

LANGUAGE IN MOTION: A FRAMEWORK FOR UNIFYING SPOKEN LANGUAGE, SIGNED LANGUAGE, AND GESTURE

SHERMAN WILCOX
University of New Mexico
Albuquerque, NM, United States
wilcox@unm.edu

ABSTRACT

In this article I describe a framework for unifying spoken language, signed language, and gesture. Called the language as motion framework, it relies on three broad theories: cognitive grammar, dynamic systems, and cognitive neuroscience. The foundational claim of the language in motion framework is that language and gesture are manifestations of a general human expressive ability which is grounded in embodied cognition and the need for mobile creatures to make sense of their environment.

KEY WORDS: signed language, spoken language, gesture, cognitive grammar, dynamic systems, neuroscience.

EL LENGUAJE EN MOVIMIENTO: UN MARCO TEÓRICO PARA UNIFICAR LA LENGUA ORAL, LA LENGUA DE SIGNOS Y EL GESTO

RESUMEN

En este artículo se describe un marco teórico para la unificación de la lengua oral, la lengua de signos y el gesto. Se denomina el lenguaje como movimiento y parte de tres teorías principales: la gramática cognitiva, la teoría de sistemas dinámicos y la neurociencia cognitiva. La afirmación fundamental del marco teórico del lenguaje en movimiento es que el lenguaje y el gesto son manifestaciones de una capacidad humana general para la expresión que está basada en la cognición corpórea y la necesidad de las criaturas móviles de dar sentido a su entorno.

PALABRAS CLAVE: lengua de signos, lengua oral, gesto, gramática cognitiva, sistemas dinámicos, neurociencia.

1. INTRODUCTION

In this article I offer a framework for the unification of spoken language, signed language, and gesture. I start with the observation that in order to communicate, animals must create perceptible signals. For human communication, the predominant way in which signals are produced is by moving parts of our bodies. For speech, the means of production is restricted to the vocal tract. For signed languages and gestural communication, much more of the body is used, including the hands, face, and body postures. As Neisser observed:

To speak is to make finely controlled movements in certain parts of your body, with the result that information about these movements is broadcast to the environment. For this reason the movements of speech are sometimes called articulatory gestures. A person who perceives speech, then, is picking up information about a certain class of real, physical, tangible (as we shall see) events that are occurring in someone's mouth (Neisser 1967: 156).

My claim is that language –all language, including spoken languages such as English, Navajo, or Portuguese and signed languages such as American Sign Language, Catalan Sign Language, or Saudi Sign Language– is the production, perception, and interpretation of biological movement. It is articulatory gesturing.

I call this the language as motion framework. Language as motion captures the fact that the physical foundation of language is movement. It also proposes that the key to unifying spoken and signed language, and gesture as well, is to begin with the real, physical, tangible events that constitute all language and gesture.

The language as motion framework is built on three theoretical pillars:

- a) A theory of language that can encompass spoken and signed languages. The theory should not rely on an abstractionist solution, one which posits an abstract set of logical symbols devoid of material substance. Instead, I insist that the proper approach is an embodied solution which unifies at the level of the physical performance of language, of usage events. The theory I use is cognitive grammar (Langacker 1987, 1991, 2008).
- b) A theory that models physical performance as skilled action, and which can be applied to a view of language as gestural performance and grammar as skill. For this I rely on dynamic systems theory (Thelen and Smith 1994; Spivey 2007).
- c) To account for how language and gesture are implemented in the brain as skilled action, a non-Cartesian, embodied theory of cognitive neuroscience is required. Such a theory is that developed by Gerald Edelman (1987, 1989), the Theory of Neuronal Group Selection or "Neural Darwinism." Compatible theories also include those offered by Berthoz (2000), Llinás (2001), and Damasio (Aziz-Zadeh and Damasio 2008; Damasio 1994, 2010).

2. THE HUMAN EXPRESSIVE ABILITY

I reject the notion that human language arose suddenly, or that it was "effectively instantaneous, in a single individual, who was instantly endowed with intellectual capacities far superior to those of others" (Chomsky 2005: 12). Rather, I claim that human language is grounded in a human expressive ability.

The human expressive ability is not unique to language. It is manifest in all forms of human expression, including dance, music, art, and gesture.

Further, I claim that this expressive ability has its ancestral source in a general comprehension ability, originating from the organism's need to make sense of its environment in order to survive, an ability that arose through natural selection. This comprehension ability to make sense is driven by the fact that we are mobile creatures. As the neuroscientist Rudolfo Llinás (2001: 38) notes, "at the behavioral level any actively moving creature must have predictive abilities in order to interact with the external world in a meaningful way." I also claim that comprehension is selectionist rather than instructionist in its nature. That is, "there is no 'voice in the burning bush' telling the animal what the world description should be" (Edelman 1987: 32).

The emergence of the ability of moving creatures to make sense of the world was a major, perhaps the primary, factor in the development of the human brain. Again, Llinás (2001: 21) sums up this position when he observes that "the capacity to predict the outcome of future events –critical to successful movement– is, most likely, the ultimate and most common of all global brain functions." Because of this, the human conceptual system is deeply embodied in perceptual and motoric interactions with the environment. Embodied cognition is the motive force driving the human expressive ability.

2.1. Embodied Cognition

Conceptualization and the human expressive ability emerged from the need for sentient and mobile creatures, possessed with brains that evolved with deep connections between motion and perception, to create meanings. This is yet another implication of the embodied solution: all meaning is embodied. This embodied view unites a fundamental dichotomy that has been deeply embedded in Western thought for centuries –the mind/body duality. The mind/body duality encompasses a host of other dichotomies– cognition/emotion, knowledge/imagination, thought/feeling, language/gesture, language/sign. The language as motion framework and the embodied solution resolves these dichotomies. As Mark Johnson (2008: 9) notes, an embodied approach suggests that "meaning is shaped by the nature of our bodies, especially our sensorimotor capacities and our ability to experience feelings and emotion." The embodied theory of meaning "sees meaning and all our higher functioning as growing out of and shaped by our abilities to perceive things, manipulate objects, move our bodies in space, and evaluate our situation" (Johnson 2008: 11).

2.2. Cognitive Grammar

Because language is conceived here as the production and perception of movement, language is intimately tied to the physical reality of our bodies and

our perceptual systems. Our body and its movements are not just the means by which language is performed, they are also the evolutionary precursors of cognition and language.

If the language in motion framework is to be sufficiently developed to account for a unified view of spoken and signed languages, we need a theory of language that embraces the embodied solution. Such an approach, and the one used here, is cognitive grammar. Cognitive grammar adopts a number of fundamental claims about language that are compatible with the language as motion framework. Cognitive grammar presents an explicitly non-abstractionist view of grammar, offering instead a model based on embodied cognition: "The picture that emerges belies the prevailing view of grammar as an autonomous formal system. Not only is it meaningful, it also reflects our basic experience as moving, perceiving, and acting on the world" (Langacker 2008: 4).

The most fundamental claim of cognitive grammar is that grammar is symbolic. It is important to recognize what is meant by the term symbolic. Within cognitive grammar, a symbol is simply the pairing between a semantic structure and a phonological structure, a meaning and a form (Langacker 2008: 5). Meanings are conceptualizations recruited for linguistic expression. Form, in the cognitive grammar perspective, is the full expressive detail of a usage event, including all the phonetic details, intonation, body language, gesture, "conceivably even pheromones" (Langacker 2008: 457). This symbolic view of language includes words and signs, but it also extends to grammar and includes morphology, grammatical markers, grammatical classes, and syntax.

Although developed as a theory of language, cognitive grammar posits only general cognitive, perceptual, and motoric abilities. In adopting cognitive grammar as one of the foundations of the language as motion framework, I suggest that all of the theoretical and analytic framework of cognitive grammar can be recruited to study gesture. Doing so will provide a unified understanding of language and gesture as manifestations of the human expressive ability.

According to cognitive grammar, one of the functions of grammar is that to impose a particular construal onto conceptual content. This ability to construe situations in myriad ways is based on imaginative and creative abilities. Cognitive grammar eschews purely propositional or truth-conditional accounts of meaning, and instead favors imagistic accounts. One type of such conceptual structure is a set of image schemas, "schematized patterns of activity abstracted from everyday bodily experiences, especially pertaining to vision, space, motion, and force" (Langacker 2008: 32).

In cognitive grammar, meaning is associated with conceptualization. Cognitive grammar uses the term 'conceptualization' to highlight its dynamic nature, "dynamic in the sense that it unfolds through processing time" (Langacker 2008: 32). Conceptualization is embodied in a biologically-implemented brain, and "as neurological activity, conceptualization has a

temporal dimension” (Langacker 2008: 31). The same is true for the phonological pole of linguistic symbols. Cognitive grammar does not adopt an abstractionist notion of phonological structure; instead, the phonological pole captures all of the physical, dynamic aspects of articulatory movements. Thus, the cognitive grammar perspective on meaning as dynamic conceptualization and on phonology as observable phenomena is entirely compatible with dynamic systems theory and the language as motion framework.

3. DYNAMIC SYSTEMS THEORY AND LANGUAGE

Under the language in motion framework, the basic units of speech, sign, and gesture are articulatory gestures which are defined as functional units, an equivalence class of coordinated movements that achieve some end (Studdert-Kennedy 1987: 77). These functionally-defined ends, or tasks, are modeled in terms of task dynamics (Hawkins 1992). In modeling speech, the task may be the formation of a constriction such as bilabial closure; this task involves the coordinated action of several articulators, such as the lower lip, upper lip, and jaw.

Articulatory phonology has significance for language and gesture that goes far beyond describing speech tasks. Other articulators may be modeled in this way as well. For example, the arm and hand can be used to reach for a cup, scratch your head, or to produce a sign or gesture. Whether for speech, sign, gesture, or motor activities unrelated to communication, tasks require the coordinated action of multiple articulators moving appropriately in time and space. These coordinated actions, called coordinative structures, are not hardwired; rather, they are emergent structures in a dynamically changing system.

Another significant aspect of articulatory phonology is that it unifies description over levels that in other theories are seen as distinct. For example, formalist theories such as Chomsky’s minimalist program assume that universal grammar “specifies certain linguistic levels, each a symbolic system” (1995: 167). One such level is a computational system that generates structural descriptions; these structural descriptions are in turn seen as instructions that are fed into another level, the articulatory-perceptual performance system, which specifies how the expression is to be articulated.

Rather than viewing the units of language, whether they are sequences of static, timeless, and non-physical (i.e., mental) units such as phonemes, syllables, morphemes, or words, or non-physical structural descriptions which must be implemented in a performance system, the dynamic view defines language “in a unitary way across both abstract ‘planning’ and concrete articulatory ‘production’ levels” (Kelso *et al.* 1986: 31). Thus, the distinction between competence and performance, which plays such a large theoretical role in generative linguistics, is collapsed into a single system described not in the

machine vocabulary of mental programs and computational systems, but in terms of a “fluid, organic system with certain thermodynamic properties” (Thelen and Smith 1994: xix). As Thelen and Smith go on to observe, the distinction between competence and performance does not make biological sense:

Abstract formal constraints are fine for disembodied logical systems. But people are biological entities; they are embedded, living process. If competence in the Chomskyan sense is part of our biology, then it must also be embodied in living, real-time process (Thelen and Smith 1994: 27).

The language as motion framework provides the theoretical basis for describing language as a dynamic, real-time process. From the language as motion perspective, language is performance.

There is one more implication of the dynamic systems approach to language and gesture. Although speech scientists who work within this theory typically restrict their study to the level of words, describing words as “coordinated patterns of gestures” (Studdert-Kennedy 1987: 78), the theory can be extended beyond words. If words are patterns of gestures, so too are larger, multiword expressions. This observation is especially significant when it is matched with the cognitive grammar claim that all levels of language, from the lexicon to syntax, are symbolic, the pairing of semantic and phonological structures. Grammar, in this view, is schematic patterns of symbolic structures that have both semantic and phonological import. The key point is that grammar always has phonological structure, even if that structure is highly schematic. So, we may now extend the claim of articulatory phonology even further, and say that if words are coordinated patterns of skilled action, and if multiword expressions are yet larger such structures, then grammar itself is coordinated patterns of cognition and action. From the language as motion perspective, grammar is skill.

4. COGNITIVE NEUROSCIENCE

The language as motion framework also has profound implications for cognitive neuroscience. While the abstractionist solution trivializes the role of embodied production and perception in language and grammar, the language as motion framework, by grounding language and gesture in physical systems, claims that cognition and perception must be intimately linked. In this view, “what we perceive is determined by what we do,” and perception is seen as a type of skillful bodily activity (Noë 2004: 1). The same dynamic models that account for the emergence of coordinative structures in skilled movement, such as fluent fingerspelling, speech, sign, or gesture, are recruited to explain cognition. Under this framework, then, “cognition –mental life– and action –the life of the limbs– are like the emergent structure of other natural phenomena” (Thelen and Smith 1994: xix).

This view is consistent with several current theories of brain phylogenetic and ontogenetic development and function. Berthoz (2000: 9), for example, observes that “perception is more than just the interpretation of sensory messages. Perception is constrained by action; it is an internal simulation of action.” In his view, the highest cognitive functions are the evolutionary result of the brain’s ability to skillfully plan movements to meet the needs of future events.

The theory of neuronal group selection or “Neural Darwinism” developed by Gerald Edelman is also consistent with the language as motion framework, including the principles of cognitive grammar and dynamic systems. Three brief examples will serve to demonstrate how Edelman’s theory can be linked to cognitive grammar and to articulatory phonology.

A key concept of neuronal group selection is reentry, “a process of temporally ongoing parallel signaling between separate maps along ordered anatomical connections” (1987: 49). For example, when we eat an apple, our experience maps across several perceptual modalities –the smell, taste, feel, color, and sound of an apple being bitten into– and motor activities –looking at, picking up and holding the apple, bringing the apple to our mouth, opening our mouth and biting, and so forth. Reentry corresponds to the cognitive grammar view of knowledge as encyclopedic, consisting of networks of concepts with no sharp boundaries between semantic and pragmatic, combining experience from multiple sensory modalities and our physical interaction with the world.

Another key concept in neuronal group selection is degeneracy. “Degeneracy is the ability of elements that are structurally different to perform the same function or yield the same output” (Edelman and Gally 2001: 13763). Degeneracy is present in many levels of language, including metaphor and polysemy; the use of different lexical/morphological/syntactic structures to accomplish the same function (e.g., verb aspect); and the aforementioned function of grammar to impose different construals on the same conceptual content. In all of these cases, we find that different structures accomplish the same function, what in speech act theory would be called the perlocutionary effect.

Edelman’s theory is also compatible with the dynamic system approach to articulatory phonology. Edelman (1987: 227) defines gesture as a “degenerate set of all those coordinated motions that can produce a particular pattern that is adaptive in a phenotype.” This view of gesture plays a significant role in Edelman’s theory in two ways. First, it ties motor activity to perception through reentry. Second, Edelman (1989) claims that the brain bases of gestural ordering played a significant role in the evolutionary emergence of language.

The coordination and planning of movement also plays a key role in the theory of the evolution of the brain and language advanced by Rudolfo Llinás. According to Llinás (2001: 17), “the evolutionary development of a nervous

system is an exclusive property of actively moving creatures.” The significance is twofold. First, Llinás links the control of movement to the development of higher cognitive functioning and the mind: “that which we call thinking is the evolutionary internalization of movement” (Llinás 2001: 35). Second, in a proposal that is entirely compatible with language as motion, Llinás also ties the significance of movement to the emergence of language. Llinás (2001: 228) claims that the coordinated movements, or gestures, required for speaking are no different than other motor actions, noting that “the premotor events leading to expression of language are in every way the same as those premotor events that precede any movement that is executed for a purpose.”

5. SETTING LANGUAGE IN MOTION

I have introduced the language as motion framework as a way to unify spoken and signed languages. But the framework does more than that. Language as motion also captures the fact that our perceptual and conceptual capabilities have been shaped by our evolution as moving creatures solving ever more complex problems of emulating and predicting the natural and social environment. Our brains evolved to make sense of the world. Cognition is grounded in perception and motion, and as a result, our conceptual abilities are deeply embodied. These embodied conceptual abilities, which developed from our need to make sense of the world in which we move about, form the basis of the human expressive abilities in all its manifestations, from music, art, and dance to signed language, spoken language, and gesture.

REFERENCES

- AZIZ-ZADEH, L. and DAMASIO, A. (2008), “Embodied semantics for actions: findings from functional brain imaging”, *Journal of Physiology, Paris*, 102(1-3), 35-39.
- BERTHOZ, A. (2000), *The brain's sense of movement*, Cambridge, MA, Harvard University Press.
- CHOMSKY, N. (1995), *The minimalist program*, Cambridge, MA, The MIT Press.
- CHOMSKY, N. (2005), “Three factors in language design”, *Linguistic inquiry*, 36(1), 1-22.
- DAMASIO, A. R. (1994), *Descartes' error: emotion, reason, and the human brain*, New York, NY, Grosset/Putnam.
- DAMASIO, A. (2010), *Self Comes to Mind: Constructing the Conscious Brain*, New York, Pantheon.
- EDELMAN, G. M. (1987), *Neural Darwinism: The theory of neuronal group selection*, New York, Basic Books.
- EDELMAN, G. M. (1989), *The remembered present: A biological theory of consciousness*, New York, Basic Books.
- EDELMAN, G. M., and GALLY, J. A. (2001), “Degeneracy and complexity in biological systems”, *Proceedings of the National Academy of Sciences of the United States of America*, 98(24), 13763.

- HAWKINS, S. (1992), "An introduction to task dynamics", in *Papers in laboratory phonology II: Gesture, segment, prosody*, Docherty, D. L. and Ladd, D. R. (eds.), Cambridge, Cambridge University Press, 9-25.
- JOHNSON, M. (2008), *The meaning of the body: Aesthetics of human understanding*, Chicago, University of Chicago Press.
- KELSO, J. A. S., SALTZMAN, E. L., and TULLER, B. (1986), "The dynamical perspective on speech production: Data and theory", *Journal of Phonetics*, 14, 29-59.
- LANGACKER, R. W. (1987), *Foundations of cognitive grammar, Volume I, Theoretical Prerequisites*, Stanford, Stanford University Press.
- LANGACKER, R. W. (1991), *Foundations of cognitive grammar, Volume II, Descriptive application*, Stanford, Stanford University Press.
- LANGACKER, R. W. (2008), *Cognitive grammar: A basic introduction*, Oxford, Oxford University Press.
- LLINÁS, R. (2001), *I of the vortex: From neurons to self*, Cambridge, Mass, MIT Press.
- NEISSER, U. (1967), *Cognitive psychology*, NY, Appleton-Century-Crofts.
- NOË, A. (2004), *Action in perception*, Cambridge, MA, The MIT Press.
- SPIVEY, M. (2007), *The continuity of mind*, Oxford, New York, Oxford University Press.
- STUDDERT-KENNEDY, M. (1987), "The phoneme as a perceptuomotor structure", *Language perception and production: Relationships between listening, speaking, reading, and writing*, D. A. Allport (ed.), London, Academic Press, 67-84.
- THELEN, E., and SMITH, L. B. (1994), *A dynamic systems approach to the development of cognition and action*, Cambridge, MA, The MIT Press.