A framework for unifying spoken language, signed language and gesture

Sherman Wilcox *
André Nogueira Xavier **

Abstract: In this paper we offer the first steps in a framework for unifying spoken language, signed language, and gesture. The framework is based on three existing theories: dynamic systems theory, cognitive grammar, and non-Cartesian approaches to cognitive neuroscience. Our article posits a human expressive ability, which we claim is based on the need of moving creatures to comprehend their environment, resulting in a conceptual system embodied in perceptual and motor systems. This human expressive ability underlies language and gesture. We explore relations between language and gesture as emergent systems.

Keywords: signed language; gesture; cognitive grammar.

Introduction

As often happens in the history of scientific progress, discoveries lead to a brief period of clarity that soon disappears as answers lead to new questions. Such is the case in the field of linguistics. When language was understood to be that which is produced by the vocal tract, distinct from gestures produced by the hands and body, the picture was clear. Discoveries by psychologists and others led to new problems, however, as the line between linguistic (speech) and nonlinguistic (gesture) became less clear. The scene was complicated once again when linguists added signed languages to the picture. Not only was the definition of language expanded beyond speech and the vocal tract, the line between language and gesture is complicated even

* Ph.D em Linguística Educacional pelo Departamento de Linguística da University of New Mexico (Estados Unidos). *E-mail*: wilcox@unm.edu.

** Doutorando em Linguística na Universidade Estadual de Campinas (Unicamp) e bolsista do Conselho Nacional de Desenvolvimento Científico (CNPq).

further since signs and gestures are produced by the same articulators and are, in some instances, difficult to distinguish.

We believe there is a way to resolve these new complications by bringing three theories to bear on the problem of unifying spoken language, signed language, and gesture:

1. First, we require a theory of language that can encompass spoken and signed languages as well as gesture. The theory should not rely on abstractionist solutions which posit an abstract symbol set devoid of material substance. Instead, we insist that the proper approach is an embodied solution that unifies at the level of the physical performance of language. The theory we use is cognitive grammar (LANGACKER, 1987, 1991, 2008).
2. We also need a theory for understanding physical performance as skilled action, and which can be applied to a view of language as performance and grammar as skill. For this we select dynamic systems theory (THELEN; SMITH, 1994; SPIVEY, 2007).
3. To account for how language and gesture are implemented in the brain as action, we require a non-Cartesian, embodied theory of cognitive neuroscience. We believe that 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 Llinás (2001), Berthoz (2000), and Damasio (AZIZ-ZADEH; DAMASIO, 2008; DAMASIO, 1994, 2010).

In this article we focus on the first step, applying cognitive grammar to the problem of the relation between language and gesture.

The human expressive ability

We believe that human language is grounded in a human expressive ability. We reject the notion that human language and the human expressive ability arose suddenly, "effectively instantaneous, in a single individual, who was instantly endowed with intellectual capacities far superior to those of others" (CHOMSKY, 2005, p. 12). Rather, we claim that this ability has its ancestral source in a general comprehension ability, based on an organism's need to make sense of its environment in order to survive, an ability that arose in Darwinian fashion through natural selection. This comprehension ability is driven by the fact that we are mobile creatures. As the neuroscientist Rudolfo Llinás (2001, p. 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". We 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, p. 32).

The emergence of the ability of moving creatures to make sense of their world was a major, perhaps the primary, factor in the development of the human brain. Again, Llinás (2001, p. 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. Such embodied cognition is the motive force driving the human expressive ability.

The human expressive ability includes a vast range of abilities and domains of expression, from dance to music, art, mathematics, and language. We focus here on only three forms of expression: spoken language, signed language, and gesture. In thinking about how to conceive of these three systems as manifestations of embodied cognition, we have come to rely on a "cloud metaphor". We envision spoken language clouds, signed language clouds, and gesture clouds. Real clouds are complex systems comprised of many factors, such as air currents, temperatures, the behavior of water molecules and dust particles, sunlight and its interaction with various landscapes, and more. These complex interactions lead to structure, not in an instructionist way, but as an emergent phenomenon. In spite of the fact that there are no instructions, no rules, no "innate" universal grammar of cloud formation specified in any part of the complex system, cloud formation leads to patterns common enough to have been classified and given names such as cumulus, stratus, cirrus, nimbus, and so forth (THELEN; SMITH, 1994, p. xix):

> *For example, in certain meteorological contexts, clouds form into thunderheads that have a particular shape, internal complexity, and behavior. There is a clear order and directionality to the way thunderheads emerge over time. [...] But there is no design written anywhere in a cloud [...] There is no set of instructions that causes a cloud or a group of plants and animals to change form in a particular way. There are only a number of complex physical and biological systems interacting over time.*

We use clouds as metaphors because, like clouds, language and gesture are also complex systems in which structure emerges in a dynamic way. Although language shares some characteristics with and interacts with gesture, it forms its own system and subsystems (language types, specific languages, specific dialects, specific idiolects, and so forth).

Extending the metaphor, we envision the process by which language and gesture clouds "condense" or emerge, and their characteristics, to be determined by a number of factors. Among these factors are: means of expression, whether defined by the articulators used (vocal tract vs. hands, face, and body); channel of transmission (acoustic vs. optic; certainly other channels are available for transmission of language and gesture, for example the kinesthetic channel for deaf-blind people); entrenchment and unit status; conventionalization; integration, schematicization, conceptual autonomy-dependency, and others.

We emphasize that although it may seem that we are describing these expressive clouds as distinct systems (a "language cloud", a "signed language cloud", a "gesture cloud" etc.), they all are manifestations of the same embodied conceptual system that drives the human expressive ability. Returning to real clouds, whether we see the formation of cumulus or nimbus clouds, cloud formation can be accounted for as the product of a complex thermodynamic system. Similarly, the same embodied conceptual system drives the formation of clouds of human expression. Thus, it is not surprising that these expressive clouds share characteristics, express similar or complementary meanings, interact extensively, merge with or transform into one another under certain conditions, and so forth.

Finally, we emphasize that, as for real clouds, no one factor is criterial in determining the type of system that emerges. Rather, the characterization of "language clouds" versus "gesture clouds" requires a prototype model. For example, while the "spoken language cloud" is prototypically produced by the vocal tract and the "gesture cloud" by the hands, face, and body, elements of the gesture cloud may enter the spoken language cloud. Consider an expression such as (1) in which a gesture using two hands to indicate size accompanies the spoken expression:

"I caught a fish this big [gesture]."

Here, not only is gesture used alongside speech, but the gesture fills a grammatical role.

In the case of "signed language clouds" this distinction fails entirely, since sign and gesture are produced with the same articulators. As is the case for real clouds, clouds of human expression do not have hard boundaries or criterial features.

Language and gesture

In the previous section, we suggested that human expression, as a manifestation of embodied cognition, can be conceived of as emergent, cloud-like structures with certain prototypical features. We also proposed a preliminary set of factors that lead to the formation of these structures. Here, we describe these factors in more detail, explore how they interact, and suggest how they lead to the formation of the identifiable clouds of human expression we call spoken language, signed language, and gesture.

Language and gesture as structured systems

A fundamental claim of cognitive grammar is that all aspects of language, from the lexicon to grammar, is symbolic in nature. Within cognitive grammar, a symbol is simply the pairing of a semantic structure and a phonological structure, a form and a meaning (LANGACKER, 2008, p. 5). Meanings are conceptualizations recruited for linguistic expression. Form, in the cognitive grammar perspective, is the full perceptible detail of an utterance, including intonation, body language, gesture, "conceivably even pheromones" (LANGACKER, 2008, p. 457). Thus, although we are not discussing musical expression here, we could equally well refer to the meaning of a musical phrase as its semantic structure, and the acoustic signal as its phonological structure; when paired, the two form a musical symbol. What is relevant for our discussion is that not only do spoken and signed languages consist of symbolic structures, so do gestures – they are the pairing of semantic structures and phonological structures. Within the framework we are developing, such symbolization is a foundational feature of the human expressive ability.

In cognitive grammar, the grammar of a language is "a structured inventory of conventional linguistic units" (LANGACKER, 1987, p. 37). We take this definition not only as the basis for understanding the grammars of spoken and signed languages, but also as the starting point for understanding the nature of gesture as a manifestation of the human expressive ability. 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 our framework, we suggest that all of the theoretical and analytic framework of cognitive grammar can be recruited to study gesture. The extent to which this strategy can be applied to gesture is an open question, but we believe exploring where it does work and where it does not will reveal aspects of gesture, and the relation between gesture and language (both spoken and signed), that have hitherto gone unexamined. Doing so will also provide an overarching framework for understanding language and gesture as manifestations of the human expressive ability.

Langacker provides a word-by-word analysis of this definition of grammar. Here, we will focus on three key concepts: structure, unit, and conventionality. We have already touched on the structured nature of language in our claim that linguistic structure is an emergent phenomenon. Langacker (1987) also provides a detailed discussion of other aspects of structure, by which he means that some linguistic units function as components of others. He describes three types of relations among the components of a complex structure: symbolization; the integration of component structures into a composite structure; and categorization (which is analyzed in terms of schematicity). We have already given a preliminary description of the central role that symbolization plays in language. Schematicity is a fundamental aspect of structure. We take up a fuller discussion of schematicity in section “Schematicization”.

Integration of component structures into composite structures can take place at either the phonological or semantic pole of symbolic structures, or the integration can be of componential symbolic units themselves into more complex symbolic units. The integration of phonological segments to form a word is one example. In signed languages, an analogous integration would be combining the componential parameters of a sign, for example a handshape, location, and movement, into a composite phonological structure. Integration of component structures into composite structures with greater complexity is a central aspect of grammar, and a topic well-studied by linguists describing spoken and signed languages. One question that could be posed is whether integration can take place across linguistic and gestural systems. We propose that such integration does indeed take place, and should be studied as such. We will discuss this in more detail in section “Structure of the gesture cloud”.

Entrenchment and unit status

A unit is a structure that has been mastered to the extent that it has become automatic, such that the user does not have to attend to the individual parts that make up the unit. Cognitive grammar describes the process of automaticization as one of entrenchment through repeated use. A complex structure that formerly consisted of a large number of individually controlled elements is now treated as a single unit, one that is “thoroughly mastered, to the point that using it is virtually automatic and requires little conscious monitoring” (LANGACKER, 2008, p. 16).

Entrenchment is quite a general concept. It is related to the dynamic systems concept of entrainment, whereby a system with a large number of degrees of freedom is reduced to far fewer degrees of freedom. Entrenchment may be mo-

toric or conceptual, and it operates across linguistic and non-linguistic domains. Learning to put a key in a car's ignition and push it in while simultaneously turning it to start the car is an example of a non-linguistic routine that at first may be learned as individually-controlled sequences of motor behavior; with practice, this complex motor structure becomes automatic. Rather than thinking about each element of the sequence, the skilled driver simply executes the entrenched unit "start the car". Learning the mathematical times table is an example of a cognitive unit. Learning to recite the alphabet is an example of linguistic and motoric entrenchment.

Conventionalization

Entrenchment takes place at the level of the individual. A linguistic unit or gesture is conventional to the extent that it is shared and known to be shared among a community of users. While conventionality pertains to the linguistic (or gestural) community, the assessment of conventionality is carried out by individuals: an expression is accepted as conventional to the extent that it conforms to the units invoked for the purpose of apprehending expressions (LANGACKER, 2008, p. 227). This determination of conventionality is the process of categorization between some aspect of a usage event and a linguistic or gestural unit: the relationship is conventional (well-formed, or grammatical) to the extent that the element produced in the usage event is an elaboration of, is sanctioned by, a linguistic or gestural unit.

A problem develops, however, in determining how the categorizing unit is selected. For example, a particular target production tends to activate a set of units, in our case either linguistic or gestural units, which could serve to categorize the target. Langacker offers three factors that encourage a unit's selection as the categorizing structure: degree of entrenchment, the influence of context, and degree of overlap with the target.

If a particular structure is highly entrenched in an individual, it will increase the ease with which that structure is activated and serve to successfully categorize the target. The influence of context works by also increasing the likelihood that a particular structure will be selected to categorize the target. The final factor is degree of overlap with the target. The more features a categorizing unit shares with the target, the more likely it will be selected to categorize the target.

It is especially important to consider effects of contextual and degree of overlap in the present context. Consider the following, true story. A group of signers were having breakfast at a popular local restaurant named Goodie's. They had been going to this establishment for many months, and they always fingerspelled the name. During breakfast conversation, one person suggested coining a sign for the name of the restaurant. At the same time, another person, unaware of this conversation, had accidentally dropped some jelly on the pinky finger of his non-dominant hand and was attempting to wipe it off with the extended pinky finger of his dominant hand. Seeing this, the group proclaimed, "Yes! That's a good one" – and in fact it became the sign-name for Goodie's for many years.

The example serves to make several important points. First, even though the sign had been newly coined and was not an existing American Sign Language (ASL) sign, it was categorized as a well-formed sign because it used conventional phonological formational patterns: the handshape, point of contact, and swiping

motion. However, the signer had not in fact produced an ASL sign. He had performed a gesture, and a gesture that was not intended to communicate but only to serve the instrumental function of cleaning his finger. In this situation, though, the overlap in means of expression (hands), and the influence of context (taking place during a discussion the point of which was to coin a new sign), caused the group of signers to select this gesture as the categorizing structure for a linguistic unit.

Schematicization

Schematization is the "process of extracting the commonality inherent in multiple experiences to arrive at a conception representing a higher level of abstraction" (LANGACKER, 2008, p. 17). Schematization plays a central role in all facets of language, grammar, and in gesture as well. Schematicization can be carried out to any degree. It works across phonological, semantic, and symbolic structures. We each have schemas for the pronunciation of ring. Learning the word ring as an item of jewelry worn on the finger, and hearing it used for a piece of circular metal worn through the nose, we develop a more schematic value for the meaning of ring. The process proceeds further as we hear it used to mean other types of circular objects.

Schemas form the basis of grammar in cognitive grammar. What would be called grammatical rules in other theories are regarded as symbolically complex schemas, or templates, in cognitive grammar. Schematicity also interacts with symbolic complexity. Symbolic assemblies can be specific or schematic. Words, phrases, clauses, and larger utterances, in the actual usage event, are examples of specific symbolic assemblies with varying degrees of complexity. Constructional or grammatical schemas are examples of more schematic symbolic assemblies.

Means of expression and channel of transmission

In order to function as communication, symbolic structures require a perceptible phonological pole, whether this is defined by the articulators used to produce the signal (vocal tract vs. hands, face, and body) or the channel of transmission of the signal (acoustic vs. optic). For spoken language, the prototypical means of expression is by vocally produced acoustic signals. However, not all vocally produced sounds are linguistic. Consider example (2):

"She was so frustrated, she was like, [CRY OF FRUSTRATION]"

The imitation of her cry is vocally produced, but would not be considered language. On the other hand, the use of deictics in face-to-face communication require an accompanying gesture, as when saying, "This one and this are almost the same". In this case the linguistic expression is incomplete, incomprehensible without the accompanying gesture.

Signed languages are prototypically produced with the hands, face, and body rather than by means of vocal sounds. However, even here the situation is not entirely clear-cut, because certain vocal sounds conventionally accompany some signs, such as the ASL sign commonly glossed by its accompanying vocal sound "Pah!"

Gestures are perhaps the most complex of the three. While most people would generally agree that gestures are prototypically with the hands, face, and body, a great deal of research has investigated "vocal gestures" – not those articulatory gestures described by speech scientist, but productions such as vocally-produced sound effects and ideophones (VOELTZ; KILIAN-HATZ, 2001). Also, in the case of ideophones, the interaction and alternation of gesture and speech is complex and still little understood. Kunene (2001, p. 183), for example, concludes that

> *[...] the ideophone is the closest linguistic substitute for a non-verbal, physical act. I find that this position is given great credence by the fact that such physical acts, by which I mean self-conscious gestures and other imitative acts, are more often than not seen to accompany ideophones and, indeed, sometimes to replace them completely.*

SIGN AND GESTURE

Although developed primarily for spoken languages, the principles of cognitive grammar described above clearly apply as well to signed languages (see FERRARA, 2012 for an example). Our claim is that these same principles, because they are based on general cognitive abilities not specific to language, can be applied as well to analyzing gesture.

Structure of the gesture cloud

We are suggesting that language clouds and gesture clouds are manifestations of the same underlying conceptual system that is the basis for the human expressive ability. Thus, we propose that the general principles of cognitive grammar can be applied to the study of gesture. Here, we provide only a first, suggestive look at how that application might proceed.

Integration: The identification and description of component structures and their integration into composite structures is far less understood for gestural expression. Some researchers even state that there is no structural integration and no complex symbolic structures to be found in gesture. McNeill (1992, p. 21), for example, claims that "gestures are noncombinatoric: two gestures produced together don't combine to form a larger, more complex gesture". Naturally, the claim hinges on what one defines as a gesture and what is meant by "produced together". For example, if we include facial gestures, it is surely the case that these combine with manually produced gestures to produce higher-level composite gestural structures. Also, a number of gesture researchers offer evidence suggesting that integration of componential into composite structures occurs in gesture (KENDON, 1997; ENFIELD, 2004; WEBB, 1996).

It seems to us not unreasonable to suggest that gesture likewise consists of such integration of componential parameters. According to Aldrete (1999, p. 36-37), Quintilian taught that "by altering the speed with which a gesture was made and its range of motion, the same gesture could have multiple meanings or purposes" and that "this strategy of modulating the speed of gesture in order to express slightly different meanings was used to give versatility of denotation to several basic gestures". We regard this process as an example of the integration of gestural components into a higher level composite structure.

Finally, we suggest that a revealing analysis of co-speech gesture would be to regard it as the integration of linguistic and gestural components into a higher level composite symbolic structure. We suggest using the analytic tools of cognitive grammar to understand how linguistic and gestural components are integrated into a higher level composite "co-speech" structure – although, from this perspective, the structure is neither "co-speech" nor "speech-accompanied gesture", but a composite symbolic unit. For example, we would want to understand the phonological integration of the component structures, noting the tight integration at both phonological poles (the temporal synchronicity of speech and gesture). We would also need to examine the full detail of the componential semantic structures. Finally, this approach requires that we understand how the integration takes place. Typically, integration involves one structure that is conceptually dependent on another. The preposition near, for example, is dependent because it makes internal reference to some type of distance relation, either real or metaphorical, involving two entities which must be elaborated. Each of those entities are elaborated by autonomous component structures. In the expression table near the door the autonomous structures are table and the door. We suggest that a similar analysis of the integration of linguistic and gestural componential structures would be revealing.

Entrenchment: As we have seen, entrenchment is a very general notion that applies not only to language but to non-linguistic actions and conceptualization. Gestures also exhibit various degrees of entrenchment. While writing this section, the first author stood up to put on his shoes. The action of putting on and tying a pair of shoes requires several entrenched gestures. We all are familiar with the process that children go through as they learn these actions, particularly the difficult sequence of actions required to perform the shoestring-tying. For the author, these actions are fully entrenched and require little if any cognitive effort. He simply performs the action "tie my shoes" as an entrenched motor skill.

The same description of entrenchment applies to symbolic gestures. The mascot of the first author's university is a wolf or lobo. Students routinely make a "lobo" gesture at sporting events, with the meaning "Go Lobos, beat the other team!" which consists of touching the thumb to the flexed middle and ring fingers while the index and pinky fingers are fully extended. We have often witnessed new students struggle to perform this gesture accurately. For these students, it is not yet an entrenched gesture.

One way that the application of entrenchment to gesture may elucidate prior problems, we suggest, is the connection between entrenchment and intentionality. Entrenchment is a matter of control or automaticization. Gestures, in a broad sense, need be neither intentional nor communicative. Tying one's shoe is intentional but not communicative. Edelman (1987, p. 227), however, defines gesture independently of both intentionality and communicativeness, as a "degenerate set of all those coordinated motions that can produce a particular pattern that is adaptive in a phenotype". Functional and adaptive gestures, such as the movements that propel a fish as it swims, may be automatic but they are not entrenched in the same sense that entrenched gestures such as the lobo gesture or tying a shoe are. In the former case, the process takes place by natural selection across individuals; in the latter, it occurs within an individual, as gestures are repeated and acquire unit status, capable of being automatically controlled.

To intentionally produce a gesture, whether for communicative purposes ("Go Lobos!") or only for functional need (tying a shoe) requires the ability to control that gesture, at least to a certain extent when it is first produced, and repeatedly producing that gesture means that the gesture will eventually become entrenched. Intentionality and entrenchment go hand in hand. We suggest that if gesture researchers, rather than defining gesture as intentionally produced communicative actions, would "unwrap" the concepts of intentionality, communicativeness, and study them in the context of entrenchment, we might better understand how these factors come together evolutionarily and developmentally in the production of communicative gestures.

Conventionality: Conventionality also applies to gesture, and gestures vary along a continuum of conventionality. While ad hoc gestures may be spontaneously "coined" (but see our discussion of schematicity and gesture the next section), many other gestures attain a high degree of conventionality, including those that have been labeled emblems (McNEILL, 1992; KENDON, 2004).

Conventionality has sometimes been proposed as the means for distinguishing between the signed language system and the gestural system (OKRENT, 2002). Clearly, this will not work. Units within each system vary along the dimension of conventionality. A newly coined sign is not by definition a gesture, and a newly coined spoken word is certainly not a gesture. Okrent's solution to the problems concerning conventionality involves positing a "modality-free" concept of gesture. We resist such an abstractionist solution. Linguistic and gestural systems grow and develop in a bottom-up fashion from usage events. Usage events are not modality-free. Even if these systems develop high level schemas they do so on the basis of fully embodied, physical actions; furthermore, even high level schemas will retain aspects of actual usage events.

Schematicity: As we have seen, schematicity is a foundational concept in cognitive grammar. Although rarely discussed explicitly in the literature, schematicity also pertains to gesture. Indeed, we suggest that some of the problems in understanding gesture and its relation to spoken language (e.g., whether gesture exhibits integration), derive from the lack of attention that schematicity has received by gesture researchers. We offer two preliminary examples of how schematicity plays a role in gesture.

Beats may be produced in a variety of ways: with the entire forearm and hand, with various types of movements, with only the hand, with the hand in various handshapes, and so forth. Whichever articulators are used to produce the beat, the temporal structure of the beat gesture coincides with the temporal structure of the speech signal. Clearly, beats have a schematic phonological structure. Likewise, beats have a conceptual or semantic structure which is also schematic. One meaning that beats convey is to mark the significance of some aspect of the semantic structure expressed by the accompanying speech signal. Thus, the semantic structure of a beat is dependent, making reference to and elaborated by an autonomous structure provided the semantic structure of the spoken language which accompanies the gesture.

A second example of schematicity involves the recurrent cyclic gesture studied by Ladewig (2011). Ladewig found that the cyclic gesture in a variety of usage events exhibited systematic variation in form and meaning. This finding provides evidence that there is a schematic cyclic gesture. Phonologically, the cyclic gesture may be produced in a variety of ways, but all share the common

feature of a circular, cyclic movement. Semantically, the gesture exhibits a schematic meaning grounded in an idealized cognitive model "consisting of the image schema cycle and its metaphoric extensions" (LADEWIG, 2011, p. 15). As for all schemas, the cycle image schema "has emerged from the abstraction of our experience of recurrent events" (LADEWIG, 2011, p. 15).

Gesture and sign, and gesture to sign

There are two ways in which researchers have suggested that gesture and sign interact, either synchronically or diachronically. In the first, gesture and sign may co-occur, either simultaneously or alternating. In the second, researchers have examined the ways in which gesture becomes incorporated into the linguistic systems of signed languages. Typically these diachronic studies come under the rubric of lexicalization and grammaticalization

Gesture and Sign

A growing body of research examines how gesture and sign interact synchronically. A review of that literature is beyond the scope of this article (but see McCLEARLY; VIOTTI, 2010). Vermeergen and Demey (2007) offer an excellent review with new data from Flemish Sign Language.

One of the most provocative proposals for the synchronic relation between gesture and sign is offered by Liddell (2003), who argues that sign and gesture are integrated in several ways, including aspects of spatialized syntax, pointing or indexical signs, and classifier constructions. In general, approaches such as Liddell's adopt a criterial model of language and gesture, assuming that linguistic material is categorical, discrete, and countable, while gestural material is gradient, analog, and uncountable. Liddell, for example, argues that the locations in pointing signs cannot be morphemic because they are uncountable. He applies the same analysis to classifier signs, arguing that while parts of these signs (e.g., handshape) are linguistic, other parts (again, locations) are "variable, gradient elements" (LIDDELL, 2003, p. 212) and should be classified as gesture.

We reject such criterial models, which we believe are doomed to failure. Where would intonation be classified under such a model – as language or gesture? Intonation is largely gradient and analog. While some might propose that intonation is correctly classified as gestural, the same question can be posed for other aspects of language. For example, Bybee (2010, p. 2) observes that "all types of units proposed by linguists show gradience, in the sense that there is a lot of variation within the domain of the unit (different types of words, morphemes, syllables) and difficulty in setting the boundaries of the unit". Hay and Bayeen (2005, p. 346) ask whether morphological structure is inherently graded, and reply, "the issue is controversial, but the evidence that is currently accumulating in the literature suggests that the answer is yes". This leads them to conclude that "accepting gradedness as part and parcel of the grammar entails a paradigm shift for linguistics" (HAY; BAYEEN, 2005, p. 346).

If we adopt the position that if any aspect of linguistic structure exhibits gradedness it must be classified as gesture, we are led to the nonsensical conclusion that all morphological structure, indeed grammar itself, is gestural.

Gesture to Sign

We can also look at the relation between sign and gesture diachronically. This view examines how gestures may become lexicalized and grammaticalized into the linguistic system.

Lexicalization of gesture in signed languages is well documented (JANZEN, 2012). We offer one example of the lexicalization of gesture in Brazilian Sign Language (Libras). The thumb-up gesture performed with one hand is an emblem for Brazilian hearing people. It is also used by Brazilian deaf people. When used by deaf people communicating to hearing people, it seems obvious that this is a conventional gesture. When deaf people use the emblem with each other to express the same range of meanings that hearing people do, we are agnostic on its status as language or gesture. However, this gesture also appears to be the source for several lexical signs in Libras. Thus, we see expressions that range from gesture to language:

- Emblem
  - THUMB-UP (1-hand)
- Emblems but with linguistic processes
  - THUMB-UP(2-hand: often seen at the beginning of a YouTube video
  - THUMB-UP(2-hand semicircular): meaning "Are you all okay?" (presumes a group of deaf people sitting in a semicircle)

  Blend of a lexical sign and an emblem
  - HOW-ARE-YOU: the sign GOOD plus THUMB-UP(1h)

  Lexical signs
  - IMPROVE
  - CRITICIZE/EVALUATE: merger of THUMB-UP and THUMB-DOWN
  - THUMB-UP-and-THUMB-DOWN: conveying the meaning "both the good and the bad aspects of something".

Once lexicalized, gesture may also undergo the process of grammaticalization. Several researchers have documented the process by which lexicalized gestures grammaticalize (JANZEN, 2012; JANZEN; SHAFFER, 2002; SHAFFER; JARQUE; WILCOX, 2011; WILCOX, 2004, 2005, 2007; WILCOX; ROSSINI; ANTINORO PIZZUTO, 2010; WILCOX; WILCOX, 1995; PFAU; STEINBACH, 2006). In general, the process starts with a manually produced gesture which enters a signed language as a lexical morpheme. That lexical sign then acquires grammatical meaning. For example, it has been proposed that a departure gesture used in the Mediterranean region entered French Sign Language (LSF) as the lexical sign PARTIR "leave". Because American Sign Language (ASL) is historically related to LSF, the sign also appeared in ASL at the turn of the 20th century with the lexical meaning "to depart". It also occurs with a more grammatical meaning, marking future.

Wilcox (2004, 2005, 2007; WILCOX; ROSSINI; ANTINORO PIZZUTO, 2010) has suggested that in addition to manual gestures becoming grammaticalized, a second route leads from gesture to language. This route begins as either facial gestures or manner of movement gestures (manual gestures, but in this case it is the manner of movement that grammaticalizes). These gestures do not enter the linguistic system as lexical signs; rather, they first appear as prosody or intonation. As they grammaticalize, they take on grammatical meaning as markers of interrogatives, topics, conditionals, verb aspect, intensification, and so forth.

Constructed action

One of the more problematic areas in which gesture and sign interact is in so-called constructed action. Constructed action has been defined as "the reporting (usually via a demonstration) of another's actions" (QUINTO-POZOS, 2007, p. 1288). Quinto-Pozos (2007, p. 1285-1286) offers an example from a story narrated in ASL by Ben Bahan, a Deaf storyteller:

> *As one example, Bahan portrays the bird's hunting skills by depicting the preparatory actions that the bird takes before swooping down from the sky in order to capture prey on the ground. In that depiction, Bahan places his hands on either side of his upper torso (just as children often do when they portray the body of a chicken) and emphatically crooks his head downward and rightward as if glaring down at the awaiting prey. The viewer is to understand that those are the head and wing actions that the bird performs in preparation for a downward swoop toward its prey.*

One of the issues concerning constructed action is its linguistic status: is it sign or gesture? Our simple answer is: "Yes". Our answer may seem facetious, but we mean that constructed action is both language and gesture in at least two senses. First, as we've noted above, the cognitive grammar view of the phonological pole of usage events is broad enough to include gesture. Also, even in the case of spoken language in which conventional linguistic units are prototypically produced by the vocal tract, it is entirely possible for gesture to serve a linguistic function, for example by filling grammatical roles, as in example (2) in section "Means of expression and channel of transmission".

However, we mean that constructed action is both gesture and language in a more complex way. Our proposal is that whereas constructed action usage events such as the example above may begin life as gestural depictions, their repeated use by signers in certain contexts and genres leads to schematization. We suggest that several factors affect how this process works.

First, we note that depictions may exhibit different degrees of symbolic or conceptual complexity. The depiction portrayed by Bahan is complex. Other depictions are less complex, such as demonstrating the action of brushing one's teeth, climbing a ladder, or riding a bicycle. When these symbolically less-complex constructed actions are repeated, they tend to lexicalize. Because the variation of these usage events is constrained, schematizing across multiple usage events results in a low-level, symbolically simple lexical schema — for example, lexical signs such as BRUSH-TEETH, CLIMB-LADDER, or BICYCLE. Symbolically-complex constructed actions resist lexicalization because the variation across usage events is too great to develop a low-level, lexical schema (see Figure 1).

The mechanism by which this differential entrenchment/conventionalization takes place is worth discussing in a bit more detail. To do so, we would like to reframe the description in terms of autonomy-dependency and to exemplify the process with the relation between handshapes, movements, and manner of movement. We note that of the three, handshape has been reported to be most susceptible to phonemicization (or to function more categorically see, e.g. (SCHEMBRI; JONES; BURNHAM, 2005, p. 286)), while movement is less so. Manner of movement has been claimed to play only a marginal role in distinguishing lexical signs (KLIMA; BELLUGI, 1979, p. 309). We attribute this to the autonomy and dependency characteristics of these three aspects of a sign. As background, we note that cognitive grammar posits experientially grounded

conceptual archetypes. Two such archetypes are the noun archetype and the verb archetype (LANGACKER, 2008, p. 103). These archetypes are grounded in the conceptual properties of objects and events respectively. Objects have a number of prototypical properties, but the one that concerns us here is that an object is "conceptually autonomous, in the sense that we can conceptualize it independently of its participation in any event" (LANGACKER, 2008, p. 104). Events are conceptually dependent; they cannot be conceptualized without conceptualizing the participants who interact to constitute the event.

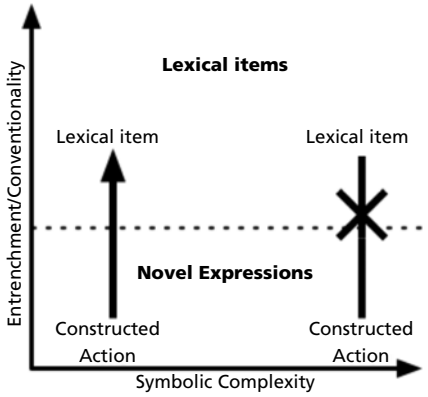

Figure 1 – Degree of symbolic complexity and entrenchment/conventionality

When applied to the articulators of signed language, we note that handshapes, as objects, are conceptually autonomous. Movements, as events, are conceptually dependent: something must move. By extension, manner of movement is also conceptually dependent: in order to move in a certain way, some type of movement is required.

We suggest that these characteristics play a role in determining susceptibility to phonemicization. In a usage-based model such as cognitive grammar, usage events are the source of all linguistic units. Entrenchment drives this process:

> *[...] units emerge via the progressive entrenchment of configurations that recur in a sufficient number of events to be established as cognitive routines. Since only recurring features are reinforced, the units that emerge are far less comprehensive and detailed than the usage events giving rise to them* (LANGACKER, 2008, p. 220).

The key factor is the frequency with which usage events occur. This might lead one to ask what the frequencies are of handshapes vs. movements vs. manner of movements. This is not, however, the correct comparison. Because of the autonomy-dependency properties of these three aspects of sign formation, movements cannot appear without handshapes, and manner of movement cannot appear without movement (and thus also a handshape). Thus, the correct comparison is of sets with increasing degrees of freedom, or variability, from [handshapes] to [movement+handshape] to [manner of movement+movement+handshape]. Location interacts with both handshape and movement, making the degrees of freedom or variability of location quite high.

This fact about usage events of formational parameters affects the frequency of recurring features, and thus accounts for the more categorical behavior of handshapes as compared to movement. The schemas that emerge for [handshapes] are necessarilymoredetailedthanthosefor[mannerofmovement+movement+handshape]. We also predict similar effects for movement type, manner of movement, and location. We suggest that this relationship is what leads Liddell to classify location as more gesture-like, less discrete and less countable, and therefore less linguistic. Johnston and Schembri (2007, p. 165) summarize the point and also note the same issue for movement in depicting verbs:

> *Liddell (2000, 2003) pointed out that the number of locations used in these verbs was potentially unlimited. As we will see below, sign language researchers encounter the same problem with all the many possible spatial arrangements found in depicting signs. It is also difficult to provide a complete list of all the movement components that are possible in these forms, because in many cases, a depicting verb of motion may imitate a large variety of types of possible movement. This makes the meaningful uses of location and movement unlike the identifiable and listable morphemes that may be found in a dictionary of English.*

We now apply this analysis to the development of constructed action schemas. When symbolically-simple constructed actions are schematized, the resulting schemas are more detailed and appear as lexical units. When symbolically or conceptually complex constructed actions occur as usage events, a different schema emerges. When signers extract the commonality inherent in multiple experiences of these symbolically complex expressions with fewer recurring features in common, the schema that emerges is at a much higher level (see Figure 2). In other words, the unit or schema is far less comprehensive and detailed than the usage events that give rise to it. We call this a constructed action grammatical schema or a constructed action scenario.

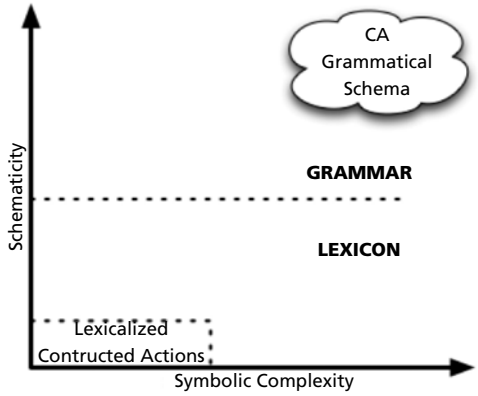

Figure 2 – Constructed Action (CA) Grammatical Schema

We take the term scenario from Langacker's (2008, p. 470) analysis of speech acts in discourse. Langacker notes that speech acts are based on standard cultural models which encompass familiar scenarios of social linguistic interaction. These scenarios are linguistic schemas: "Schematically, they represent any pre-

condition required for the performance of the act" (LANGACKER, 2008, p. 471). One such speech act scenario is the promising scenario. For example, if a home-owner provides specific instructions on how a remodel job should be performed, the contractor might reply with the promise scenario: [Usage Event [Promise Scenario [I'll do that]]], where the specific expression in the actual usage event will vary: If that's how you want it done, that's what I'll do; I'll do that; or simply, Sure!

Constructed actions, like speech act scenarios, are based on standard cultural models of narrative style and they represent the preconditions for the performance of the constructed action. If "complex scenarios … have the status of conventional linguistic units" (LANGACKER, 2008, p. 472), so does the constructed action scenario. The constructed action scenario is a schematic conventional linguistic unit.

Thus, our argument is that the schemas which develop across multiple occurrences of constructed action usage events are necessarily high-level grammatical schemas or discourse scenarios. If a speaker wants to make a promise she uses the promise scenario, a conventional linguistic schema. Her grammar will also provide grammatical schemas for sanctioning different promise expressions. If a signer wants to depict the actions taken by some person, she uses the constructed action scenario. Again, her grammar will provide grammatical schemas for sanctioning different constructed action expressions.

There would appear to be a difference, however. When the contractor instantiated the promise scenario in the previous example, he did so with language: If that's how you want it done, that's what I'll do. Notice, however, that the promise scenario can as easily sanction a gesture (or, put the other way around, a gesture can instantiate the promise scenario). The contractor might simply give a thumb-up gesture, promising to follow the home-owner's specifications. Still, though, the alternation seems to be categorical: either language or gesture can fulfill the promise scenario. When signers use constructed action, the expressions can range from those that are quite gestural to those that are more linguistic, for example, the use of polycomponential verbs (QUINTO-POZOS, 2007) or the productive lexicon. How do we account for this?

Recall that in our framework, signed language clouds and gesture clouds have fuzzy boundaries and are defined by prototypical and not by criterial features. Because of this, they can overlap and merge in various ways. Our conception of the relation between language and gesture is thus analogous to the way Langacker views the relation between semantics and pragmatics. Langacker posits four possible positions on the relation between semantics and pragmatics: 1. separate components; 2. the non-existence of one or the other component; 3. non-differentiation; and 4. gradation. Langacker rejects "the strictly dichotomous view …, with a fixed and definite boundary between two separate components. … The claim … is that semantics and pragmatics form a gradation … with no sharp boundary between the two. But toward either extreme of the scale lie phenomena that are indisputably either semantic or pragmatic" (LANGACKER, 2008, p. 40).

These four positions have all been assumed at various times to describe the relation between signed language and gesture. Questions about whether this or that feature of an utterance is "language or gesture" assume the first position, that they are separate components. Some scholars have even questioned whe-

ther signers gesture, thus suggesting the non-existence of gesture (EMMOREY, 1999). We know that during certain points in the history of deaf education, signed languages were regarded merely as gesture. During the Milan conference in 1880, when advocates of signed language fought against and ultimately lost to the proponents of speech-only oral education, Marius Magnat, an oralist, took a non-existence position on the linguistic status of signed language, pronouncing that, "sign cannot convey number, gender, person, time, nouns, verbs, adverbs, adjectives" (LANE, 1984, p. 388). Giulio Tarra, summed up the non-differentiation position, categorizing sign as nothing more than gesture: "Gesture is not the true language of man which suits the dignity of his nature. Gesture, instead of addressing the mind, addresses the imagination and the senses. Moreover, it is not and never will be the language of society" (LANE, 1984, p. 391, 393).

We adopt the fourth position, gradation, and view the relation between language and gesture as a continuum, with certain phenomena prototypically either language or gesture. Not only this, but as we saw in section "Gesture and sign, and gesture to sign", gesture becomes incorporated into signed languages in several ways, further blurring the boundary between the two systems (Figure 3).

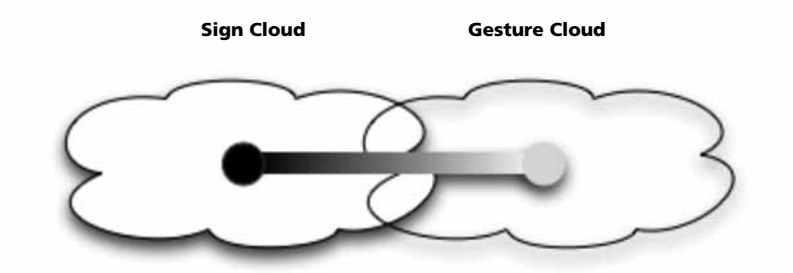

Figure 3 – Sign and cloud continua

Instantiations of the constructed action grammatical schema or scenario may fall at any point along the continuum from language to gesture. As evidence, we offer a preliminary analysis of three examples from signed retellings of the Pear Story. Two examples are signed in Libras, and one in Chinese Sign Language (CSL). Two deaf female fluent signers of Libras saw the Pear Story video and then were filmed retelling the story to another deaf person who hadn't seen the video. They were filmed without the presence of any hearing person. One deaf male fluent signer of CSL watched the Pear Story and retold the story; a hearing, fluent signer of CSL was present while the video was made. We have selected the pear-picking event for analysis. In this event, a man is in a tree picking pears (see Chart 1).

We suggest that while all three signers use constructed actions or lexicalized constructed actions to describe the pear-picking event, their expressions vary along the language-gesture continuum. Signer R and signer L use handshapes which more realistically depict the shape of the pear. Signer S uses a more stylized handshape. Signers R and L also use hand arrangements that are more depictive. Signer R alternates her hands to show the man picking pears from different branches of the tree. Signer L uses his non-dominant hand to hold the branches steady while his dominant hand removes the pears. Both of these sig-

ners use repeated picking movements, in different locations, with signing space expanded to depict the height of the branches above the man's head, again more realistically depicting the actual action of the man picking pears. They also use manner of movement to realistically depict the resistance of the pear as it is removed from the branch.

Chart 1 – Pear-picking Event

	Libras Signer S	**Libras Signer R**	**CSL Signer L**
Handshape	More stylized	More depictive	More depictive
Hand Arrangement	No alternation of hands	Alternation of hands depicting real event of pear picking	Dominant hand is picking pears Non-dominant hand holds the branch
Manner of Movement	Movement is stylized (default manner of movement)	Manner of movement shows the physical resistance of the pear being removed from the branch	Manner of movement shows the physical resistance of the pear being removed from the branch
Repetition	4 times with decreasing height of movement path, in same location	6 times, in different locations	8 times, in different locations
Face	Stylized/ grammaticized	Depicts effort involved on the part of the pear picker	Neutral
Signing Space	Default	Enlarged, picking occurs high in signing space	Enlarged, picking occurs high in signing space
Eye Gaze	Predominantly at interlocutor	Predominantly not at interlocutor	Predominantly not at interlocutor

Signer S does not alternate her hands; her dominant hand movement is highly stylized, moving from a position near the top of neutral signing space downward to a position where her non-dominant hand receives the pears. The movement path decreases in its extent across the four repetitions. The impression here is one of "on-the-spot" grammaticalization of the movement. She also uses a default, stylized manner of movement that does not evoke the actual force dynamics of the event, and the least number of repetitions.

While signer L uses a fairly neutral face throughout, signer R's face depicts the effort the man expends in removing the pears from the branches. She also uses what might be seen as an "embedded constructed action": she uses her non-dominant hand to fan herself, also sticking out her tongue, showing that the man was hot and had worked up a sweat. We note, however, that this gestu-

re is also the Libras sign for HOT or SUMMER; thus, this is may be an example of the lexicalization of a gesture. Signer S produces a very subtle, grammaticalized facial gesture, a very slight puff of air, that coincides with the manual path movement; these facial gestures also decrease in intensity across the four repetitions of the path movement. Thus, while it may relate to the event, it does not depict the man's face.

In constructed action, eye gaze is prototypically shifted. Signer S predominantly looks at the interlocutor. Sometimes, however, she shifts her eye gaze away from the interlocutor. These shifts occur both during what we are calling lexicalized constructed actions and at other times as well, such as when she is producing polycomponential signs and normal discourse. Signer R and signer L predominantly do not look at the interlocutor. Especially during construction actions, their eye gaze is shifted for long periods of time.

We should point out that eye gaze shifts not only in constructed action but in many other circumstances as well. For example, eye gaze shift occurs in constructed dialog when no constructed action is present (JOHNSTON; SCHEMBRI, 2007). Even signer S, who predominantly looks at the interlocutor, shifts her eye gaze throughout the pear-picking event. For example, when she uses a polycomponential sign to depict the man climbing a ladder into the pear tree, she looks at her hands while he is climbing. Later, when the man starts climbing down, she looks at her hands depicting the beginning of the climb down, then at the interlocutor at the mid-point of the descent, and again at her hands at the end.

Eye gaze is clearly an important and complex factor in constructed action, and one which deserves further research. In dynamic systems terms, we would call eye gaze a strong attractor. It is not, however, a criterial factor in determining what is and is not constructed action.

Delexicalization

One final issue is the matter of delexicalization. Following the work of others, Johnston and Schembri (2007) distinguish between the core native lexicon and the non-core native lexicon (or productive lexicon). They note that

> *[...] the core native lexicon consists of those completely and incompletely specified lexicalised forms which are frequently used and highly standardised in the language, while the non-core native lexicon is made up of meaningful units which are only partly specified* (JOHNSTON; SCHEMBRI, 2007, p. 163).

Clearly, then, one feature that distinguishes core from non-core or productive lexicon is degree of schematicity: the productive lexicon is more schematically represented.

They further suggest that in delexicalization "the components of a sign are modified to depict characteristics of the referent". Their example is modifying the extent of the movement of a size-and-shape specifier to indicate the extent of the size. They describe other ways in which lexicalized signs can be modified (JOHNSTON; SCHEMBRI, 2007, p. 164):

> *Auslan, like other signed languages, has a wide range of such meaningful units in the non-core native component of the lexicon: meaningful uses of handshape, orientation, location and movement, as well as a variety of nonmanual signals, are available in the mental lexicon of the fluent signer. These units can be used by the signer to extend or modify the meaning of lexicalised signs, as we have seen with the use of space in indicating verbs.*

A concrete example is the embedded constructed action described above, in which signer R depicted the man fanning himself. One interpretation of this expression is that a lexicalized sign has been delexicalized in the way described by Johnston and Schembri. Another interpretation is that signer R did not produce a delexicalized sign, but inserted a gesture into her stream of signs. Our position is that making such a determination can only arbitrarily be made. Like Janus, expressions like this have two faces: they are both lexicalized gestures, and gestures that are linguistically structured (COSTELLO; FERNÁNDEZ, LANDA, 2006).

We return to our original question regarding constructed action: is it language or gesture? We would argue that in the realm of language and gesture – and probably in a great many more scientific domains as well – such either/or questions are ill-formed, based on erroneous assumptions about the phenomena under investigation. In the present case, we stand by our answer. No matter where constructed action expressions fall along a language-gesture continuum, no matter where they reside in the ever-dynamic interface of sign language clouds and gesture clouds, they are sanctioned by the constructed action scenario, which is itself a complex but nevertheless conventional linguistic schema. Thus, in our view, constructed action expressions are always linguistic, a part of the structured inventory of conventional linguistic units, the grammar, invoked by signers when they construct expressions.

Conclusion

We have offered the metaphor of clouds of human expression. We have labeled and discussed these as language clouds, sign language clouds, gesture clouds, and so forth. The cloud metaphor was chosen to focus attention on important aspects of language and gesture. Language and gesture are dynamic, emergent systems, the product of a human expressive ability that is grounded in embodied cognitive abilities. We might extend the metaphor once more and note that clouds form from moisture that often travels up from the ground, and then that moisture rains back down to earth. Expressive units of language and gesture are also interactive in the vertical dimension. Expressive units – we use the term to include both linguistic and gestural units – form out of actual usage events; these units then sanction new usage events. Usage events are modality-rich. The units formed are of various types and differ along a range of characteristics, including more specific and more schematic, more or less symbolically complex, more or less entrenched in any one individual's system of expressive clouds, and more or less shared across the expressive clouds of other individuals. Because they are built from the bottom up on the basis of modality-rich usage events, expressive units are hybrid entities, aggregates of any and all aspects of the multimodal usage event that gave rise to them.

There is no doubt that specific language systems emerge – spoken and signed languages, English, Portuguese, Brazilian Sign Language, Chinese Sign Language, and so forth. There is also no doubt that these systems have fuzzy boundaries and are as ephemeral as real clouds, that they are constantly changing, and that each individual has different systems. The implications of such a dynamic, embodied view are profound. We will note only one final ramification: determining whether a unit is language or gesture may ultimately be futile. It will depend on the expressive unit's occurrence in a particular usage event, by a par-

ticular individual, who is communicating in a particular context, who knows and uses particular languages. Even then, because these systems are in constant interaction and undergo constant change, what is gesture today may be tomorrow's language.

References

ALDRETE, G. S. *Gestures and acclamations in ancient Rome*. Baltimore: Johns Hopkins University Press, 1999.

AZIZ-ZADEH, L.; DAMASIO, A. Embodied semantics for actions: findings from functional brain imaging. *J Physiol Paris*, v. 102, n. 1-3, p. 35-39, 2008.

BERTHOZ, A. *The brain's sense of movement*. Cambridge, MA: Harvard University Press, 2000.

BYBEE, J. L. *Language, usage and cognition*. Cambridge; New York: Cambridge University Press, 2010.

CHOMSKY, N. Three factors in language design. *Linguistic inquiry*, v. 36, n. 1, p. 1-22, 2005.

COSTELLO, B.; FERNÁNDEZ, J.; LANDA, A. The non-(existent) native signer: sign language research in a small Deaf population. In: PROCEEDINGS FROM TISLR9: forty five papers and three posters from the 9th Theoretical Issues in Sign Language Research Conference. Florianopolis. 2006

DAMASIO, A. R. *Descartes' error*. New York: Putnam, 1994.

DAMASIO, A. R. *Self comes to mind*: constructing the conscious brain. New York: Pantheon, 2010.

EDELMAN, G. M. *Neural Darwinism*: the theory of neuronal group selection. New York: Basic Books, 1987.

EDELMAN, G. M. *The remembered present*: a biological theory of consciousness. New York: Basic Books, 1989.

EMMOREY, K. Do signers gesture? In: MESSING, L. S.; CAMPBELL, R. (Ed.). *Gesture, speech, and sign*. Oxford: Oxford University Press, 1999. p. 133-159.

ENFIELD, N. J. On linear segmentation and combinatorics in co-speech gesture: A symmetry-dominance construction in Lao fish trap descriptions. *Semiotica*, v. 149, p. 57-124, 2004.

FERRARA, L. *The grammar of depiction*: exploring gesture and language in Australian Sign Language (Auslan). 2012. Thesis (Doctoral in Linguistics)–Macquarie University, Sydney, 2012.

HAY, J. B.; BAAYEN, R. H. Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Science*, v. 9, n. 7, p. 342-348, 2005.

JANZEN, T. Lexicalization and grammaticalization. In: STEINBACH, M. et al. (Ed.). *Handbook on Sign Languages*. Berlin; New York: Mouton de Gruyter, 2012.

JANZEN, T.; SHAFFER, B. Gesture as the substrate in the process of ASL grammaticization. In: MEIER, R.; QUINTO, D.; CORMIER, K. (Ed.). *Modality and structure in signed and spoken languages*. Cambridge: Cambridge University Press, 2002. p. 199-223.

JOHNSTON, T.; SCHEMBRI, A. *Australian Sign Language (Auslan)*: an introduction to sign language linguistics. Cambridge: Cambridge University Press, 2007.

KENDON, A. Gesture. *Annual Review of Anthropology*, v. 26, p. 109-128, 1997.

KENDON, A. *Gesture*: visible action as utterance. Cambridge; New York: Cambridge University Press, 2004.

KLIMA, E.; BELLUGI, U. *The signs of language*. Cambridge, MA: Harvard University Press, 1979.

KUNENE, D. P. Speaking the act. In: VOELTZ, E. F. K.; KILIAN-HATZ, C. (Ed.). *Ideophones*. Philadelphia: John Benjamins Publishing Company, 2001. p. 183-191.

LADEWIG, S. H. Putting the cyclic gesture on a cognitive basis. *CogniTextes. Revue de l'Association Française de Linguistique Cognitive*, v. 6, n. 23, 2011.

LANE, H. *When the mind hears*: a history of the deaf. New York: Random House, 1984.

LANGACKER, R. W. *Foundations of cognitive grammar*: theoretical foundations. Stanford: Stanford University Press, 1987. v. I.

LANGACKER, R. W. *Foundations of cognitive grammar*: descriptive application. Stanford: Stanford University Press, 1991. v. II.

LANGACKER, R. W. *Cognitive grammar*: a basic introduction. Oxford: Oxford University Press, 2008.

LIDDELL, S. K. *Grammar, gesture, and meaning in American Sign Language*. New York: Cambridge University Press, 2003.

LLINÁS, R. *I of the vortex*: From neurons to self. Cambridge, MA: MIT Press, 2001.

McCLEARLY, L.; VIOTTI, E. Sign-Gesture Symbiosis in Brazilian Sign Language Narrative. *Meaning, Form, and Body*, p. 181-201, 2010.

McNEILL, D. *Hand and mind*: what gestures reveal about thought. Chicago: University of Chicago Press, 1992.

OKRENT, A. A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics (or at least give us some criteria to work with). In: MEIER, R.; CORMIER, K.; QUINTO-POZOS, D. (Ed.). *Modality and structure in signed and spoken languages*. Cambridge: Cambridge University Press, 2002. p. 175-198.

PFAU, R.; STEINBACH, M. *Modality independent and modality specific aspects of grammaticalization in sign languages* (Linguistics in Potsdam 24). Postdam: Universitätsverlag Potsdam, 2006.

QUINTO-POZOS, D. Why does constructed action seem obligatory? An analysis of classifiers and the lack of articulator-referent correspondence. *Sign Language Studies*, v. 7, n. 4, p. 458-506, 2007.

SCHEMBRI, A.; JONES, C.; BURNHAM, D. Comparing action gestures and classifier verbs of motion: evidence from Australian Sign Language, Taiwan Sign Language, and nonsigners' gestures without speech. *J Deaf Stud Deaf Educ*, v. 10, n. 3, p. 272-290, 2005.

SHAFFER, B.; JARQUE, M. J.; WILCOX, S. The expression of modality: conversational data from two signed languages. In: NOGUEIRA, M. T.; LOPES, M. F. V. (Ed.). *Modo e modalidade*: gramática, discurso e interação. Fortaleza: Edições UFC, 2011. p. 11-39.

SPIVEY, M. *The continuity of mind.* Oxford; New York: Oxford University Press, 2007.

THELEN, E.; SMITH, L. B. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: The MIT Press, 1994.

VERMEERBERGEN, M.; DEMEY, E. Sign+ gesture= speech+ gesture? In: VERMEERBERGEN, M.; LEESON, L.; CRASBORN, O. (Ed.). *Simultaneity in signed languages*: form and function. Philadelphia: John Benjamins Publishing Company, 2007. p. 257-282.

VOELTZ, E. F. K.; KILIAN-HATZ, C. *Ideophones*. Amsterdam; Philadelphia: John Benjamins, 2001. (Col. Typological Studies in Language, v. 44).

WEBB, R. A. *Linguistic features of metaphoric gestures*. University of Rochester. Department of Brain and Cognitive Sciences, 1996.

WILCOX, S. Gesture and language: Cross-linguistic and historical data from signed languages. *Gesture*, v. 4, n. 1, p. 43-75, 2004.

WILCOX, S. Routes from gesture to language. *Revista da ABRALIN – Associação Brasileira de Linguística*, v. 4, n. 1-2, p. 11-45, 2005.

WILCOX, S. Routes from gesture to language. In: PIZZUTO ANTINORI, E.; PIETRANDREA, P.; SIMONE, R. (Ed.). *Verbal and signed languages*: comparing structures, constructs and methodologies. Berlin: Mouton de Gruyter, 2007. p. 107-131.

WILCOX, S.; ROSSINI, P.; ANTINORO PIZZUTO, E. Grammaticalization in sign languages. In: BRENTARI, D. (Ed.). *Sign languages*. Cambridge: Cambridge University Press, 2010. p. 332-354.

WILCOX, S.; WILCOX, P. The gestural expression of modality in American Sign Language. In: BYBEE, J.; FLEISCHMAN, S. (Ed.). *Modality in grammar and discourse*. Amsterdam: John Benjamins Publishing Company, 1995. p. 135-162.

WILCOX, S.; XAVIER, A. N. Um quadro teórico para um tratamento unificado das línguas faladas, línguas sinalizadas e gestos. *Todas as Letras*, São Paulo, v. 15, n. 1, p. 88-110, 2013.

Resumo: *Neste artigo apresentamos os primeiros passos para o desenvolvimento de um quadro teórico capaz de tratar unificadamente as línguas faladas, as línguas sinalizadas e os gestos. Esse quadro teórico se baseia em três teorias já existentes: a teoria dos sistemas dinâmicos, a gramática cognitiva e abordagens não cartesianas à neurociência. Consideramos a existência de uma habilidade expressiva humana, que argumentamos ser baseada na necessidade de seres que se movem compreender seu ambiente e resultante de um sistema conceitual corporeado em sistemas motores e perceptivos. Essa habilidade expressiva humana subjaz a língua e os gestos. Exploramos, neste trabalho, as relações entre língua e gesto como sistema emergentes.*

Palavras-chave: *línguas de sinais; gestos; gramática cognitiva.*

Recebido em fevereiro de 2013.
Aprovado em fevereiro de 2013.