Hrvatska revija za rehabilitacijska istraživanja 2022, Vol 58, (Special Issue) Sign Language, Deaf Culture, and Bilingual Education str. 121-142

GESTURES AND ADAPTIVE NICHES: AN EVOLUTIONARY PERSPECTIVE ON CO-SPEECH GESTURES

DONOVAN RICHARD GROSE
Hong Kong SAR, China
contact: donovangrose@me.com

Received: 25.10.2021.
Accepted: 13.09.2022.

PRELIMINARY REPORT
UDK: 81'221
doi: https://doi.org/10.31299/hrri.58.si.7

Abstract: *This proposal presents an evolutionary analysis of three types of co-speech gestures: symbolic emblems, indexical pointing gestures and iconic representational gesticulations. Synthesizing insights from a range of published sources in gestural studies, general linguistics and sign language linguistics, primate studies and analyses of biological evolution, these gestures are analyzed as evolved traits adapted to particular niches or roles within broader systems. Niche boundaries are comprised of an element's distinct properties and functions, routes of learning and transmission and degrees of innateness and evolvability within populations. Rather than elements distributed along a flat productive-analytical continuum or as stages along diachronic pathways, these gestural traits are analyzed in terms of adaptive peaks and valleys with a landscape representing the broader system comprising human gesture and language. The same evolutionary processes are used to analyze gestures in speaking populations and the linguistic traits derived from gestures in signing populations. This approach offers new ways of approaching proposed linguistic universals and long-standing issues such as listability in sign languages, while offering a formal approach to gestures.*

Keywords: *co-speech gesture, evolution, selection, sign language*

1. INTRODUCTION

Grammatical language is the most complex and powerful communication system to have ever evolved, and natural languages are only acquired and used by one species, *Homo sapiens*. But human gesture is also unique to our species. Although gesture is often viewed as simpler and less expressive than grammatical language, the two are used together in the same utterances, develop together and they have co-evolved. The sorts of gestures discussed here are voluntary intentional communicative movements and configurations, primarily of the hands and arms, that tend to accompany the grammatical language of the speech stream in composite utterances. These gestures are often collectively referred to as co-speech gestures, but many of the same behaviors can be used without speech and the same kinds of gestures occur in sign languages (Goldin-Meadow & Brentari, 2017). These are: emblems which have stable conventionalized forms and meanings, like the *thumbs up* gesture, and can be used in the absence of speech; pointing gestures or points that use a body part to project a vector towards a referent to which they refer; and representational gestures or gesticulations that depict their meanings with highly variable and flexible context dependent forms, including representations of an entity's size and shape, and motion through space (Özyürek, 2012; Kendon, 2004). These gestures are functional but generally optional and infamously idiosyncractic, and so do not always yield themselves easily to description and analysis using linguistic tools and techniques. Their forms and meanings are holistic and cannot be decomposed into units of form like phonemes or units of meaning like morphemes, and they are not bound by syntactic constraints. They make contributions to composite utterances, but in a sense are 'outside' of grammar and so have generally been set aside in phonological and syntactic analyses. Recently attitudes towards gestures have shifted, but relative to grammatical language, gesture is much less well-understood.

This proposal contributes to gestural analyses using an framework of cultural evolution, adapted

from Jablonka and Lamb (2014). Emblems, gesticulations and points are analyzed as evolved traits that are adapted to three particular roles or niches within the broader system that comprises human language and gesture. A niche is distinct from a grammatical category in that they are defined by their functions and interactions within the broader system, but their boundaries are gradient and an individual gestural trait may be more or less adaptive to a particular niche. These niches can be associated with distinct routes of development and learning as well as different evolutionary histories and degrees of evolvability within a population. Gestural and linguistic evolution is conceptualized here as the movement and location over time of a trait through a metaphorical landscape composed of adaptive peaks and less adaptive valleys. Notions such as productivity, conventionalization and grammaticalization that in linguistic analyses are often conceptualized in terms of continuums or stages along diachronic pathways are understood here products of ongoing evolutionary processes pushing traits up an adaptive peak and maintaining them there. The same evolutionary processes apply to both gestural and linguistic traits and occur in both speaking populations where these traits are channeled through different modalities and in signing populations where they are channeled through one.

This brief proposal begins with a discussion of gestural research with a focus on those areas that are relevant for an evolutionary analysis. Section 3 outlines the current framework of evolution in multiple dimensions, introducing general terminology from evolutionary theory that can usefully be applied to linguistic and gestural analyses. Section 4 discusses emblems, points and gesticulations with reference to data from studies of sign languages where relevant. Section 5 presents conclusions and discusses possible future applications of evolutionary approaches to gestural and linguistic research.

2. GESTURE VS. LANGUAGE

2.1 (Co-Speech) Gestures, in Brief

Gestures that occur together with speech, despite being meaningful and functional in context have generally been excluded from analyses of language, even when their contributions to a composite utterance might be relevant for the research questions at hand. It was often assumed that semantically gestures served as visual paraphrases of meanings already expressed within the speech stream, but Kendon (2004) demonstrates that this assumption is incorrect in two ways. Gestures not only express visual information, but they also express meanings visually in ways that grammatical language does not. This can be illustrated with the notions of path and manner. Any meaning may be lexicalized in theory, but languages do not lexicalize path and manner meanings together in individual verbs (Talmy, 2000). Since these meanings can be combined within a single predicate, this does not constrain expressive power.

Verb-framing langauges prefer to lexicalize paths in verbs; satellite-framing languages tend to lexicalize manners in verbs. These two strategies are illustrated in English (otherwise a satellite-framing langauge) with: (a) *They ascended the hill by car;* and (b) *They drove up the hill*. Both predicates express a path (*ascended/up*) and a manner (*by car/drove*). Gesticulations are not conventionalized or lexicalized units and may depict a manner, a path or both simultaneously. As a consequence, a composite utterance with both speech and gesticulations may express both path and manner even when only one of the two is expressed in the speech stream (*they ascended the hill/they went up the hill*). These gestural contributions are missed in linguistic analyses based solely on spoken or written utterances.

Kendon (2004; 2000) argues that gesture and language should be treated as parts of the same broader system within a theoretical framework that encompasses both. This raises questions. If Gesture as a system is considered distinct from grammatical Language, there are issues regarding the nature of their interfaces and how these two systems evolve and develop in order to operate together. However, if Language is conceptualized more broadly to include both Gesture and Grammar, the same sorts of issues arise. Developing semantic analyses that take gesture into account is one issue, but it is not clear how or if the same should be done for syntactic analyses. Gestures

do not have grammatical properties, so grafting something like Gesture Phrases onto syntactic trees is not a systematic solution.

McNeill (2016; 2012; 2005) argues, from a distinct phenomenological perspective, for an approach in which gesture and speech (or more generally grammatical language) are unified semantically rather than syntactically. Gesture and speech are treated as parallel outgrowths in semiotic opposition that emerge simultaneously through a Growth Point from a shared minimal idea unit. Each stream is specialized for expressing facets of meaning that the other is not, but they are unified by a shared mental representation. The productivity and optionality of co-speech gestures arise from speakers' evaluation of the discourse context as they determine if and when to gesture and what forms any gestures will take. The speech stream influences the gestural stream and vice versa as they expand from the Growth Point, allowing for utterances containing only speech and those containing only gesture.

McNeill (2012) argues that any model of the evolution of language must account for the unity of thought, gesture and speech (again grammatical language) in modern human cognition and communication. He further argues that because they are in semiotic opposition, gesture and speech must have originated together and neither could evolve without the other. There is much to take onboard from McNeill (2016; 2012; 2005)'s analyses but claims from a phenomenological approach must first be translated into an evolutionary framework for the current proposal. For example, Gesture and Language, as abstract entities, do not represent units of evolution, nor do they evolve in a technical sense.

2.2 Gestures: Development vs. Evolution

Human language and gesture leave no direct evidence in the archeological record, so there are some questions that will never be answerable with certainty. Without direct evidence the most plausible models of language and gesture evolution make inferences from a wide range of sources, including our closest living relatives, development in modern human infants and the emergence of *de novo* sign languages. What these strands of evidence indicate is that some aspects of language and gesture are largely innate and so must have evolved biologically. Other aspects are much more variable and dependent upon the environment, exposure and learning. Deaf individuals without sufficient access to the spoken languages in their developmental environments and without exposure to a sign language tend to develop elaborated gestural home-sign systems for communication with family members, who tend not to use these systems themselves. These systems lack those properties of a language that must be acquired from within a linguistic environment, or fragile properties, but yet exhibit other resilient properties of language. These resilient properties appear without appropriate language input and include the use of stable form/meaning pairs in regular patterns (Goldin-Meadow, 2005). These resilient properties in home-sign systems have been treated as proto-syntax and proto-morphology but with a population of only one, these systems cannot evolve. These findings show that modern humans have evolved some innate capacities to acquire and use language and gesture, and also that modern natural languages are products of cultural evolution within populations building on these biologically evolved capacities. When groups of previously isolated home-signers come together to form new populations, language evolution proceeds extremely rapidly, as documented in the *de novo* emergence of Nicaraguan Sign Language over less than a decade (Sengas, Özyürek, & Goldin-Meadow, 2013; Arbib, 2012; Senghas, Senghas, & Pyers, 2005; Senghas, 2003).

Home-sign systems show what is possible with a modern 'language ready' brain without access to linguistic input in a population. *De novo* sign languages show what is possible with a population in which linguistic evolution can take place. Studies of our closest living relatives the chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*), provide the best available models of what is possible without any of the human biological adaptations relevant for gesture and language. These species are social and extremely intelligent, but our lin-

eage diverged from theirs between 7 and 5 million years ago (Seddon, 2020), and so they lack any of the innate capacities for language and gesture that evolved in our lineage after that split. For example, typically developing human infants begin pointing imperatively and informatively around 10 months of age (Tomasello & Call, 2014; Goldin-Meadow, 2005). Wild chimps do not produce human-like points, but captive chimps can learn to use points imperatively with humans, but do not appear to be able to use points purely for informative functions (Francis, 2015; Tomasello & Call, 2014). Without the anatomy for human speech, several chimps including Nim Chimpsky (1973-2000) and Washoe (1965-2007) and a lowland gorilla, Koko (1971-2018) were trained to communicate using signs from American Sign Language (ASL). These individuals were able to learn hundreds of units but were not able to acquire a grammar, either of ASL or of spoken English. This suggests that the units they actually learned lacked grammatical properties and were similar to emblems rather than to the linguistic signs in their input. On this view, the capacities of sign-trained apes are comparable to those of Kanzi, a bonobo who has learned to communicate by touching hundreds of lexigrams symbols on a board or computer screen (Francis, 2015; Tomasello & Call, 2014; Tomasello, 2008). These individual apes did not learn these units in the ways that human children learn lexical items and they did not pick them up as rapidly, but these results suggest that at least under the unnatural conditions in which they were trained, some apes are able to stretch their innate cognitive capacities to encompass many symbolic units. Their achievements are consistent with their relatively smaller brain sizes, about a third that of modern humans, and the evolutionary relationships between our species and theirs. They offer hints at what fragile traits of human language and gesture may be acquired without any of the human-specific adaptations for doing so.

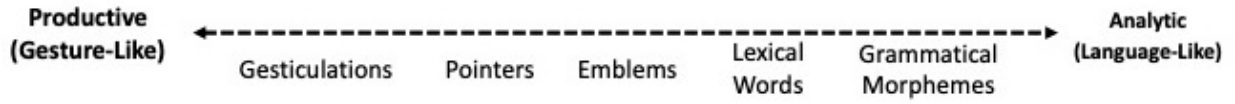
Evolution occurs at the level of populations. Development refers to the changes that an individual undergoes over their lifetime, but Goldin-Meadow (2012; 2005; 2003)'s work on early childhood gestures and language acquisition is relevant for evolutionary analyses. Developmental pathways are also products of evolutionary change. It has been argued that most evolutionary changes in multi-cellular organisms are due to changes in developmental pathways rather than mutations in single genes (Francis, 2015). It may be the case that most of the important evolutionary changes relevant for human language and gesture arose from changes in how and when our ancestors acquired these systems, rather than to changes within these systems themselves. If this is the case, then the relative sequence in which traits emerge in development may not reflect the sequence in which they evolved in our ancestral species.

This is perhaps most important regarding points. Chimpanzees are limited in their capacity to use human-like points, but Goldin-Meadow (2012; 2005; 2003) shows that early gestures scaffold and facilitate language acquisition with points appearing before first words, allowing infants to develop shared attention with caregivers. Later, points allow for rapidly expanding lexical inventories as children point to ask for names and lables for unfamiliar referents. These studies indicate not only that humans have evolved the capacity to point but also the developmental pathways in which points appear very early. In other words, points seem simple and easy to us because we have evolved adaptations that make them so. They appear first in development for modern humans, but they did not necessarily evolve before emblems and gesticulations. Whenever the capacities for points evolved, later changes in developmental pathways pushed the emergence of points very early in development, facilitating earlier language acquisition.

2.3 Gestural and Linguistic Elements: Continuums and Pathways

At a word or gesture level of granularity, in mature systems gestural and linguistic elements appear to vary along continuums from the most productive and gesture-like to the most analytic, grammaticalized and language-like (Langacker, 1999; and many others):

Figure 1. *A Continuum of Language and Gestural Elements*

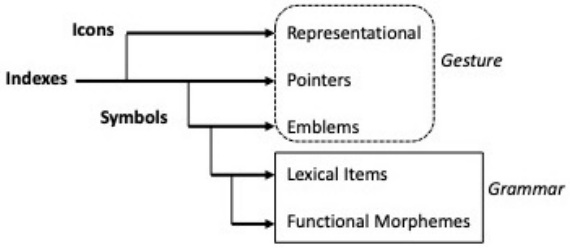


In systems comprising a spoken language and gesture, language-like elements towards the right are channeled through the aural/oral modality and gesture-like elements towards the left are channeled through the visual/motor or visual/spatial modality. In systems comprising a sign language and gesture, all elements are channeled through the same modality. If modern systems have evolved to contain both grammatical language and gesture (McNeill, 2012), and if humans have evolved to acquire and use systems with the properties of modern language and gesture, this would seem to create challenges for signing populations when language and gesture, specialized to express distinct sorts of meanings, must compete for space within the single modality. The solutions that the overwhelming majority of sign languages have arrived at, independently in completely unrelated languages and groups, include linguistic signs that appear much more gesture-like and further to the left in Fig. 1 that might be expected compared to the situations in speaking populations. These include the personal pronominal signs and classifier predicates and the linguistic use of space (Cormier, 2014; 2012; Liddell, 2003: Schembri, 2003; and many others) discussed below. The nature of these productive signs is debated, but from the current evolutionary perspective, we can assume that the distributions of elements along the continuum represented in Fig. 1, are products of the same evolutionary processes operating on different kinds of variation in speaking and signing populations, resulting in distinct but functionally equivalent systems.

Synchronic analysis utilizing continuums like that Fig. 1 are attractive in part because they seem to reflect the results of diachronic processes. These processes arise from the creative and productive use of language and gesture and are cumulative and directional (Heine & Kuteva, 2012; Brinton & Traugott, 2006; Pfau & Steinbach, 2006; Bybee, Perkins, & Pagliuca, 1994). Changes within an element are likely to push it rightwards with increases in degrees of conventionalization, lexicalization or grammaticalization. Functional morphemes may be derived from lexical words, as in the derivation of the aspectual *have* in English from the verbal *have,* but the reverse either does not occur or is very rare. New emblems may be derived from more productive gestures and gestures may become more regularized in particular contexts, but in speaking populations they are not lexicalized or grammaticalized. When speech is absent, gestures can be pulled into linguistic roles, for example in monastic communities that use manual communication systems (Pfau, 2012). Signing populations demonstrate that in the right conditions it is possible to derive both lexical signs and functional elements from gesture (Pfau & Steinbach, 2006; Wilcox, 2004). Yet, even with systems in which gestures can be grammaticalized, signers still gesture (Goldin-Meadow & Brentari, 2017).

These facts are consistent with claims that modern gestures and grammatical language operate together within the same system, even if grammatical language evolved from largely gestural precursor systems deep in our evolutionary history (Arbib, 2012; Corballis, 2009; and others). In broad strokes, the modern system is assumed to have evolved along a general semiotic pathway with indexes emerging first, likely only points, followed by icons, including early gesticulations, and then symbols, including holistic proto-signs and proto-words (Arbib, 2012) similar to modern emblems. Symbols may be iconic or indexical as well as arbitrary, and symbolic thinking would have allowed for a huge leap in the expressive power of these systems. Whether they emerged relatively gradually or abruptly, grammatical systems later evolved to manipulate and productively recombine symbols:

Figure 2. *A Semiotic Pathway*



This figure, adapted from Arbib (2012)'s Expanding Spiral Model, illustrates the changes that unfolded perhaps over hundreds of thousands of years through which gestural and linguistic systems were progressively expanded and elaborated in the lineage that produced *Homo sapiens*. These are changes on a much longer scale from the diachronic changes that occur within modern languages and were likely to have been associated with biological changes in the size and organization of pre-modern human brains. This pathway, as it is represented in Fig. 2 assumes icons emerged in a system that already contained indexes, without assuming that icons are derived from indexes or symbols from icons. More recently evolved traits would not necessarily replace older existing traits. Instead, the emergence of new types of traits altered the interactions within the system so that gestural and linguistic traits co-evolved.

There are a number of proposals arguing that gestures evolved before speech and that the earliest grammatical languages were signed rather than spoken (Armstrong & Wilcox, 2007; Armstrong, Stokoe, & Wilcox, 1995). The *de novo* emergence of new sign languages demonstrates that full grammatical languages can evolve entirely from existing gestural systems in the absence of speech (Sengas, Özyürek, & Goldin-Meadow, 2013; Arbib, 2012), but these languages have all emerged in populations of modern humans with fully modern brains. The evolution of grammatical language itself may have been quite different and certainly was not as abrupt as the emergence of documented *de novo* sign languages. Since any account of the evolution of modern languages has to address the evolution of grammatical systems regardless of modality, the evolution of speech and the evolution of modern gestural systems, sign-first models face some important questions. For example, it is not clear why speech should evolve relatively late in the presence of an existing grammatical sign language followed by a modality shift from grammatical language in one modality to another. Scenarios involving the co-evolution of traits in both modalities throughout the emergence of human gesture and langauge do not require a massive modality shift, while still allowing for earlier systems relying more heavily on gesture and signs than modern spoken populations. Whatever the case, a clear sequence of steps towards fully modern systems is unlikely. The indirect archeological evidence suggests long periods of cultural stability within our archaic ancestors over tens of thousands to millions of years, punctuated by abrupt changes often associated with drastic climatic swings during the Pleistocene epoch (2.8 million to 11,400 years ago).

These shifts in climate would have imposed intense evolutionary pressures on the evolution of human brains, cultures and communication systems (Seddon, 2020). Each pre-modern and archaic communication system used by our ancestors would have been functional and adaptive for the people who used it, but but at this distance we cannot know for certain what any of those systems were like. With the exception of modern *Homo sapiens*, all other human species and their gestural and linguistic systems disappeared long ago.

3. GESTURE, LANGUAGE AND EVOLUTION

3.1 An Evolutionary Framework for Gesture and Language

Jablonka and Lamb (2014) argue that evolution is a general process that occurs in any system involving the differential success of heritable variation. Evolution is most familiar in biological systems in which genes are the units of inheritance that are copied and transmitted from parents to offspring. Genes are reshuffled during sexual reproduction and copying errors (mutations) can occur, producing variation within a biological population, defined as the summation of all of the members of a group or species living in a region and reproducing

(Mayr, 2001). As an individual genotype interacts with its environment it produces a unique phenotype, the total of all of an organism's physical and behavioral traits (Francis, 2015; Mayr, 2001). Phenotypic traits may vary in their relative adaptiveness to a particular environment, and because they interact with the environment directly they are the units that selection can target. Traits are adaptive only if they provide their owners with reproductive advantages that allow them to contribute relative more copies of their genes to the next generation than their competitors with less adaptive traits produced by different genotypes. This is differential success and over time it produces biological evolution, or changes in the frequencies of gene variants (alleles) within a population over time (Francis, 2015; Jablonka & Lamb, 2012; Mayr, 2001).

Physical phenotypic traits are not directly biologically heritable, but humans have evolved very large brains that allow for elaborate cultures and cultural evolution. In this dimension, some behavioral and cultural traits become functionally heritable because they are learned and copied as stable units that persist within a population (Jablonka & Lamb, 2014; Ritt, 2004). A cultural trait is not a physical entity like a gene, but rather a small-scale mental competence that can be reconstructed from observation and exposure and then expressed in future performances. These traits may be transmitted within a population in any direction, between parents and offspring, between peers and between more and less experienced individuals (Jablonka & Lamb, 2012). For linguistic and gestural traits to be transmitted, they must be reconstructed from observed performances. Reconstructions are often inexact copies, introducing variation within a population. Humans use gesture and language creatively and productively and in goal-directed ways. This is another source of variation, but this allows observers to evaluate the relative success of a performance. Traits reconstructed from successful performances are more likely to be used again in the future. Because traits are expressed at different rates will also differ in the rates at which they are learned, this produces differential success among acquired cultural traits. Traits will also vary in their relative fragility and resilience, to use Goldin-Meadow's terms. Traits with a heavier innate contribution require very little exposure to be learned, appear earlier in development and are more resistant to modifications based on learning. Traits with lighter genetic contributions require more exposure and learning, but are more malleable and evolvable within a population (Francis, 2015; Jablonka & Lamb, 2014).

For the current discussion, the relevant dimension of cultural evolution involves changes in the frequencies of linguistic and gestural traits within a population over time. The relevant sort of populations here are summations of all the individuals who communicate with each other using the same code, acquired by new members from within the population. Languages, gestural systems and their varieties are conceptualized as pools of all of these acquired traits contained within the population.

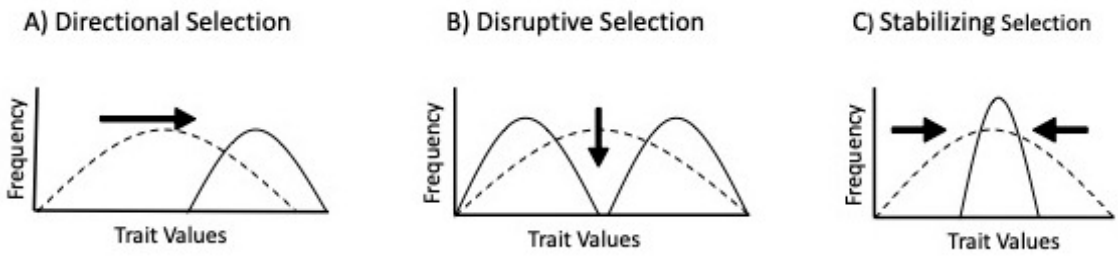
In any dimension, only populations evolve (Mayr, 2001). The changes in an individual over their lifetime, including changes in linguistic and gestural behavior represent development. Aggregations of these individual-level changes over time, and population turn-over produce evolutionary changes at the population level.

3.2 Traits and Selection Processes and Pressures

Evolutionary changes are products of only four basic mechanisms (Mayr, 2001). When adapted for analyses of linguistic/gestural evolution these are: (1) mutation, in which a new trait is introduced from inside a population, through inexact copying, modifications of existing traits based on learning or creative recombination such as coining a new word; (2) flow, when a trait is introduced from another population, as in lexical borrowing from one language into another; (3) drift, when trait frequencies shift by chance; and (4) selection, when trait frequencies change due to differing degrees of adaptiveness. All evolutionary changes are products of one or more of these basic mechanisms, but most linguistic/gestural evolution is attributable to mutation, flow and drift rather than selection.

Selection pressures arise from interactions between performances and the environments in

Figure 3. *Selection Processes*

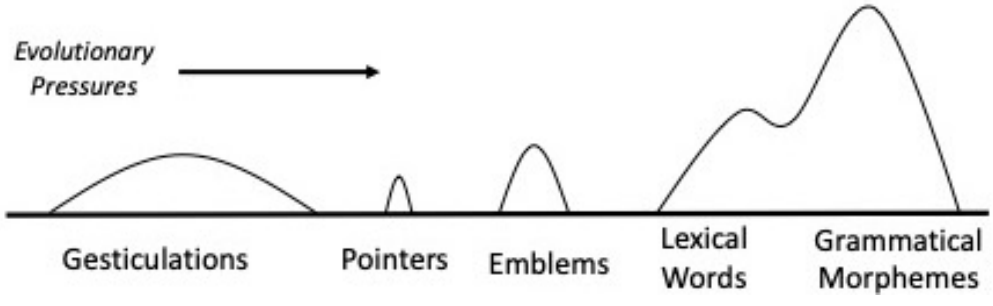


which those performances occur and can only target existing variation. These pressures are distinct from grammatical constraints or rules and are not parts of these systems themselves. Each trait is expressed across performances with a range of values, referring to variations in form, meaning and function. A positive selection pressure favors a variant, a negative pressure does the opposite. A pressure may target only one aspect of trait, for example form, but produce indirect impacts on function and meaning. In general, variants that are easier to learn will have advantages over those that are more difficult. Variants that are more easily produced or perceived will have advantages over those that are harder. Variants have help to constrain interpretations in some way will also tend to be favored over those that do not. When they occur, selection pressures produce changes through three processes: directional, disruptive, or stabilizing selection (Francis, 2015; Jablonka & Lamb, 2014; Mayr, 2001).

Directional selection (Fig. 3a) favors values at one extreme of a trait's range, pushing the mean value for the trait in a direction. Disruptive selection (Fig. 3b) involves selection for extreme values over intermediate ones, splitting an initial trait into two or more resulting traits. Stabilizing selection (Fig. 3c) involves selection against extreme or divergent values of a trait, reducing variation and maintaining the trait within a narrow range.

Despite their optionality and variability, the gestures discussed here fall into three groups, each with a set of properties that make it distinct from linguistic traits and from the other two groups of gestures. The current analysis argues that the distribution and behaviors of gestures are products of evolutionary processes that adapt these gestural traits to particular niches within the broader landscape comprising gestural and linguistic systems. Extending a notion from biological adaptation of organisms within ecosystems (Mayr, 2001), a niche is conceptualized as an adaptive peak, with distinct peaks separated from each other by less adaptive valleys. Evolutionary pressures push and pull traits up these peaks through directional and disruptive selection, and then tend to maintain them there through stabilizing selection against divergent variants. In the general illustration below, variation within a niche is represented with width, steepness represents the intensity of selection pressures and height represents evolvability, or possible changes within a population.

Figure 4. *Gestural and Linguistic Traits and Adaptive Peaks*



The peak to which gesticulations are adapted is wide and low, reflecting the broad variation among highly flexible traits, in a niche where variation is either selectively neutral or pressures are weak. Points are adapted to a narrower steeper peak, where selection pressures tolerate a narrower range of variation. The peak for emblems is steeper and narrower than that of gesticulations, reflecting strong selection pressures on these much more evolvable traits. Lexical and grammatical traits are represented together, with a lower peak for lexical content categories and a higher peak for grammatical traits and functional morphemes. Linguistic traits as a group are more evolvable than gestural traits. These peaks do not represent stages along a pathway. Evolutionary pressures are illustrated here as pushing traits rightwards and upwards against adaptive peaks.

In speaking populations lexical and grammatical peaks are populated with spoken linguistic traits. In signing populations, the same peaks are populated with signed linguistic traits. When no linguistic traits are available to fill an existing niche, selection pressures pull gestural traits already adapted to other niches rightwards and upwards to fill empty niches. In a new niche, traits undergo processes of adaptation and hill-climbing again as they are modified to fit into new roles. As we will see below, these processes are reflected in sign languages in the lexicalization from gestures of new lexical signs, the grammaticalization of points into pronominal roles and the evolution of gesticulations into classifier predicates.

4. GESTURES AS EVOLVED AND EVOLVING TRAITS

One way to help conceptualize the differences between gestural and linguistic traits within systems is to imagine two pairs of vast metaphorical spaces. One pair includes a space representing all possible word-forms, given the phonological inventory and constraints within a language, and another space representing all possible word-meanings. These two spaces are linked together by actual words within a language represent acquired form/meaning pairs. Because their meanings are always contextual, we will set aside pronouns and points here. Phonological or semantic similarities can be conceptualized within these spaces in terms of proximity. For example, the forms *pat* and *bat* as minimal pairs, are immediately adjacent to each other in English phonological space, but their distinct meanings are far apart in the corresponding semantic space. In both signed and spoken languages, there are dense links between phonological and semantic spaces, but each link is learned as a unit.

The pair of equivalent gestural form and meaning spaces are populated quite differently. Gesticulations do not represent individual learned form/meaning pairs, but flexible linkages so that gradations in meaning, for example in path, manner, shape or spatial relationships correspond to gradations in form. All possible pairs of form and meaning may appear as actual well-formed gesticulations. In other words, we do not learn that a particular gestural form means ‘round’, but rather how to use gesticulations to depict any round entity and how to correctly interpret depictions of round entities across contexts including completely novel forms. Emblems, like lexical items, are stable conventionalized form/meaning pairs that are learned individually, and inventories of these symbols are open-ended. Their meanings may be general or specific, and like lexical items, the same form may be linked to multiple meanings. The V handshape (✌) with the palm outwards, for instance may express the meanings *2*, *peace* or *victory* depending on context. Both emblems and lexical items must be learnable as form/meaning pairs and also continuously used and learned to maintain themselves within a population. This is not the case for gesticulations which may be completely novel and yet perfectly functional as depictions. Emblems are free from phonological and syntactic constraints, so the expectation might be for gestural space to be filled with large numbers of emblems. But this is not the case.

4.1 Emblems as Stable Shared Units

Emblems as symbols, are distinct from gesticulations and points also and from lexical words by their lack grammatical properties. The form of an emblem may be arbitrary or motivated in

some way but the form/meaning pair must be stable and shared between the gesturer and the interlocutor. An emblem is successful and persists within a population to the degree that it is shared. The most successful are shared by virtually everyone, for example the *thumbs up* (👍) gesture. Emblems may persist at low frequencies when they are used only by smaller sub-populations, allowing for secret gestures known only to the initiated. Gesticulations and points do not need to be learned as units in order to express their meanings. Copies of emblems may be inexact, and the expressions of these traits will vary in context, so to persist for any length of time as units, emblems must be maintained within narrow ranges of form and meaning. Divergent variants that fall outside of the conventionalized range for a particular emblem may be uninterpretable, interpreted as ill-formed or as an instance of some other gesture. In evolutionary terms the pressures maintaining an emblem represent stabilizing selection (Fig. 3c).

Emblems are the most word-like of the gestures, but their adaptiveness arises in part from being free from grammatical constraints. Their meanings that can be paraphrased linguistically, and their forms are interpretable with or without accompanying speech. They may be used as substitutes for words and phrases within otherwise spoken utterances or in contexts in which speech is not possible or not desirable. With possible combinations of handshapes, orientations, locations in space and movements and their symbolic nature emblems as a group are highly evolvable and their potential inventories are very large, but this potential seems underexploited in speaking populations. Actual inventories of emblems tend to be rather constrained. The gestural equivalents of minimal pairs are uncommon. In gestural form space, the V form for number *2* with the index and middle fingers extended, has two neighbors close by in the emblems representing *1* with an extended index finger and a form representing *3* with the addition of the extended ring finger. Gestures for *4* with the addition of an extended ring and pinkie fingers and *5* with all five digits extended are slightly further away. Pairs like *thumbs up* (👍) and *thumbs down* (👎) differ only in orientation, and they have directly opposing meanings. Between these two poles, gradations in meaning can be expressed with variations in orientation, but these forms represent gesticulations modifying an existing emblem depicting dials or gages turned up or down, not individual emblems learned as units. Rotating other emblems does not produce the same variations in meaning, even when the meanings might seem to allow for this. The *okay* emblem (👌) cannot be rotated to indicate degrees of okay-ness or its opposite.

Inventories of possible emblems are enormous, but inventories of actual emblems are limited, especially relative to lexical inventories. This is attributable in part to their holistic nature and a lack of grammatical properties. Lexical words are phonologically structured and are composed from a limited inventory of phonemes that conform to general phonotactic and prosodic constraints of a language. Categorical distinctions, for example between /p/ and /b/, allow for minimal pairs that are immediately adjacent in phonological space (i.e. *tap/tab* and *pad/bad* in English) yet perceived as entirely distinct words. Words also have syntactic properties and fall into a very limited number of grammatical categories which determine their syntactic behaviors. Words have internal structures that are not present in emblems, but this makes them easier to learn and use. Much about a lexical item can be gleaned from relatively little exposure, and lexical items are continuously recombined with other words through productive grammatical operations. This helps to maintain their semantic boundaries, but also produces new niches for new words with related (i.e. *dog/puppy*), contrasting (*dog/cat*) and opposing (*pedigree/mongrel*) meanings and ordinate (*dog*), subordinate (*poodle, pug*) and superordinate (*pet/mammal*) semantic relationships. Within grammatical systems, the more words there are the more words there can be.

It would be possible to coin emblems with the meanings *dog, puppy, poodle* and *mammal*, but outside of a productive grammatical system none of these emblems are likely to be used enough to maintain themselves within a population. Emblems are highly evolvable, but as inventories ex-

pand individual emblems are squeezed closer and closer to each other in gestural form space making their forms harder to distinguish. The chances for each emblem to be continuously used and learned so that it persists within the system also declines. These are constraints apply to gestural traits within a symbolic but a-grammatical niche, when they are moved out of this niche, former emblems can evolve very quickly.

4.2 From Gestures to Lexical Signs

Emblems and lexical signs in sign languages may appear similar, but signs are phonologically or syntactically structured. Lexical signs may be derived from non-linguistic gestures, obtaining linguistic properties during the process (Pfau & Steinbach, 2006; Sandler & Lillo-Martin, 2006; Wilcox, 2004). In the abstract, this process involves the reanalysis of a gestural form. It is assigned a grammatical category label and its holistic form is decomposed into phonological elements of handshape, place of articulation (POA) and movement consistent with the phonological constraints of the sign language. In fully lexicalized signs, these three components are fixed parts of a lexical unit that vary little with context. The forms and meanings of many lexical signs suggest that they were derived from gestures early in the history of their respective languages. Examples from ASL that appear similar across many sign languages are glossed EAT and DRINK (Meir, Padden, Aronoff, & Sandler, 2007) derived from gestures mimicking putting food in the mouth with a hand and drinking a liquid from a container respectively. Whether the original gestures these signs were derived from were stable emblems or gesticulations, the derivational processes would be similar.

An evolutionary approach makes it possible to fill out more of these processes. There is an initial state in a signing population in which there is a concept that can be expressed gesturally but for which there is not a lexical sign. At least once, but likely multiple times, a signer reconstructs a sign from a gestural performance as described above. By doing so the signer both adds a new item to their individual lexicon and also introduces a new linguistic trait into the population. When this trait is expressed in future utterances, it appears as a lexical sign that is then learned and copied by others and added to their own lexicons. As this happens the trait spreads within the population. The new lexical item is only successful if it is copied as a linguistic trait. It may not entirely displace the otherwise similar gestures, but to be successful it must eventually supplant these gestural traits in its linguistic role. This is a process of disruptive selection (Fig. 3b), splitting an initial gestural trait into a gestural trait and a new linguistic trait that conforms to all of the grammatical constraints within the language. If there are multiple new linguistic traits competing for the same role, one will eventually displace the other competitors either because it is selected for or simply by chance and drift. The relative size of the lexicon and a trait's interactions within the emerging grammatical system will determine the relative strength of stabilizing selection (Fig. 3c). When lexical inventories are smaller, early in the emergence of a new sign language, stabilizing selection is weak allowing for lexical items with wide variations in function and meaning. As the lexicon expands and grammatical systems are elaborated, lexical traits increasingly bump into each other and the relative strength of stabilizing selection increases, reducing variation in function and meaning for individual traits.

Lexical signs may be eventually derived into functional morphemes. This grammaticalization is understood as a change from an initial state with only lexical item, followed by a subsequent stage with both the lexical item and a more grammaticalized variant derived from it coexisting until the newer grammaticalized form displaces the lexical form or the two diverge into more distinct forms. Again, the aspectual and modal variants of *have* in English where derived from the verbal *have* through this process. It is possible to elaborate on these general processes from an evolutionary perspective. When a grammatical niche is present but otherwise empty, the role will be temporarily filled by lexical items.

Any lexical item with a functional range that extends into the empty grammatical role may

trigger a process of disruptive selection in which variants with purely lexical values and those with purely functional values are favored but any intermediate values between the two are selected against. This splits a single initial trait into two, with a lexical variant and a new functional variant. This has been observed in the derivation of an aspectual variant from an ASL verb FINISH (Janzen, 1995). The forms of these signs remain phonologically similar with a flip of one or both hands from a palm up to a palm down orientation. After an initial trait is split into two or more, resulting functional traits will undergo semantic bleaching and phonological reduction over time, grammaticalization processes that correlate with directional selection (Fig. 3a). In many ways, the processes by which new lexical signs with fixed forms are derived from gestures are similar to processes of lexicalization and grammaticalization that have been observed in spoken languages, but these are not the only ways that gestures may become linguistic.

4.3 Pointing Gestures

Points, uniquely, project a vector as a mental construct from a body part or grasped object towards a referent (Kita, 2003). Not all gestures that do so are traditionally grouped together as points, but all of these gestures involve the same underlying heavily innate trait. For example, a gesture with the V handshape and a downward palm directed at a referent with a meaning like '*I see you/ that'* or '*Look at that'* represents a combination of a point with other representational elements. Points may be used either imperatively to make requests or demands, or declaratively when they are used to inform, but they all involve the establishment of shared attention with an interlocutor (Kita, 2003). The prototypical point gesture involves an extended index finger, but many other gestures may utilize this handshape without pointing, for example to trace shapes or motions. These gesticulations are distinct from points in that they depict rather than point to a referent.

Points may be produced with any finger or handshapes, with grasped objects like pens, with legs and feet, head tilts and chin thrusts. The referents of points may be any concrete, abstract or imagined entity or location in any direction relative to the gesturer. Directions themselves may be referents, with absolute values like north or south, or relative values like right or left, forwards or backwards. Entities that are not present or are imagined can be pointed to as locations in space. Points forward or backward may represent the future or the past and points downward may refer to the here and now. The referents that points project vectors towards are not parts of the gesture themselves, suggesting that these forms are all variants of a single flexible trait, not multiple traits with related meanings as would be the case with emblems or words. Evolution can target these variations in form and other patterns of use, but it cannot alter the directionality of pointing gestures themselves. The projection of vectors towards referents appears to be heavily innate and cannot be unlearned, and the locations of referents in space do not represent heritable variation that evolution can target.

The notion that pointers are heavily innate is supported by the very early appearance of both imperative and informative points in human infants by around 10 months of age (Goldin-Meadow, 2005; 2003), by the resistance of points to modifications based on learning, and the capacities of captive chimps for learning imperative but not informative human-like pointer gestures (Francis, 2015; Tomasello & Call, 2014). Points may also suggest something of how gestures can evolve from purely learned traits into heavily innate traits. Early in our evolutionary lineage, our pre-modern ancestors would have had to learn points much as modern chimps do, without any particular adaptations for doing so and late in development through repeated exposure. Individuals with brains that happened to allow them to learn points faster and earlier would have had reproductive advantages, even if the ability to acquire and use points was not the direct target of the selection pressures. It has been hypothesized that changes like this produced positive feedback loops, selecting for the genes to build brains adapted for learning and using gestures, and so increasing the frequency of gestures used in the population. This

in turn increased the selective pressures on brains and so on (Arbib, 2012). Over many generations, the effect was to push the emergence of pointing increasingly early in development, until they reached the modern state where their early emergence facilitates subsequent language acquisition.

This heavy innate component also has consequences for the evolvability of points within populations. Any changes to points must be learned on top of the innate vector projection function. The remaining evolvable aspects include particular hand configurations and orientations associated with particular referents and patterns of appropriate use. It is possible to learn to inhibit points or particular forms if for example, it is considered impolite to point to people or to point to people with an extended index finger. Other forms such as an open hand with the palm up may be substituted. Beckoning gestures, involving one or more movements from an open to a partially closed position projected at a referent may be produced with the palm upwards with the movement from the base knuckles of the hand. In East Asia, the palm is oriented downward when beckoning a person. Beckoning gestures may be amplified in context by shifting the movement from base knuckles to the wrist, elbow or shoulder, or minimized by shifting the movement to the middle joint of the index finger. These minimized forms may then take on more direct personal or intimate meanings. The underlying vector projection function of pointers cannot be altered, but within this range of evolvability, sign languages have evolved pronominal signs from pointing gestures.

4.4 From Pointer Gestures to Pronouns

Sign languages include a number of sign types that are phonologically and otherwise grammatically structured but with flexible gesture-like properties. Unlike lexicalized signs in which handshapes, movements and POA are fixed parts of the sign (Sandler & Lillo-Martin, 2006; Brentari, 1998), one or more of these components are flexible and context-dependent. In singular pronominal signs referring to non-1st person referents, the handshapes and movements are fixed but POA is not. Like pointing gestures these signs are directed towards present referents or at locations in space assigned to absent referents within the discourse (Cormier, 2014; 2012; 2007; Cormier, Schembri, & Woll, 2013). Unlike points however, pronominal signs function as grammatical pronouns. These and other flexible signs do not represent intermediate stages along diachronic pathways towards lexicalization and grammaticalization and fully fixed forms. They have evolved from gestures, but under different selective pressures than those that derive lexical signs with fixed forms.

The flexible properties of some types of signs have given rise to what is referred to as the listability issue in sign languages (Wilbur, 2013; Quer, 2011; Rathmann & Mathur, 2011; Liddell, 2003; and many others). If lexical items are understood as stable form/meaning pairs 'listed' in a mental lexicon as units, pronominal and other flexible signs create problems. There is no way for all of the possible variations in form and meaning to be listed in any plausibly-sized lexicon. Many parts of language are not lexical items, so the kernel of this issue is not that these elements cannot be listed individually as lexical items. Rather these signs are dependent on context which cannot be grammaticalized and to which syntactic operations are not sensitive. In other words, personal pronominal signs function syntactically like pronouns in spoken languages, but their forms behave like gestural points and depend on factors outside of syntax.

The treatments for these signs with flexible forms that have been proposed include: treating them as gestural in a framework with a gradient distinction between gesture and grammar (Liddell, 2003; and many others); as combinations of grammaticalized and gestural components unique to sign languages (Rathmann & Mathur, 2011; and others); or as entirely linguistic, arising through interfaces between semantics and phonology again unique to sign languages (Wilbur, 2013; Quer, 2011; and others). From the current evolutionary perspective, the details of these proposals and debates can be set aside. The pronominal systems in sign and spoken languages are functionally equivalent evolved responses to similar problems using different resources. What is interesting

here are the evolutionary processes that derive linguistic traits with flexible forms from existing gestural traits in signing populations. Importantly, the pronominal systems in spoken languages vary far more than those in even completely unrelated sign languages, all of which seem to have converged on very similar solutions. This is an indication of selection, rather than chance.

The contrasts among sources, addressees and non-participants are grounded in discourse roles (Harley & Ritter, 2002), and are invariant across all populations. What has evolved in all spoken languages are grammaticalized ways of distinguishing among 1st, 2nd and 3rd persons, as well as grammatical number, gender and case that may also be involved (Harley & Ritter, 2002). These systems are diverse, but they are evolved solutions to the same problem. Without accompanying gestural points, the possible interpretations of utterances containing personal pronouns would be insufficiently constrained without at least this three-way contrast. In sign languages, singular 1st person forms may be fixed, as their referents are always the signer as the source, even in cases of role-shift and quotation. Non-1st singular personal pronouns use the same handshape with an extended index finger directed towards both addressees and non-participants. This has led to claims that sign languages, as a group, grammaticalize only 1st/non-1st contrasts rather than a three-way contrast in apparent violation of a linguistic universal (Cormier, 2014; Harley & Ritter, 2002; McBurney, 2002).

If this apparent universal is an evolved property of spoken languages, rather than an innate component of human language itself, the pronominal systems of sign languages make sense: they have not evolved solutions to problems they do not have. It is in fact possible for signing populations to evolve ways of marking distinctions between 2nd and 3rd person referents (Alibasic Ciciliani & Wilbur, 2006), but they are not under the same selective pressures to constrain interpretations that spoken languages are. The relevant distinctions are already sufficiently clear for addressees, the only persons for whom the distinctions matter. But sign languages face a different problem. Points are highly resilient and infants acquiring a sign language will always learn pointing gestures first. Points cannot be displaced from signing populations and they will always 'work' as points, projecting vectors towards their referents. Pronominal signs may be derived as functional elements directly from gestural points, rather than through an intermediate lexical stage (Pfau & Steinbach, 2006; Wilcox, 2004), but for this to occur gestural points have to be inhibited within grammatical utterances leaving only pronominal forms that conform to the relevant grammatical constraints.

The current evolutionary framework suggests mechanisms by which this may happen. Natural langauges seem to need pronominal systems, but in the very early emergence of sign languages, there are no functional morphemes to fill these niches. Points are already present in these initial states with the necessary referential functions but they do not conform to the emerging grammatical constraints. Early on, they will be used in pronominal roles, but also everywhere else that points are useful. Eventually, the system must evolve to a point where signers can identify well and poorly formed instances of pronominal signs according to the constraints of their language. A plausible evolutionary scenario involves a form of disruptive selection followed by directional selection. It is not necessary for disruptive selection to select for pronominal points, all that is required is for variants that do not conform to linguistic constraints to be inhibited within grammatical utterances. Points cannot be displaced entirely, but this would leave only those variants filling pronominal roles. Directional selection can then target the remaining pronominal signs, reducing variation in handshape and orientation, and pushing these signs towards a narrow range of variants. Due to the innate and flexible nature of points, forms that are eventually grammaticalized are already present in the population prior to selection. Instead of being conventionalized as new units as new lexical signs would be, a subset of the existing variants are selectively retained as well-formed pronominal signs, while other variants are inhibited. The resulting pronominal signs have been

'domesticated' by linguistic selection and have evolved constraints on their interpretations that are not present in the 'wild' gestural points not bound by linguistic constraints.

The pronominal systems derived from pointing gestures evolve very quickly and early and seem to be resistant to some sorts of additional changes. Once a range of values for handshape and orientation have been conventionalised, other handshapes and/or orientations may evolve for possessives, locatives and demonstrative pronouns, introducing further constraints on possible interpretations. At least hypothetically, it would be possible for a sign language to evolve distinct handshapes to distinguish addressees and non-participant referents of non-1st person signs, but such forms do not seem to emerge. These traits would be functional, but they would offer few additional constraints on interpretations and the selective pressures favoring them over the existing traits would be weak. Changes towards pronominal systems like those of spoken languages with distinct non-directional fixed forms for 1st, 2nd and 3rd persons are also possible, but extremely unlikely. Directional pronominal signs are able to pick out individuals from groups. Pronouns in spoken languages cannot do this without context or pointing. So for sign languages, any changes away from flexible directional pronominal signs would represent a loss of adaptiveness, or a slide down from a more adaptive peak into a less adaptive valley.

4.5 Gesticulations

Gesticulations or representational gestures iconically depict their referents (McNeill, 2016; Özyürek, 2012; Kendon, 2004; Liddell, 2003; and many others). Emblems may be iconic, but gesticulations must be and they must also be flexible. This broad group of gestures includes depictions of properties and changes of properties of entities; locations, orientations, activities and motions of entities and parts of entities relative to wholes; and interactions and relations among entities including manipulation and caused motion. The entities, properties and events that are depicted may be concrete, abstract or imagined. Within composite utterances, the interpretations of these gestures are constrained by the accompanying speech stream, but similar meanings can be depicted in more or less equally effective ways by different people as well as by the same person at different times and in different contexts.

In other groups of gestural traits, selection reduces variation, but the variation within gesticulations is vast. This can be explained if some of the variation within these gestures is not heritable, and also if some of this variation is adaptive and a product of selection. A circle for example may be depicted by tracing the shape with an index finger from any point within its circumference, moving clockwise or anticlockwise, with two fingers moving in opposite directions, or with one or both hands held in space with extended thumbs and curved fingers depicting edges among other possibilities. Similar strategies are available to represent spheres and columns and other shapes and combinations of shapes of different sizes and orientations in space. As far as gestural evolution is concerned, these properties of the world are invariant; they do not evolve. What has to evolve are ways to depict all of these meanings with a learnable inventory of highly flexible traits. Emblems are highly evolvable as a group, with new emblems corresponding to new meanings. To express the same range of meanings expressed by gesticulations with emblems would require an unlearnable, and therefore unusable, number of traits. Gesticulations employ an entirely different strategy of form/meaning mapping, but their flexibility limits their evolvability rather than increases it. Instead of a huge number of separate traits, gesticulations seem to involve relatively few.

Providing an inventory of the relevant traits within gesticulations is not possible here, but they have a deep evolutionary history, based in the imitative capacities of primates and especially monkeys and apes. The primate systems of mirror neurons allow an individual to understand and replicate the behaviors of others in terms of their own bodies. Simians are famous for these abilities (*monkey see, monkey do*; *to ape*), but these capacities were far more elaborated in the human lineage to a point here an individual can use ges-

tures to intentionally represent entities and events in the world displaced in space and time (Arbib, 2012). The physical properties of the world do not evolve, but our capacity to iconically depict those properties in ways that are reliably interpreted with their intended meanings certainly has. There may be more than one way to depict any entity, activity or event, but as long as those forms are interpretable as depictions of the intended referents, different variants may be selectively neutral relative to each other.

Contextual factors also limit the evolvability of gesticulations, including the context of composite utterances with speech. Studies in which non-signers were asked to describe events of motion and location using only speech, speech and gesture, or only gesture show that participants use more elaborate and detailed gesticulations when they are not also speaking (Schembri, Jones, & Burnham, 2005). These gestural productions without speech were similar in many ways to classifier predicates produced by signers describing the same events. Iconic representations of concrete events might be expected to be broadly similar across the gestural productions of non-signers and the linguistic productions of signers, but for this to be the case, non-signers must already have the capacity to adjust their gestural productions to carry the entire meaning of an utterance without speaking. They must also be able to partition meanings between speech and gestural streams relative to each other within composite utterances. As contextual factors do not evolve, what had to evolve were traits flexible enough to be adaptive across a very wide range of conditions.

An additional source of variability beyond the reach of gestural selection may come from the environments in which these gestures develop and the routes by which they are learned. Gesticulations cannot be learned as stable units. Imitation plays a role but what is imitated are general strategies and heuristics rather than individual performances. A child can assume that gestural performances are adaptive, but they will be exposed to different strategies employed by different role models to reach similar goals. A child learns by evaluating the relative successes of different strategies relative to goals, based on their own behavior and those of others. Jablonka and Lamb (2014) refer to these processes as exploration and selective retention and socially mediated learning. A learner may attempt multiple strategies to reach goals, retaining those that appear more successful and pruning away those that are less so. The resulting idiosyncratic combination of strategies are products of development, not evolution, but they will be adaptive for the individual. As long as the resulting gestural performances are iconic and their interpretations are reliable in context, there may be no strong selection pressures pushing gesticulations to evolve in any direction at a population-level. This seems to be the case at least for co-speech gesticulations within composite utterances, where the speech stream constrains their interpretations. This is not the case for classifier predicates in sign languages.

4.6 The Evolution of Classifier Predicates

The group of signs referred to here as classifier predicates is diverse, but includes productive signs that denote states and events of spatial motion and location of concrete entities. Signs that depict sizes and shapes (SAS signs or descriptive classifier predicates) will be set aside here. These signs are also referred to as depicting verbs (Liddell, 2003) and poly-componential signs (Schembri, 2003). They indeed depict spatial meanings in ways similar to gesticulations but their handshapes, movements and POA components are independently productive. Only the handshapes are grammaticalized elements with fixed forms; the movement and POA are flexible. Multi-morphemic words tend to be multi-syllabic but classifier predicates, like lexical verbs with fixed forms, are overwhelmingly monosyllabic built around a single phonological movement (Sandler & Lillo-Martin, 2006; Brentari, 1998). Goldin-Meadow (2005) reports that home-signers tend to reanalyze holistic gestures into proto-morphological parts that can be recombined to produce new forms with meanings. In new signing populations this process or 'fracturing' of gesticulations (Arbib, 2012) leads to the very early emergence of classifier predicates (Aronoff, Meir, & Sandler,

2005). These signs are nearly universal in sign languages, although direct linguistic equivalents have not been reported in spoken langauges (Sandler & Lillo-Martin, 2006; Aronoff, Meir, & Sandler, 2005; Emmory, 2003).

Classifier predicates are the most iconic of verbal signs. They have been analyzed as adaptations to the visual/spatial modality, on the assumption that languages will tend to exploit iconicity when possible (Sandler & Lillo-Martin, 2006; and many others). Yet, iconicity is not the only factor. If it were, the expectations would be not only that the majority of verbal signs are iconic in some way, but that more iconic forms would always be favored over less iconic ones. If greater flexibility allows for greater iconicity than fixed forms, the expectation would be that all verbs would evolve to become increasingly productive and more like classifier predicates. This does not seem to be the case. Many lexical verbs are derived from iconic gestures, like EAT and DRINK, and many other are derived from classifier predicates, including READ, WRITE, DANCE and CAR-CRASH from ASL. These 'frozen' verbs remain iconic, but their movements and POA are no longer productive. While performances of these signs can be reanalyzed online as productive forms, this does not displace the frozen lexical verbs within the population (Sandler & Lillo-Martin, 2006). Classifier predicates continue to co-exist with fixed lexical verbs, within the same systems. Classifier predicates are adaptations to the visual/spatial modality, but they are also adaptations to linguistic systems containing verbs with fixed and less productive forms, and despite their flexible forms, the meanings expressed by classifier predicates are tightly constrained in ways that gesticulations are not. Why should this be the case?

Classifier predicates appear early in emerging sign languages from an initial state in which gesticulations are used to describe events of motion and location but also other events for which there are not yet lexical verbs. Some of these gestures are lexicalized as single fixed signs, as discussed above. Classifier predicates evolve through a different route, one by which holistic gestures are fractured into parts that can be productively recombined. Some insights into this process are found in Nicaraguan Sign Language. The founding members of this population (cohort 1) use signs that may depict spatial paths, manners of motion or both together, as gesticulations can. Members who entered the population later and at younger ages (cohorts 2 and 3) are much more likely to produce signs that depict only path or manner, not both (Arbib, 2012; Sengas, Özyürek, & Goldin-Meadow, 2013; Senghas, Kita, & Özyürek, 2005; Senghas, 2003). Forms depicting only paths or manners were always present, but as the language emerged, disruptive selection favored those variants over intermediates between the two that express both, effectively splitting initial gestural traits depicting motions into path-depicting and manner-depicting traits within this language.

The movements that result from this fracturing remain iconic, but the resulting signs are less iconic than those that depict path and manner together. Languages do not lexicalize both paths and manners together within single verbs and thus constrain the size of the lexicon but these signs do not have fixed lexical forms, so the size of the lexicon is not relevant. The important factor here seems to be constraints on interpretation. The interpretations of co-speech gesticulations are constrained by the content of the speech stream so there is little pressure to limit forms to depicting only either paths or manners. Without an accompanying speech stream, the selective environments for classifier predicates are different. The flexibility of gesticulations to depict any real or imagined entity or event is a problem. In response sign languages have evolved ways to reliably constrain interpretations within the signs themselves.

This is accomplished with grammaticalized handshapes, of which there are several groups: those that refer to the participant in an event as whole entities but as a member of a grammaticalized category (thus the term 'classifier'), including vehicles, aircraft, animals and upright entities; those that refer to the participant as part of a whole, including body-parts like heads or wings; and those that refer to one participant manipulating another with a grasping limb, such as a hand, claw or mouth (Sandler & Lillo-Martin, 2006;

Emmory, 2003). Each classifier predicate has a handshape which constrains the ways in which the form's movement and POA are interpreted. Whole entity classifier handshapes only refer to concrete entities, constraining the interpretations of the phonological movements to spatial paths or manners of motion. Body-part classifier handshapes constrain the interpretations of movements to changes of relative location or orientation, or manners of motion, of a part relative to a whole. Handling classifier handshapes restrict the interpretations of movements to caused motion and location. These handshapes inventories are limited and incongruent combinations of handshapes with movements and POA are considered ill-formed. The constraints on interpretations provided by grammaticalized classifier handshapes help also allows for otherwise similar movements and POA to be interpreted differently when associated with types of signs. In SAS signs, movement and POA are interpreted as depictions of physical dimensions and surfaces rather than motion and location. Flexible directional movements and POA may also be interpreted in non-spatial and more abstract ways, including in the markers associated with directional or agreeing verbs (GIVE, SEND) that indicate the recipients, and sources of transfers without describing actual spatial paths of transfers.

On this view, classifier predicates are not just adaptations in sign languages to iconically depict spatial motion and location. Non-signers are capable of producing similar forms with similar meanings, even if they tend not to do so in the composite utterances with a co-articulated speech stream. Classifier predicates have also evolved grammatical constraints on the interpretations of these flexible forms so that they are reliably interpreted with their intended meanings in utterances channeled through a single modality.

5. CONCLUSIONS AND DISCUSSION

This proposal has outlined an evolutionary approach to three sorts of gestures identified in the literature, treating them as adaptations to distinct niches within the broader system that includes grammatical language. Emblems are the more evolvable within populations as they are symbolic form/meaning pairs learned as units. The cultural evolvability of points within a population is limited to those components that are not contextual or innate. Because their adaptiveness depends entirely on their iconicity and flexibility, gesticulations are the least evolvable within a speaking population. Each group of gestures is adapted to the expression of particular kinds of meaning within the broader system that includes language, but outside the reach of grammatical constraints. The relationships and interactions among gestural and linguistic systems is conceptualized here, at least in outline, in terms of a landscape composed of adaptive niches representing peaks, separated by less adaptive valleys. Linguistic and gestural evolution can be understood as movement of a trait within least landscapes over time, with the most drastic movements being the shifts of gestural traits to linguistic niches in the emergence of sign languages.

There are many issues and questions here that will have to be left to future research, but this framework does allow for some testable predictions. When initial conditions are similar, the products of linguistic and gestural evolution will be similar, even across widely separated populations. This is consistent with observations of some broad similarities across unrelated sign languages in terms of pronominal systems and classifier predicates evolved from points and gesticulations respectively. However, there are sign languages with otherwise 'atypical' properties relative to more well-studied community sign languages like ASL or German Sign Language, used by populations of mostly unrelated Deaf people clustered in urban areas. The languages with atypical properties tend to be village sign languages, used by populations of closely related Deaf and hearing people in more rural areas (Pfau, 2012). So why do some sign languages evolve properties that are otherwise atypical? The nature of the relationships within a community, density, and the proportions of native, late-learning and non-native members may be factors, but evolutionary theory suggests that population size is the most important. Drift, or changes due to chance rather than selection,

can play a larger role in smaller populations. The prediction would be that smaller and more isolated signing populations are more likely to evolve in ways that are atypical compared to larger signing populations.

This brief analysis is only able to scratch the surface of some of the potential applications of evolutionary theory to human gesture and language. Evolution theory deals with issues of function, but without assuming functional explanations for every property or change within a system. It comes with a well-established formal system that can be applied to both gestural and linguistic traits. This offers new ways to approach long standing issues such as proposed universals and listability in sign languages. Because it applies only at the population level, it is not inconsistent with analyses at other levels from other frameworks, and may draw from functional and formal analyses of gesture and language, and acquisition and development at the level of individuals. There is saying that nothing makes sense in biology without evolution. Perhaps it is time to consider how true this might be for human gesture and language.

REFERENCES

Özyürek, A. (2012). Gesture. In R. Pfau, M. Steinbach, & B. Woll, *Sign Language: an international handbook.* Berlin: Mouton de Gruyer.

Alibasic Ciciliani, T., & Wilbur, R. (2006). Pronominal system in Croatian Sign Language. *Sign Language and Linguistics, 9*(1-2), 95-132.

Arbib, M. (2012). *How the Brain Got Language: The Mirror System Hypothesis.* Oxford: Oxford University Press.

Armstrong, D. F., & Wilcox, S. E. (2007). *The Gestural Origin of Language* . New York, USA: Oxford University Press.

Armstrong, D. F., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the Nature of Language.* Cambridge, UK: Cambridge University Press.

Aronoff, M., Meir, I., & Sandler, W. (2005). The paradox of sign language morphology. *Language, 81*, 301-344.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology* . Cambridge, Mass.: the MIT Press.

Brinton, L. J., & Traugott, E. C. (2006). *Lexicalization and Language Change* . Cambridge , UK: Cambridge University Press.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World.* Chicago: The University of Chicago Press.

Call, J., & Tomasello, M. (2007). Ape gestures and the origins of language. In J. Call, & M. Tomasello, *The Gestural Communication of Apes and Monkeys.* Mahwah, New Jersey : Lawrence Erlbaum Associates.

Corballis, M. C. (2009). From Hand to Mouth: The Gestural Origins of Language . In *Language Evolution* . Oxford, United Kingdom: Oxford University Press.

Cormier, K. (2007). Do all pronouns point? Indexicality of first person pronouns in BSL and ASL. In P. Perniss, R. Pfau, & M. Steinbach, *Visible Variation: Comparative studies on sign language structure* . Berlin: Mouton de Gruyter .

Cormier, K. (2012). Pronouns. In R. Pfau, M. Steinbach, & B. Woll, *Sign Language: an international handbook.* Berlin: Mouton de Gruyer.

Cormier, K. (2014). Pronouns, agreement and classifiers: What sign languages can tell us about linguistic diversity and linguistic universals. *UCL Working Papers in Linguistics, 26*.

Cormier, K., Schembri, A., & Woll, B. (2013). Pronouns and pointing in sign languages. *Lingua, 137*, 230-247.

Emmorey, K. (2003). *Perspectives on Classifier Constructions in Sign Languages.* Mahwah NJ: Lawrence Erlbaum Associates .

Francis, R. C. (2015). *Domesticated: evolution in a man-made world.* New York: W. W. Norton & Company .

Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think.* Cambridge Mass: Harvard University Press.

Goldin-Meadow, S. (2005). *The Resilience of Language: What gesture creation in deaf children can tell us about how all children learn language.* New York, USA: Psychology Press.

Goldin-Meadow, S. (2012). What modern-day gesture can tell us about language evolution. In M. Tallerman, & K. R. Gibson (Eds.), *The Oxford Handbook of Language Evolution* (pp. 545-557). New York, USA: Oxford University Press.

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, Sign and Language: the coming of age of sign language and gesture studies. *Behavioral and Brain Science*.

Harley, H., & Ritter, E. (2002). Person and number in pronouns: a feature-geometric analysis. *Language, 78*(3).

Heine, B., & Kuteva, T. (2012). Grammaticalization theory as a tool for reconstructing language evolution . In M. Tallerman, & K. R. Gibson (Eds.), *The Oxford Handbook of Language Evolution* (pp. 512-527). New York, USA: Oxford University Press.

Jablonka, E., & Lamb, M. J. (2014). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral and Symbolic Variation in the History of Life* (Revised Edition ed.). Cambridge, Massachusetts, USA: A Bradford Book The MIT Press.

Janzen, T. (1995). *The polygrammaticalization of FINISH in ASL.* Thesis (MA). University of Manitoba .

Kendon, A. (2000). Language and gesture: unity or duality? In D. McNeil, *Language and Gesture* (pp. 47-63). Cambridge UK: Cambridge University Press.

Kendon, A. (2004). *Gesture: Visible Action as Utterance.* New York : Cambridge University Press.

Kita, S. (2003). Pointing: a foundational building block of human communication. In S. Kita, *Pointing: Where Language, Culture and Cognition Meet* (pp. 1-8). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Langacker, R. (1999). *Foundations of Cognitive Grammar: Theoretical Prerequisites* . Stanford CA: Stanford University Press.

Liddell, S. (2003). *Grammar, Gesture and Meaning in American Sign Language.* Cambridge, United Kingdom: Cambridge University Press.

Liddell, S. K. (2011). Agreement disagreements. *Theoretical Linguistics, 37*(3/4), 161-172.

Lillo-Martin, D., & Meier, R. (2011). On the linguistic status of 'agreement' in sign languages. *Theoretical Linguistics, 37*(3/4), 95-141.

Mayr, E. (2001). *What evolution is.* New York: Basic Books.

McBurney, S. L. (2002). Pronominal reference in signed and spoken language: Are grammatical categories modality-dependent. In R. P. Meier, K. Cormier, & D. Quinto-Pozos, *Modality and structure in signed and spoken languages* (pp. 329-369). Cambridge: Cambridge University Press.

McNeil, D. (2012). *How Language Began: Gesture and Speech in Human Evolution.* Cambridge UK: Cambridge University Press.

McNeill, D. (2005). *Gesture & Thought* . Chicago: University of Chicago Press.

McNeill, D. (2016). *Why We Gesture: The surprising role of hand movements in communication.* Cambridge UK: Cambridge University Press.

Meir, I., Padden, C. A., Aronoff, M., & Sandler, W. (2007). Body as subject. *Journal of Linguistics*, 531-563.

Pfau, R. (2012). Manual commuication systems: evolution and variation. In R. Pfau, M. Steinbach, & W. B, *Sign Language: an international handbook.* Berlin: Mouton de Gruyer.

Pfau, R., & Steinbach, M. (2006). *Modality-Independent and Modality Specific Aspects of Grammaticalization.* Potsdam: Universitäs-Verlag.

Quer, J. (2011). When agreeing to disagree is not enough:Further arguments for the linguistic status of sign language agreement . *Theoretical Linguistics , 37*(3/4), 189-196.

Rathmann, C., & Mathur, G. (2011). A featural approach to verb agreement in sign languages. *Theoretical Linguistics, 37*(3/4), 197-208.

Ritt, N. (2004). *Selfish sounds and linguistic evolution: A darwinian approach to langauge change.* Cambridge, UK: Cambridge Univeristy Press.

Sandler, W., & Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals.* Cambridge, UK: Cambridge University Press.

Schembri, A. (2003). Rethinking 'classifiers' in signed languages. In K. Emmory, *Perspectives on Classifier Constructions in Sign Languages.* New York: Lawrence Erlbaum Associates.

Seddon, C. (2020). *Humans: from the beginning* . London: Glanville Publications.

Sengas, A., Özyürek, A., & Goldin-Meadow, S. (2013). Homesign as a way-station between co-speech gesture and sign language: the evolution of segmentation and sequencing. In R. Botha, & M. Everaert (Eds.), *The Evolutoinary Emergence of Language: Evidence and Inference* (pp. 62-76). New Yokr, New York : Oxford University Press.

Senghas, A. (2003). Intergenerational influence and ontogenic development in the emergence of spatial grammar in Nicaraguan Sign Language . *Cognitive Development*.

Senghas, A., & Coppola, M. (2001). Children creating language: how Nicaraguan Sign Language acquired a spatial grammar. *Pyschological Science, 12*(4).

Senghas, A., Kita, S., & Özyürek, A. (2005). Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science, 305*.

Senghas, R., Senghas, A., & Pyers, J. (2005). The emergence of Nicaraguan Sign Language: questions of development, acquistion and evolution . In J. Langer, P. S. T., & M. C. (eds), *Biology and Knowledge revisited: From neurogenesis to psychogenesis.* Mahwah N.J.: Erlbaum .

Talmy, L. (2000). *Towards a Cognitive Semantics.* Cambridge Mass: MIT Press.

Tomasello, M. (2008). *Origins of Human Communication* . Cambridge , Massachusetts, USA: The MIT Press.

Tomasello, M., & Call, J. (2014). Ape Gestures and the Origins of Language. In J. Call, & M. Tomasello (Eds.), *The Gestural Communication of Apes and Monkeys* (pp. 221-240). New York , USA: Psychology Press.

Wilbur, R. (2013). The point of agreement: Changing how we think about sign language, gesture and agreement. *Sign Language and Linguistics* , 221-258.

Wilcox, S. (2004). Gesture and langauge: cross-linguistic and historical data from sign languages. *Gesture*.