experiences, will stage the difference between the continuous and the variable, paving the way for lexical categorization (nouns, verbs) and syntactic distributions (nominals, predicates).

The goal of lexical semantics is to define the nature of what is reminisced by lexical units and how those programs are formed in personal experience. That of morphology is to define the nature of the connective procedures forming the resulting network. That of syntax is to model the ritualized temporal sequence of vocally marked intellectual operations followed in a given language in this assembling process. That of prosody is to couple the mechanistic dimension of the preceding items with the personal choices about any of them at any moment in the general psychological context. This involves the syntactic steering and the emotional coloring of the vocal motoring (whether private or intimate).

10.4 The Lexicon

The word has been defined as a vocal rite commanding an extended and ever-growing network of heterogeneous, hierarchized recorded experiences

in a hybrid architecture that can be mapped against immediate experience to retrieve relevant decisions and plans, and, on occasion, to make them evolve or add new ones. By definition, the word *cat* does not point to the animal that I may or may not be encountering when using the word (except if the prosody clearly directs the syllable to the cat);¹¹ it revives the network of knowledge acquired in my past experience of cats. This is why determiners and deictics are required to establish the link (if any) between this ensemble and ongoing experience (*a cat, the cat, this cat*), and why it is possible to envisage a cat freely with no connection with the situation, as I am doing in these lines. A lexical unit is *concrete* when it tokenizes a recurring set of experiences that can easily be retrieved by perceptual means without using the word, like a *cat*. It is *abstract* if the collection of events it tokenizes is not perceptually federated by a center that could be retrieved directly, like *society*; words may refer to notions that cannot be treated as entities or events, but to sensations, emotions, and abstract categories born out of highly heterogeneous experiences with no material core. This is due to the fact that the network of experiences out of which the notion is born is not only personal, but also collective and constructivist in its genesis: the notions of *cat* and *probity* do not stem from *my* experience, but from that of all the persons who used the word publicly enough to have a say in the kind of situation recorded in the word's usage.¹² As a result, this collection is atemporal, in that the personal history of the personal encounters with the thing is blurred, to be replaced by a hierarchy of relevant features organized in terms of frequency, reciprocal links and contextual necessary conditions to be activated (as in Rastier's interpretive semantics; see Rastier 2009).¹³ This collection may turn out to be inconsistent, or even contradictory, paving the way for polysemy,¹⁴ but the very principle of the word is that the vocal token is used as an operator of reminiscence to federate disconnected sets of experiences, originating the nominalistic illusion that any word does revolve around a firm conceptual core.

Metaphors, for example, are usually associative and based on analogy (*a bluebottle*), but they may also appear as downright misnomers (French *bouse* "cow dung," for a car) by which the interpreter is forced to redefine his favorite possession in the terms imposed by the speaker (which can only be done in unequivocal situations requiring a demonstrative or a possessive rather than an indefinite article: **une bouse > cette bouse, ta bouse*). Such distortions, if reiterated, are recorded by all speakers in their individual experience, so that the deviant uses of the word are inscribed in the notional network along with the "undistorted" features: transgression is included in the standard. For this reason, it is impossible to draw

a general semantic theory on the basis of a simple subject-world relation: what is at stake is a *world-based subject-subject relation* and, ever since the very "beginning," the lexicon has been determined by the conditions and intentions motivating its use—so that the individual-centered theories of conceptual metaphor (CMT) and blending (BT) (Grady, Oakley, and Coulson 1999) in which individual representational operations overrule social intercourse are highly suspicious.

The way in which the various languagings of the fluctuating lexical units in a given language can be taken as an empirical basis for raising the conventional enactive questioning: how does one describe the sensorimotor experience of the motoring, that is, the acoustic disturbance as it is enacted by the coupling of the vocal output with the tactile and auditory input? And by correlating this description with the interpretive effect experienced by the competent and intuitive linguist in the process of languaging the object of his own study, what can one infer about the architecture of the network¹⁵ of recorded experience that is controlled by the vocal key? Does one really isolate cultural prototypes, or is the format of the "prototyping" universal? Typological diversity strongly suggests that this is not the case and the enormous body of data made available by typological description makes it possible to undertake a systematic modeling of the accurate profile of the structural loops characterizing what nonvocal experience is tokenized by what vocal experience, and how the profile of the latter and the architecture of the former foster each other in the long-term process of personal experience and language formation (rather than acquisition).

The analysis of the properties of the lexical word, both formal and semantic, play a major role in orienting the question about its origin, or, more accurately, the word-forming process in human experience. Current models state that language was originally imitative (Merlin Donald [2001]: human actions mimicking the properties of the object), deictic (Corballis [2003]: silent gestures pointing to objects or animals), emotional (expressions of frustration, effort, pleasure, pain, disgust, fear, anger, etc.). All we can do is guess; there is no theoretical reason why one should decide that one approach is more relevant than any other in the first place. Any repeatable vocal action that happens to correlate with mimicking, pointing, or reacting is a potential candidate for tokenizing: if a caveman walks on a snake, gets bitten and yells, the event may be remembered in any relevant form ranging from the imitation of the yell to that of the hissing animal slithering away. In *Tlön, Uqbar, Orbis Tertius*, Borges [1940] 1956 fancied a language coined by idealist philosophers in which no stable word referred

to objects like the moon—any relevant description of sensation and emotion had to be coined in the spur of the moment, making the word an unpredictable, fluctuating *haiku*. Second, there is no reason either why there should have been an equation between words and things in the first place: if the yell tokenizes the whole event, it is neither a noun nor a verb, but a sentence, that is, a single syllable connected with a whole event involving both actors and action. Only later does the yell get specialized in referring to one of the event's components: the biting or the shouting, the man or the snake; in Wolof, *niao* is not the meowing of the cat, but the animal itself, tokenized by the imitation of one of its most striking audible attitudes: a *symbol* in the etymological sense. The syntax of some languages like Inuktitut and Mayan is suggestive that the sentence is not actually composed by assembling free lexical units, but, on the contrary, that specialized lexical cores tend to emerge out of complex clusters decomposing entire scenes or events into fragmentary subparts. In much the same way, bilabial syllables uttered by infants are associated with breastfeeding, the milk, the mother, and only later does *mum* become associated with one feature of the network, the mother herself, while other vocal candidates are proposed by the surrounding adults for the other elements; a vocal token may get increasingly attached to one of the features of the global scene that originally motivated it.

10.5 Syntax

Linguistic typology identifies word orders in which the prototypical sentence is asserted in a given language: the verb is initial in Irish, medial in English, and final in Basque. Most languages have a prototypical pattern ordering lexical units (excluding pronouns): SVO, SOV, VSO, and so on. This raises the question of the relevance and universality of the categories used for the description, and it falls short of elucidating how syntactic categories marked by morphemes and/or location are experienced by the language user, formed in the experience of languaging, and what role they play in assembling meaning.

Syntax is an algorithm collectively and culturally formed through personal experience in the course of individual experience and collective history of language forming, learning, using, and teaching. This algorithm is an ordered sequence of actions: grammatical relations are to be reconstructed in real time in the process of interpreting sentences (Valin 1981). According to a variety of traditions, the *agenda* (literally, "the actions to execute") of the prototypical English sentence schedules a nominal section

known as the subject or NP and a verbal section known as the predicate or VP.

Together, they constitute the *orthosyntactic agenda*, the core of the syntactic program, the general network of the languaging unit. This pattern is doubly versatile and can vary in two ways, globally and locally. Globally, it may be complemented by optional expansions before, between, and after the two steps of the orthosyntactic agenda; the *parasynthetic addenda*, that is, adverbial phrases. And locally, each step of the agenda is to be executed by a versatile local network whose expansion ranges from the lexical singleton to the entire clause: the subject is realized either by a pronoun, or an NP, or a nonfinite clause, or a finite clause introduced by a complementizer. In the same way, the predicate is either saturated by an intransitive verb, or expanded by a transitive verb, a whole range of complements, and so on. The parasynthetic addenda are ruled by the same global versatility at the local scale: an adverbial clause is saturated by an adverb or expanded by a prepositional phrase, a nonfinite or a finite clause introduced by a conjunction. If expanded and complex enough, the execution of the subject in the orthosyntactic agenda may incorporate an internal parasynthetic addendum in the form of a relative clause or be suspended by an insertion; the same incorporative or suspensive diversions may be improvised within a parasynthetic addendum (e.g., a relative clause within a finite temporal adverbial clause). The interpreter of the utterance is trained in the practice of tracing the syntactic route in real time just as the speaker is trained to profile it, so that dialog can effectively alternate individual roles in semantic orienting, with metasyntactic prosody anticipating directional decisions for both speech and interpretation (for an enactive analysis of prosody, see Auchlin et al. 2004).

Example In Basque (*Euskara*), an ergative and agglutinative non-Indo-European language spoken in the Basque country (the western tip of the Spanish and French Pyrenees), the prototypical sentence begins with nominal and adverbial arguments and ends with the verb. For each argument, the *lexical phrase* falls into two sections, a notional phrase receiving all notional elements (adverb, noun, adjective, demonstrative) suffixed by a functional phrase receiving all grammatical specifications (case markers, determiners, number); the functional phrase is attached to the last lexical element present (the noun, the adjective or the demonstrative): [((oso)) *andre* (*polit*)]-a "[((very)) woman (*pretty*)]-the."

Adverbials are obtained by suffixing notional phrases with one or several spatial or relational case markers forming "simple" cases (the inessive, the allative, the ablative, the genitives, the instrumental, and so on: *etxera* "to

the house") and complex cases (the prolative, the comitative, the destinative, and so on): *etxerakoan* "in the of to house" > "while going home." Nominal arguments of the verb receive a simple case marked by no or one suffix: the absolutive (A, -Ø), the ergative (E, -k) and the dative (D, -i). It can be demonstrated that this specific morphological trio is in fact a grammatical distributor of gestaltian positions coupling notional phrases with functional roles, namely a *base* (the obligatory and unmarked absolutive), a *dynamic peak of instability* (the optional ergative, marked like a kind of genitive in relation to the absolutive base) and a *trough of stabilization* (the dative, similarly optional and marked). This results in four argumental configurations (A, EA, DA and EDA) in which the unmarked absolutive is enacted as a site, a patient, a cause or a trajector depending on whether it is isolated or accompanied by the dominating ergative, the dominated dative or both: *emaztea*^A "the wife" (isolated base); *senarrak*^E *emaztea*^A "the husband, the wife" (agent, patient); *emaztea*^A, *senarrari*^D "the wife, to the husband" (cause, experiencer); *senarrak*^E *lorea*^A *emazteari*^D "the husband, the flower, to the wife" (source, trajector, target). In this system, valency converges on the uninflected absolutive NP, with the optional and inflected ergative and dative NPs connected with the former like genitival or adjectival adnominals: (E-k) > A (Ø) < (D-i), that is, [(peak) > base < (trough)] or [(+) > 0 < (-)].

As regards the VP, the final finite form incorporates a series of bound pronouns retrieving the previously coupled notional/functional NPs, thus echoing by multiple agreement the nominal, absolutive-centered argument structure. This agglutinative conjugation either revolves around the lexical verb stem (*daramakiot* "it-bring-to him-I" > "I bring it to him") or accretes into an autonomous final auxiliary (devoid of any lexical root), while the verb stem receives aspectual suffixes: *eramango diot* "bring-genitive it-to him-I" > "I'll bring it to him." In this system, (1) verbal valence, incorporated in the form of the bound pronouns by multiple agreement, tends to adjust to the initial nominal configuration (even if some minor verbal categories do tend to export their own requirements), and (2) placing the same notional argument in two distinct functional roles is prohibited as in **senarrak*^E *senarra*^A the husband, the husband (agent, patient), so that reflexive double agreement is strictly banned, albeit morphologically predictable and interpretable (Bottineau and Roulland 2007). The complete sentence comes out as follows:

<i>Senarrak</i>	<i>emazteari</i>	<i>lorea</i>	<i>emango</i>	<i>dio</i>
Husband-the-E	wife-the-D	flower-the	give-of	it-her-(he)

"The husband will give the flower to his wife"

The interpretation of the system is the following (Bottineau 2005a, 2005b): sentencing in Basque consists in producing separately an analysis and a synthesis of the “scene”: a *dramatis personae* or *cast* (with the husband starring the source, the flower the trajector, the wife the goal) in which notional actors (the husband, the flower, the wife) are analytically coupled with functional roles (the peak, the base, the trough). This results in the formation of one to three arguments: three, in this example. In the verb phrase, the plot consists in recombining them (the agglutinative multiple argument *dio*) in the context of the time, mode and dialogical specifications (marked by other affixes: *zion* in the past, *lioke* in the conditional) around the verb stem (if it is free, *daramakiot*) or in an emancipated “auxiliary” (if the verb stem receives aspectual parameters, *emango*).

This seemingly mechanistic algorithm is to be relativized by the orthosyntactic, parasyntactic, and metasyntactic couplings; not all arguments need be explicitly stated, some may be fore-fronted or back-grounded for specific motives, the verb may “disagree” with the nominal argument structure (passive intransitivization, antipassives, allocutive inclusion), all of which generates the same amount of local and global versatility as in English. This piece of algorithm operates both locally, in each part of the clause, and globally, over the whole sentencing process. This could be analyzed in terms of Robert’s fractal grammar, possibly an emerging formalization of structural loops.

Each language is characterized by a dynamic versatile orienting algorithm of the kind exemplified earlier; those syntactic models illustrate the way in which languaging, according to Maturana 1978, is the reciprocal coordination of actions between orienters and orientees, with both functions constructively carried out by both participants.

10.6 Forming Lexical and Grammatical Semantics

10.6.1 Forming Notions and Relations: An Authentic Example

The following excerpt was actually uttered by a mother playing with her two-year-old daughter at a French skiing resort:

Regarde! C’est de la neige. Regarde! C’est blanc, c’est froid, ça colle, on peut en ramasser et en faire une boule, et la jeter sur papa, tiens, regarde, poum! [Look! This is snow. Look! It’s white, it’s cold, it sticks, you can pick some and make a ball, and throw it at Daddy, look, splash!] And they all burst out laughing.

A traditional vision of the meaning of *est* in “c’est blanc” and so on is that the verb *refers* to a continuous state, a transitional or permanent

property, or attaches a unit to an ensemble. But what is overlooked is the general architecture of this piece of collective experience in which the playing and the languaging are intermingled. Before this passage is given, the child is not supposed to have even begun to detect the snow, let alone identify it as such, let alone remember its “noun.” In the aftermath of the utterances, the child has been given the opportunity, all in one, to experience the snow in all its sensorimotor dimensions (the color, the tactile effects, and what one can actually do with it on account of those properties), and each of these elements has been matched with a vocal sensorimotor event in the form of the very words *snow*, *white*, *cold*, *sticks*, *pick*, *throw*: the child has been presented (in Brentano [1874] 1944’s sense) with the cultural properties of the snow such as the word, a reproducible external memory, records them.¹⁶ The simultaneous creation of the semantic and lexical fields are part and parcel of the same bundle of experiences in which no theoretical reason should justify that one draw a clear-cut distinction between the verbal and the nonverbal. Semantic features and phonological ones are both sensorimotor predictions formed in the course of long-running experience and training; the articulatory and auditory features should be placed among the semantic features of the notion rather than separated because they are of the same nature—a position that is reminiscent of *phonology as a human behavior* (Diver 1979).

This is also true of the abstract verb *be*, whose connective value is experienced whenever the child is confronted with this kind of situation, and of the demonstrative *this*, which is associated with a deictic windowing of the child’s attention oriented by the mother’s bodily movements. The child needs no previous knowledge that *ce* is deictic, nor does he even have any inborn cognitive “deictic functional slot” to be filled (*parametered*) by a language-specific marker.

In the present day, nobody could decisively demonstrate that postulating an inborn universal grammar is a downright error, but it certainly is an extremely costly, implausible, and unnecessary hypothesis in an historical scientific context in which the exploration of nonvocal and vocal couplings (even before birth) is only beginning, paving the way for modeling how cognition is embodied in experience.

10.6.2 The Network-Forming Value of Grammatical Operators: Some Examples

In Basque, each notional phrase like *etxe* “house” is suffixed by a functional phrase like *-a* “a/the.” The orthosyntactic pattern, at every level, consists

of identifying one or several notions and then incorporating them in the forming network by means of a relation marked by the one or several postpositions involved in the process. *-a* validates the correlation between the notional network tokenized by the vocal experience of the word and nonvocal experience of the situational reality; *etxe* applies to both the memory of experiences past and to the newly added present experience (including, precisely, memory as a present experience). This is not symbolical, it does not mean that there is an iconic correspondence linking the "represented" abstract house with an identifiable one, but that the same network of heterogeneous features is co-activated by the vocal experience of the voiced word and the nonvocal experience of "reality"; *-a* enacts a convergence, the present merger of experiences retrieved from different moments in personal history. When *-a* is absent, immediate experience is excluded from the converging process, which only mobilizes recorded history.

In Guillaumean terms, this is called *représentation* and *actualisation*, a binary model with a strong enactive bias. The core value of Basque *-a* is the convergence of experiences acquired in the past and in the immediate present concerning the set of features tokenized by the immediately preceding notional unit (or set): (1) *etxe* restores the network (by intelligence amplification), (2) *-a* restores the converging process. This is how the intimate, private or public orientee is made to "network" meaning and obtain a semantic event—something experienced as real happens to be a house, or, more precisely, in the opposite order; the previously experienced set of events vocally and conventionally tokenized as *etxe* happens to zero in on some fragment of immediate experience, *-a*. *-a* may seem to point to an external reality (referential semantics), but what with the definition of experience, medium, and enaction, it should be clear that what *-a* stages cannot be the binary pointing of an object by a subject, but a merging, unifying process.

As regards the role of *-a* in the orthosyntactic algorithm, analysis/synthesis in Basque, the "determiner" provides the answer for the problem raised by the notion: "*etxe*, house. (Yes, I know what this is, so what, here and now? Yes, house here and now, *-a*.) Some authors anchor syntactic patterns in narrative procedures and semantic ordering in cognitive constructional processes carried out by the speaker (Guillaume 1964; Adamczewski and Delmas 1982; Cotte et al. 1993; Rousseau 2005); others investigate the hypothesis that they may fossilize or sediment ritual dialogic patterns involving alternating turns of speech at different phrasing scales in the sentencing process (Givon 1994; Douay 2000; Bottineau and

Roulland 2007). Although Langacker's (1987) cognitive semantics is not primarily devoted to morphemes as the markers of orienting procedures, there do emerge occasional analyses which unmistakably involve this kind of approach.¹⁷

In the case of Basque, the question/answer ritual—which normally spans two sentences spoken successively by two interlocutors, not one—seems to have been embedded in at least three increasingly concentrated levels in the orthosyntactic algorithm:

1. At the global level of the clause, in the form of the analysis/synthesis pattern discriminating the argumental cast and the verbal plot: *x* is *E*, *y* is *A*, *z* is *D*, > (implicit intermediary question: so that, knowing each actor's respective roles, how do I plot their encounter in my experience?) > answer for the purported question: verb + all specifications.
2. At the local level of each argument, the notional phrase, by summoning an amplified network of recorded experiences forming a notional network, raises the question of how the latter relates with other networks (summoned by other notional phrases) and immediate experience (both vocal and nonvocal), which is the same thing (as notional networks are summoned by words). The answer to this question is provided by the functional phrases (postpositions) comprising case markers (connections linking phrases with other phrases: the ergative with the absolutive, the inessive and instrumental with the verb, etc.), and determiners and quantifiers (connections linking recorded networks with the environment of immediate experience at large).
3. In-between, on each side of the clause (NPs and VP, in the analytical and synthetic phases), the lexical cast (nouns, verb) is interconnected by the grammatical network (postpositions, bound pronouns and temporal, modal, and allocutive markers).

All this is far too complex for illustration, but the principle remains that at every stage, the binary algorithms are federated by an intermediary implicit question raised for the orientee by the first segment and answered by the second, which amounts to saying that the algorithm, as a general rule and whatever its position, composition, and range, sediments and embeds the dialogic experience of asserting, questioning, and answering, making the ritual of each languaging act the prediction of the experienced narrative structure of the social intercourse. In this respect, Basque appears to be narratively iconic in that the [statement-(question)-answer] procedure experienced in the alternating [speaker1-(speaker2)-speaker1] dialogic sequence is recorded in the grammatical pairs staging the alternations

[orienter-(orientee)-orienter] in the same order at the global, intermediary, and local levels (clause, phases, and phrases).

In contrast, the French and Breton languages forefront the determiner *la maison*, *an ti*: in both cases, the answer concerning how the notion is to be interconnected with the general assembly is provided even before this question is actually raised by the lexical specification of the notion. Basque anticipates the need for an answer; Breton and French anticipate the question itself in the act of predetermining the answer, which is an example of how a structural loop originates predictive decisions in the phrase and sentence planning of interpretation. This holds true for prepositions, auxiliaries, and conjunctions at the scale NPs, VPs, and clauses, all of which are forefronted in Breton and French, back-fronted in Basque.

The orthosyntactic algorithm sediments the dialogic experience (incorporates in it the languaging experience at all its levels by the structural looping)—what O. Fischer calls “diagrammatic iconicity (Fischer and Nänny 2001).”¹⁸ But the algorithms are language-specific; one of the factors of variation is the range of the prediction (the answer or the question preceding it), and the dialogic pattern may not be universally definable in the terms of questions and answers (a track unexplored as of now). The network-forming value of a grammatical morpheme cannot be understood outside the frame of the orthosyntactic algorithm defined as a languaging rite sedimenting the dialogic experience.

10.6.3 Submorphology

The operational value of a lexical or grammatical marker is not supposed to be found in the phonological features making up the phonemes involved in the syllable(s): *dog* tokens a set of experiences, but there is no obvious analogical connection between the non-vocal ones (the barking, the drooling, the fretting, and so on) and the vocal one (the wording). In binary terms, the arbitrary word is not phosymbolic in that the sensorimotor experience of the linguistic unit is not akin to that of the nonvocally experienced “thing”; in unitary terms, the vocal features forming the linguistic experience incorporated in the general network are not akin to the other features originated by direct interaction with the phenomenon. All that federates the network is repeated co-occurrence.

In the same way, a grammatical morpheme is not supposed to be *suggestive* of the operation to be carried out, *reminiscent* only, just as there is no resemblance between Proust’s tea and madeleines and the *Piazza San Marco*. And yet! Onomatopoeia is based on an imitation of nonvocal experience by vocal action, in some cases implying multimodal synesthesia. In

many languages,¹⁹ the lexicon is at least partially underlain by internal, submorphological consistency (phonesthemes, ideophones, *racines*, *matri-ces*, *étymons*) that may at its own level be either “arbitrary” (fully contingent, like *sp* for centrifugation in English) or, in some cases, possibly motivated (like *st* for immobilization in English), at least in a very remote past, and psychologically relevant (Bergen 2004). Lexical classifiers in Bantu languages are similarly organized in a consistent way (Reid, Otheguy, and Stern 2002).

The semantic effects of this property of the lexicon have long been recognized (Wallis [1653] 1969). A reasonable hypothesis in this domain is that for the efficiency of the languaging performance, it is profitable to elaborate a limited number of submorphemes that are used as experiential classifiers of lexical networks, and even more so if one happens to come up with a combination of phonemes whose sensorimotor experience coincides with that of the nonvocal features of the assembly: suggestion is more powerful than reminiscence, and cultural convergence around this fact will make imitations easier to convene upon than random couplings (which neutralizes the traditional opposition between natural motivation and social convention). So if imitation is by no means necessary in diachronic lexical emergence, it is a useful catalyzer for the individual as well as the community; this fits well with the data, as although human language is widely documented to be unmotivated, there tends to emerge locally remarkably organized subsystems—*islands of order* in oceans of chaos—which are strongly suggestive of an underlying ordering principle.

The same holds true for grammatical morphology. The word *be* is by no means vocally suggestive of its combinatory role in the orthosyntactic train. But in English, there exists a finite paradigm of morphemes including *the*, *this*, *that*, *there*, *then*, *thus*, *though*, all of which begin with *th-*, and all of which signal the retrieval of something immediately available from memory, whose category is differentiated by the rest of the operator (*there*: a locus, *then*: a moment, *the* + N: a notion, *this* and *that*: situated nonvocal experiences that may be specified by a notion, etc.). The submorpheme *th-* manifestly alternates with *wh-* signaling the unavailability of any relevant preestablished knowledge in working memory and in the field specified (*who*, *which*, *what*, *where*, *when*, *why*). On top of the submorphological pair, some operators do alternate as wholes (*where/there*, *when/then*, *which/this*, *what/that*), revealing other semantically relevant phonological minimal pairs (*i/a*, *s/t*, *r/n*) that happen to have long and massively been evidenced in similar functions in altogether unrelated linguistic types

(Atlantic, Bantu, Altaic, Semitic, Caucasian languages) and across diverse grammatical categories (Robert 2003).

While some pairs appear to be purely un- or demotivated (how should *th-* be suggestive of memory?) in the current synchrony at least, others are easy to connect with their operational value: *i*, a sensorimotor experience of conjunction and contact, is frequently used as an operator for the same relation between semantic phases in the algorithm (examples are by the thousands in natural languages). *a*, enacting vocal disjunction and distance in aperture, is apt to operate the same kind of semantic connection. *s* for (sibilant fricative) continuity, *t* for late (plosive) interruption (as in Tossaint's 1983 model), *k* for precocious interruption, *r* (apical or velar) for forceful launching (*passage en force*), *n* for (nasal) by-passing (of the oral tract) as in negation in Indo-European languages (but also Japanese). In a theoretical frame known as cognemetics (Bottineau 2003, 2008), I have proposed a model according to which, within the formal morphophonological constraints of voicing (syllable structure, phonemic compatibilities) such as they have evolved to the present day (cf. the great vowel shift in English), there does exist, in grammatical morphology, a strong subterranean trend or force dynamics presiding over a quest for relevance in phonological selection and placing in the formation of morphemes; cognemetics appears to be at least highly compatible with motor theories of the origins and current workings of language (Studdert-Kennedy and Goldstein 2003; Allott 1995; Lieberman 1991).

Morphosemantic connections are not necessary: (1) not all morphemes need be componential, and (2) not all phonemes need have been selected for such a relevance. A strongly catalyzing factor is when a limited number of markers form a finite paradigm in which they can echo one another both vocally and semantically, forming a kind of virtual, discontinuous poem in which family resemblances are easy to detect, contrary to what happens in the larger lexicon—except in local semantic-lexical fields, precisely because they are federated by the same principle. In short, the very idea that in local subsystems the sensorimotor patterns of voicing may be directly utilized to ensure semantic connections is not to be censored for the sake of safeguarding the orthodoxy of randomness. This appears as an extremely deep case of embodied cognition in which the sensorimotor action controlling the rhetorically distinguished voicing and meaning actually fully coincide in one single process, entirely revoking binary symbolism. The process is not a ruling principle for all lexical and grammatical units, far from it; it is only a potential principle that will emerge only when viable, that is, productive and efficient, or at least relatively more so in relation to existing strategies.

10.7 Perspectives

This chapter has not discussed some crucial topics. The enactive bias to be observed in ancient and modern linguistic thought has been merely alluded to, but not explored. One should itemize the various features making up the enactive paradigm (autopoiesis, allopoiesis, action, structural loops, sensorimotor experience, the medium, embodiment, life, cognition, anticipation and decision, individuation processes, membrane, self, boundary, and the revocation of binary thinking and wording: innate/acquired, interiority/exteriority, subject/object, stable/dynamic, permanent/transitional, symbolism, etc.), assess whether they do form a self-sustaining paradigm in which all components are organically attached and mutually necessary, or whether they form a consistent but relatively loose collection of concepts with no real gravitational centers as some critics suggest. This is essential for the linguist (and any discipline indeed) as the qualitative variation of the "enactive bias" in models can be explored in terms of which component of the enactive model is retained, in isolation or in collaboration with others, and incorporated in a preexisting frame that may be compatible or not, depending on whether one operates with a loose or strict construction of the enactive paradigm. The perfect example is embodiment and force dynamics in symbolic models, probably the most obvious topic regarding the linguistic/enaction connection.

Other related questions are the acceptability of some of these concepts in a field studying a living behavior rather than the being itself. Indeed, the living is the action, and the loop is looped: the enactive approach will inevitably reform the "object" and the "subject" of scientific enquiry into a medium including the discourse. A reader relying on symbolical orthodoxy may have judged some terms metaphorical, which amounts to rejecting the structural loop mutually affecting the language and the languaging, the reduction of binarism. If the notion of individuating autopoietic viable development is accepted for social bodies, it is applicable to the emergence of dialects, literary genres, conversational rites, stylistic conventions, grammatographies, linguistic theories, and formalisms.

As regards the question of the origin of language, everybody feels concerned with the topic and has something to say on the matter—the social link par excellence, the "sport" turning the growing individual into a full-fledged cultural person. The question, of course, is ill-worded, and some misunderstandings are to be removed. Language is not an object; it cannot be acquired. The forming of personal languaging is part and parcel of the forming of the person, just as walking, jumping, or flying. In the same

way, but at a different scale, the forming of a community's social language is part and parcel of the cultural forming of the tribe. Through continuous structural looping, a whole range of heterogeneous actions with diverging functions gradually accrete, in a self-individuating autopoietic emergence, into what we presently enact as a unified function. Why should language have *one* origin, or be based on a single primitive perceptual or agentive channel, such as vision in Givon 1994, or connections between biological and cognitive evolution (Leroi-Gourhan 1964), or tool-manipulating and language-sequencing (Greenfield 1991)? Why should not the *bow-wow* theory and all the others coexist separately in their own experiential domains, or be linked in a looser connection that falls short of deserving the name *language* such as we enact it today? Why should not iconographic engravings have emerged from prehistoric painting separately from the articulated shouts, to merge later into what one now calls writing? Why should not the various social functions (expressions of thwarting, threat, seduction, order, emotions, religious spelling, etc.) have emerged through unrelated channels, vocal or not, to precipitate later into a unified system? The prevailing controversy opposing mono- and polygenetic views might well turn out to be irrelevant: for it to be valid, language has to be unified in the first place, no matter whether it "appears" in one or several places, simultaneously or in succession.

And is language a unified phenomenon by any other means than the very word that crystallizes the cluster of individual, social, instantaneous, eternal, human experiences? Is it a property of the being—a competence emanating from biological evolution—or of the species organizing itself into co-allied assemblies (Dessalles 2000)? Does the question not arise from the very word? How many languages do not have any such word, and is the question universally relevant in human thinking? Language looks like the Amazon rather than the Nile—to find the origin, seek the sources.

Language is clearly a question of vocal action tokenizing experiential events and actions in real life. When the ape comes across the leopard, it does not respond by a specific, recording- and reminiscence-triggering yell that may be later reproduced in the absence of the predator but in the presence of the conspecifics; the human being does. This amounts to a 0/1 opposition in which an insignificant change in behavior, yelling or not yelling at important times, will change the species' destiny forever: either the animal does not create vocal tokens for the diversity of experience, or it does, and this slight change may either result from evolutionary continuity, or stem from a purely accidental change in habits that turned out to provide good results.

It is very likely that language has never been simple; it may have suddenly emerged in a kind of critical change when some individuals embarked upon a new behavioral pattern, the vocal accompaniment of experience, with some individuals performing it efficiently (in terms of selection for mating for example, and then transmitting the habit to the offspring), others not. It may also have gradually shifted from the voicings of one's emotions and feelings (pain, fear, desire, anger) to that of the external realities causing them (predators, food, mates) in a kind of hypallactic shift. One cannot know how long it took to eliminate the individuals who did not participate in the change, but one may hypothesize that for one given individual, the change was not gradual: either one did not vocalize, and used zero words, or one did, and tokenized as many experiences as appeared relevant in real life.

If it was so, the sheer number and diversity of voicings raised the question of syntax right at the start: how did two tokens relate with one another, how did the receivers tackle joint occurrences? As for the cognitive capacities required for the efficiency of vocal tokenizing, especially in terms of memory and centralizing the diversity of heterogeneous experiences under one label, either the brain was somehow ready for the change, or it got trained to become competent in this task, stimulated by the recurring occurrences of the task, just as one is not born with big muscles but acquires them through training; if all the offspring undergoes the same process, it becomes "natural," that is, inevitably attached to the developing process of each individual, no matter whether this translates into some form of biochemical genetic encoding. Parrots can imitate the human voice with a vocal organ that is entirely different, but do not speak. Some primates do have the required auditory capacities to develop phonological aptitudes, but do not do so even if immersed in a human community for a long time (Karmiloff and Karmiloff-Smith 2001). Language is not an individual behavior transcending the species; the form of language we know characterizes our species and is *imposed by social life* in the course of individual development at a time when the individual is not yet a full-fledged person: growth breathes and feeds on a chatty world. Chomsky's 1965 hypothesis of a universal grammar is as unnecessary as it is arbitrary, and Pinker's 1994 individual-centered suggestion of a language instinct, whether or not formalized by a UG and materialized by an "organ," requires that the person be fully autonomous and undergoing a social experience-tight development in the first place to acquire the knowhow as an external activity, opting to do so rather than stay on the sidelines.

Language is, all in one, an enactive all-selves-encompassing way of organizing knowledge and controlling action through direct intervention in the world.²⁰ Linguaging alters the environment and accretes the selves into a cultural body that self-defines itself as one of the living species—mankind. The symbolical denial of this unity stems from the failure to acknowledge that a signifier is not exclusively physical, just as a concept is not purely mental, and that none of them control the other in a one-way relation. Notions, or social-cultural enactive knowledge, is information gained through reciprocal and collective perception-action in the environment.

Language is the school of human life and individuation; it enables the species to survive socially, thanks to the concerted production of food and artifacts, *out of its ecological niche, or make the entire accessible universe an ecological niche of its own*, spread over the planet and beyond; it is even to be feared, to overwhelm other species within the expanded niche and destabilize the henceforth fragile balance of biodiversity as is currently witnessed, with an increasingly dramatic number of living species *and human languages* collapsing even before we have become aware that they exist. How language is turning into the instrument of universal survival and doom might be worth investigation.

Notes

1. Talmy's project (2000) is to study the linguistic representation of conceptual structures. In theory, sentences prompt listeners to construct mental representations, but communication appears extremely late (vol. 2, chap. 6, 337–369) and in strictly pragmatic terms; despite the symbolic connection between the two levels, the dynamic procedure is not made explicit.

2. For example, the Basque and Breton languages had not developed the technical vocabulary of new technologies until academic authorities decided that their lexicons had to fill the gaps that appeared when the languages were mapped with Spanish, French, and English. Over the past few years were published a manual of teaching methods in mathematics in the Breton language, along with a *Geriadur ar Fizik* (An Noalleg 2006), an impressive collection of scientific and technological terminology, have recently been published. Lexicons are *potentially* universal, but they *actually* cover the areas their community has substantially experienced.

3. For a sample of conversation in Sochiapam Chinantec (Mexico), see <http://www.sil.org/mexico/chinanteca/sochiapam/131-Conversacion-cso.htm>.

4. The pitch variation of African talking drums makes it possible to imitate the languages' tones and convey some relatively complex messages, including the "drum names" of the sender and the addressee.

5. Toyi-toyi, a southern African dance used in demonstrations during the apartheid period, is occasionally called "foot language," but its "semantics" is limited to the expression of the hostility of a social group in the face of the enemy, like the Maori *haka* war dance.

6. Joly in Cotte et al. 1993, 46.

7. What one reads is fully understood if and only if the relevant intonation and rhythm guide the mental voicing. Otherwise, itemizing a list of disconnected words is semantically unproductive.

8. This is not to say that the structure of the word is iconically and phonosymbolically motivated by the sensorimotor properties of the phonemes and/or graphemes: the spelling *wr-* is no less arbitrary than the whole word *write*. Simply put, English has a strong tendency to retain a whole gamut of such arbitrary consonant clusters, nonetheless specifying the semantic field to which the lexical belongs. These fields happen to reflect the way in which the "object" or "action" is experienced through sensorimotor dynamics.

9. The most striking example is, no doubt, Lakoff and Johnson's title *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought* (1999), a book in which Varela, Thompson, and Rosch's *The Embodied Mind* (1993) is not mentioned.

10. Basque has an extraordinary way of displaying this dynamics in morphology. "I was" is said *nintzen*. *Zen* is "was" in the third person singular. *Ni* is the first person pronoun "I," and *-n* is the marker of the past, but also the genitive (possessor, as in *aitaren kapela* "Daddy's hat") and the inessive (location in a frame or background as in *etxean* "at home"). So the whole cluster is to be interpreted as *nin* + *zen* (*t* is epenthetic), that is, *ni-n*, "the past version of me" (taken as a background against which the present instance of the speaking I stands in the here and now) "was" (in the third person). If the subject pronoun is given explicitly, *ni nintzen*, as in French "moi j'étais," "I" is envisaged both as a present reflecting self (*ni*) and as a past reflected instance (*nin*), and in *nintzen*, the past is marked twice: once for the event (*zen*) and once for the occurrence of the self experiencing it (*ni-n*). Verbal paradigms consistently evidence this kind of analysis.

11. Cf. this anecdote: I once saw a two-year-old boy kill a bird accidentally. Trying to bring it round, the boy desperately called it, saying "Oiseau! Oiseau!" using the noun as a proper name.

12. I myself remember reading dozens of occurrences of the word *probité* in Jules Verne's novels when I was a teenager; in my experience of this word in French, this