Appearing in: *Cognitive Semantics*, volume 4, number 2. Leiden-Boston: Brill

Combinance in the Evolution of Language: Overcoming Limitations

Leonard Talmy
Department of Linguistics and Center for Cognitive Science,
University at Buffalo, State University of New York
talmy@buffalo.edu

Abstract

For early pre-language hominins, the vocal-auditory channel of communication as then organized may have been unable to accommodate any enhancement in the transmission of conceptual content due to three limitations: comparatively low degrees of parameter diversity, iconicity, and fidelity. We propose that these limitations were overcome by an evolutionary development that enabled an advance from the fixed holophrastic calls of earlier species to the open-ended spoken language of our own species. What developed was a "combinant" form of organization.

Such combinance is a system in which smaller units combine to form larger units. At its smallest scale, this process yields a "clave". In a clave, generally, units from an inventory at a lower tier combine to form the units of an inventory at the next higher tier in accord with a particular set of constraints. In turn, such claves function as the smaller units that combine to form a larger unit, a concatenation, where the higher tier of one clave serves as the lower tier of the next. The longest such concatenation in language, to simplify, consists of these six successive claves: Phonetic features combine to form phonemes under the constraints of feature assembly; phonemes combine to form morphemes under the constraints of phonotactics; morphemes combine to form complex words under the constraints of morphology; morphemes and complex words combine to form expressions under the constraints of syntax; expressions combine to form a single speaker's "monolog" under the constraints of discourse rules; and such monologs combine to form an exchange between speakers under the constraints of turn-taking.

Our analysis characterizes communication at its most general and contrasts different channels of communication. In particular, the vocal-auditory channel of spoken language is extensively contrasted with the somatic-visual channel of signed language, whose classifier system largely lacks the three limitations of the former. To show this difference, the limitations are analyzed in detail (e.g., iconicity is shown to be based on six properties: prorepresentation, covariation, proportionality, proportional directness, cogranularity, and codomainality). In accord with this difference, the signed classifier system demonstrates the cognitive feasibility of communicating advanced conceptual content with little combinance, but the vocal-auditory channel is seen to have needed the incorporation of combinance for spoken language to evolve.

Keywords

language evolution – communication channel – sign language – classifier system – iconicity

1 Introduction[1]

For early pre-language hominins, the vocal-auditory channel of communication, as it was then organized, may have been unable to accommodate any enhancement in the transmission of conceptual content. This circumstance may have in effect constituted an evolutionary bottleneck. On the one hand, there may have been a selective advantage in such enhancement. On the other hand, that channel of transmission then had certain limitations impeding such enhancement. Our proposal is that these limitations, and hence the bottleneck, were overcome by a particular evolutionary development. This development–whether itself new or an elaboration of an already existing system–occurred in the organization of the cognitive system that underlay communication through the vocal-auditory channel. In this proposal, what evolved in that channel was what we will call a **combinant** form of organization or, in a single word, what evolved was **combinance**. In the broadest strokes (refined in section 1.4), combinance is a system in which smaller units combine to form larger units.[2]

Our proposal, then, is that as the cognitive system that used the vocal-auditory channel for communication increasingly came to incorporate a combinant form of organization, the type of communication it could support advanced from the holophrastic calls of earlier species to the spoken language of our own species.

In this article, we focus on the limitations on communication originally present in the vocal-auditory channel. Sections 2, 3, and 4 respectively address three major limitations: comparatively low degrees of parameter diversity, iconicity, and fidelity. We here thus mainly set the stage for the proposal that the development of combinance in the channel has overcome these limitations. A planned follow-up article will expand in great detail on what such combinance consists of.

The remainder of this Introduction provides summary overviews of communication (1.1), channels of communication (1.2), conceptual content (1.3), and combinance itself (1.4).

1.1 *Communication*

The following paragraph proposes, in synopsis, our proposed characterization of communication in its most general scope. The reasons for the choice of certain words in the synopsis are then addressed. The subsequent discussion soon narrows down from the general to a specific level—to the linguistic forms of communication that are our main focus of attention.

Communication is a process that has evolved or been contrived in which one entity, the **sender** executes certain actions that have the function of inducing particular responses in a certain other entity, the **receiver**. The sender's actions produce certain physical effects that the receiver can detect across the separation between them, the **medium**. Where such a communication process has evolved in a species, the two entities can be different components of a single cell, different cells within a single multicellular organism, or different individuals, whether themselves single celled or multicellular. This third alternative further includes the case where the sending individual is of one species and the receiving individual is of a different species. Especially where the communication is between two individuals of multicellular animal species, the physical effects produced by the sender's actions are understood as "stimuli" within a particular sensory modality, and the receiver's detection of them as "perception" within that modality.

The use of the word "contrived" in the first sentence above includes within our proposed notion of communication the case in which some aspect of the process has been effected nonevolutionarily. Examples include the case of writing and the case in which one instrument sends signals to another. Such forms of communication, though, are derivative in that certain aspects of them must be intentionally constructed by a sentient species in which communication has already evolved.

The kinds of "actions" that a sender can execute have a wide extent. They can range from the release of chemicals, to the movement of limbs, to vocalizations. And a receiver's "responses" can range from the production of chemicals, to fleeing, to the experiencing of a concept in consciousness. The first action and response examples just cited occur together, for instance, where a plant attacked by certain insects releases chemicals that induce other nearby plants of its species to produce certain other chemicals deterrent to those insects (e.g., Babikova et al. 2013).

The use of the word "function" excludes certain cases from communication. This word here indicates that the sender's actions have evolved due to the selective advantage—or were contrived for the purpose—of inducing particular responses in a certain receiver. This excludes cases where the actions have evolved or been contrived for another function. An example of such exclusion is the case where the vibrations of a fly caught in a web alert a spider to its presence and location. Such actions executed by a fly presumably evolved in its species for the function of facilitating escape, not of inducing a spider to approach—hence, not of "communicating" with the spider to that end.

The word "function" further excludes cases of "eavesdropping". For example, a bird of species Y might flee on hearing an alarm call from a bird of species X. However, the species X alarm call evolved, say, to induce mobbing in birds of species X, but not fleeing in birds of species Y—hence not to "communicate" with birds of species Y to that end.

The word "induce" also excludes certain cases from communication. Its sense of indirectly eliciting or fostering a certain response in what can be regarded as an otherwise autonomous entity excludes direct physical causation from our characterization of communication. Thus, a wild horse kicking a wolf away has gotten the wolf to move off through direct physical causation, not by inducing the wolf to exercise its own neuromuscular control—hence, not by "communicating" with it to that end.

The cases of display and camouflage can test the boundaries and appropriateness of our characterization of communication. With respect to display, certain static shapes or surface markings on a particular plant or animal as sender can have evolved with the function of inducing certain responses in a receiver, such as attracting a pollinator or, as a warning, inciting avoidance in a potential predator. However, insofar as the sender does not perform an active display, it does not execute an action. Still, this phenomenon seems to have enough of the cited characteristics to be regarded as communication.

Camouflage, on the other hand, seems to fall below a critical mass of characteristics to qualify as communication. Although the "sender" may execute certain actions to blend in with its surroundings, the function is to *fail* to induce a response from certain other individuals, that is, to avoid there being a receiver.

1.2 *Channels of Communication*

A particular **channel** of communication between two individuals can be distinguished on the basis of the kind of actions that the sender executes and the kind of sensory modality that the receiver uses to perceive them. The medium of the channel must be able to support the transmission of the stimuli from the sender's actions to the receiver's relevant sense organ.

1.2.1 Types of Communication Channels

Communication channels can, for example, be tactile, chemical, or electrical. But in the particular communication channel we focus on in this study, the actions that the sender executes consist of vocalizations; the medium supports the passage of sound through it; and the sensory modality that the receiver perceives in is that of hearing. This will be called the **vocal-auditory channel**, or "VAC".

Another communication channel will serve as our main reference point for comparison. In it, the actions executed by the sender consist of visible movements of particular body parts in certain configurations (other than for vocalization); the medium supports the passage of light through it; and the sensory modality that the receiver perceives in is that of sight. This will be called the **somatic-visual channel**, or "SVC".

These two communication channels, the vocal-auditory and the somatic-visual, thus comprise two different pairings of production and perception modalities. Contrasting them is instructive because, as will be argued, the three limitations cited for the VAC are exhibited to a substantially lesser degree in the SVC (see Talmy 2003 for more on the differences)

1.2.2 Settings for Communication Channels

In part separately from the preceding issue of what constitutes them, communication channels can differ with respect to their settings along certain scales. Such scales include those of conceptuality, informativeness, volitionality, and immediacy. Our focus here is on channels set at the high end of these four scales. Such channels, that is, are conceptual, informative, volitional, and immediate. In a "civi" channel of this sort, the two individuals act sentiently, and what is communicated is conceptual content.

In a civi channel, the communicative process begins with the sender coming to experience certain conceptual content. She wants the receiver to experience the same content. To bring this about, she voluntarily executes certain actions. These actions produce sensory stimuli that travel through the medium and immediately reach the receiver. As his response, the receiver perceives and cognitively processes these stimuli. This processing leads to an experience of conceptual content approximately the same as that experienced by the sender.

Paralleling traditional linguistic usage, the sender's actions or the stimuli they produce can be said to "represent" the conceptual content experienced by the sender or the receiver.

We can help clarify the character of civi scale settings in a communication channel through contrasts with different settings. First, an example of a channel with a nonconceptual setting is that cited earlier in which a plant attacked by insects releases certain chemicals that induce a nearby plant to produce deterrent chemicals.

A channel has a noninformative setting if the sender intends the receiver to generate conceptual content that is not the same as the original, but differs in a specific way. An example is the case of threat, where the sender experiences aggressiveness but intends the receiver to experience fear. Another example is the case of surprise, where a speaking sender experiences a particular phenomenon as already known but intends the receiver to experience it as unanticipatedly novel.

A channel has a nonvolitional setting if, say, a (nonhuman) sender produces an involuntary alarm call or display as a signal of danger, and the receiver's processing of this leads to reflexive flight.

And a channel has a nonimmediate setting where the sender's actions occur substantially before the receiver perceives stimuli from them, for example, in the cases of scent marking and writing.

1.3 *Conceptual Content*

Several observations can be made about the conceptual content in a communication channel in which the setting is conceptual rather than nonconceptual.

First, conceptual content is here understood to consist not only of intellective material but of experiential material of any sort, including that of affect.

Second, although conceptual content is here and below often characterized in terms of the conduit metaphor as being "sent" by one individual, "received" by another individual, and "transmitted" from the one to the other, only the physical stimuli actually follow a path of this sort. The conceptual content per se does not move. Rather, it is separately generated internally within each individual through cognitive processing of two correspondingly different types. In the sender, a particular portion of conceptual content can arise in consciousness through such cognitive processes as a free flow of thought or a reaction to something perceived. In the receiver, roughly the same portion of conceptual content arises in consciousness through the cognitive process of interpreting a communication from the sender.

Third, given its form of organization in a particular species, a conceptual communication channel can transmit conceptual content at a certain **level of advancement** per unit of time. Conceptual content can be assessed as being at a more advanced level to the degree that it includes concepts that are finer-grained, more complex, or more abstract; that belong to a greater diversity of categories or participate in more complex interrelations with each other; or that undergo more complex processing in ideation or thought. As the species evolves, that communication channel exhibits **enhancement** if it becomes able to accommodate the transmission of a more advanced level of conceptual content per unit of time.

1.4 *Combinance in Synopsis*

We present a synopsis of combinance here and plan to detail it in a later article. A preliminary component of combinance is **discretization**, in which an originally gradient aspect of cognition becomes at least in part discretized into **units**, or in which new discrete units appear in cognition beside the gradients there.

In combinance at its most basic, then, smaller units combine to form larger units. In any single case of such combining, the two sizes of units are treated as occurring on different adjacent tiers. The smaller units are on the **source tier** and the larger units are on the **goal tier**. The process of combining will here be generally represented as if moving upward, so that the source tier is the lower tier, while the goal tier is the higher tier.

The main type of such adjacent tier-to-tier combinance consists of a certain bounded sequence of steps. This main type will be called a **clave** (described further below). This new term is derived from the word "claviature", a keyboard, to suggest another bounded sequence of steps: the piano keys from one note to the note one octave higher.

What we will call the **general clave** is the combinance sequence whose steps are characterized in their most generic form and are realized in their maximal number. In turn, a **specific clave**—there are many different ones of these—is a particular instance of the general clave. Its steps are particularized for it, and it may not employ all the steps of the general clave.

While a clave exhibits combinance internally with its own combining of smaller into larger units, claves themselves can in turn function as smaller units that combine to form larger units. Three types of such combinance among claves can be identified: concatenation, association, and alignment.

In concatenation, the goal tier of one specific clave becomes the source tier of another, and so on in a succession. Five types of concatenation can be identified. Of these, the longest is the **mainline concatenation** that, to simplify, progresses from phonetic features to an interchange between speakers. Maximally, seven claves can link up to form this mainline concatenation. Most of these claves begin and end with an **inventory**—a relatively closed set of units—and include a step of **arrangement**.

We can sketch out the mainline concatenation as it might proceed in a particular language. The source tier of the initial clave has an inventory of phonetic features that, in accord with the rules of feature combination (its step of arrangement), combine to form an inventory of phonemes, the "phonemic inventory", at the goal tier. This goal tier becomes the source tier of the second clave. The phonemes of the inventory there, under the rules of phonotactics, combine to form an inventory of morphemes, the "morpheme lexicon", at the new goal tier. Next, the morphemes of this inventory, now at the source tier of the third clave, under the rules of morphology, combine to form an inventory of multimorphemic words, the "multimorphemic-word lexicon", at the new goal tier. In turn, the morphemes and multimorphemic words of their respective inventories, now at the source tier of the fourth clave, under certain rules of syntax, combine to form an inventory of idioms, the "idiom lexicon", at the new goal tier. Then the morphemes, multimorphemic words, and idioms from their respective inventories, now composing the "total lexicon" at the source tier of the fifth clave, under the rules of syntax, combine to form expressions at the new goal tier.

Neither these expressions nor subsequent combinations of them form inventories. But on a given occasion, a speaker can execute a sixth clave, whose source tier consists of expressions that, under the rules of discourse, combine to form what might be called a **monolog** at the new goal tier. In turn, to or more speakers can execute a seventh clave whose monologs in the source tier, under the rules of turn-taking, combine to form an interchange at the new goal tier.

Besides this mainline concatenation, there is also a semantic concatenation whose combinations proceed from semantic components to the concepts associated with individual morphemes or idioms, and on to the compositional meanings of multimorphemic words, expressions, monologs, and interchanges. Further, there is a grammatical concatenation that can be understood as progressing from elementary grammatical features to the grammatical complexes associated with individual morphemes, and on to morphological and syntactic structures and constructions.

We can further posit a fourth and a fifth concatenation (other analyses might include them within the mainline concatenation). In the fourth, a succession of claves proceeds from stress features to syllabic stresses, from there to stress patterns over multisyllabic words, and from there on to stress patterns over expressions. And the fifth case might be analyzed as a single clave or a two-clave concatenation that proceeds from pitch features to intonation patterns over expressions.

We had stated that concatenation is just one type of inter-clave combinance, that is, where whole claves are combined. A second type of combinance, as noted, is association. In this type, a unit from a particular clave in one concatenation combines as a whole with a unit from a particular clave in another concatenation to form a larger cross-concatenation unit. The main instance of such association occurs between the mainline and the semantic concatenations. To characterize it, though, we must emend an earlier simplified description. The clave within the mainline concatenation in which phonemes are combined actually yields morphemic *shapes*, not morphemes, in its goal tier inventory. In the simplest case, each such morphemic shape then as a whole becomes associated with a particular whole concept. This concept is from the goal tier inventory of a particular clave in the semantic concatenation, the clave that begins with semantic components.

A further case of association is between the mainline and the grammatical concatenations. Each morphemic shape also becomes associated with a particular complex of grammatical features. This complex is from the clave that begins with individual grammatical features. The association of a morphemic shape, a concept, and a grammatical complex then yields a genuine morpheme, This is a larger unit, formed by the combination of units from three different concatenations. It is actually post-association units of this type that compose the goal-tier inventory earlier called the morpheme lexicon.

In the third type of combinance, alignment, a whole unit from one concatenation again becomes associated with a whole unit from another concatenation, but in addition, the smaller units that make up the two larger units must become associated with each other in a particular way. For one instance of this type, a particular intonation pattern–a unit from the top inventory of the pitch concatenation—first becomes associated as a whole with a particular whole sentence—itself a unit within the clave of the mainline concatenation that yields expressions. But in addition, the particular ups and downs that make up the intonation pattern must align with particular constituents within the sentence.

To address now the steps that make up a single general clave internally from its source tier to its goal tier, we will use the specific phonemes-to-morphemes clave in English to both explicate and illustrate. The source tier here consists of the relatively closed inventory of English phonemes. A heuristic and schematic analysis then holds that, from this starting point, a succession of six steps occurs. The first five steps lead to the formation of a goal-tier unit, and the sixth step leads to the formation of the goal-tier inventory.

The first step is that of **selection**, itself consisting of several successive operations. An initial operation of **singling out** picks one of the phonemes, for example /s/, and places a copy of it on the goal tier.

An operation of **iteration**, in conjunction with the condition of **alternativity**, can repeat the singling out operation, not necessarily with the same phoneme, and here might, say, also select the phonemes /æ/ and /k/.

An operation of **limiting** ensures that the number of times that the iteration operation occurs tends to fall within a certain approximate range. In the present case, the maximum may tend to be some dozen (e.g., the eleven phonemes in *hippopotamus*).

The second step is that of **arrangement**, which organizes the selected phonemes. The initial operation here is **clustering** which, for this specific clave in English, ensures that the selected phonemes will occur adjacently (without intervening material) and sequentially (not concurrently).

Then an operation specifically of **arranging** orders the selected phonemes in a particular sequence in accordance with this language's rules of phonotactics. This operation works in conjunction with the condition of alternativity as well, which permits it to order the same phonemes in different sequences. The result of any instance of such arranging is a potential morphemic shape. With respect to the three phonemes selected in our illustration so far, the sequences phonotactically allowed are /æsk/, /æks/, /sæk/, and /kæs/, while those disallowed are /skæ/ and /ksæ/. The allowed sequences are then potential morphemic shapes for the language.

The third step is that of **licensing**, which grants a potential linguistic entity the status of actually being part of the language. Specifically here, only some of the morphemic shapes potential for a language are licensed at any given historical point, while the rest are not. Of the four potential morphemic shapes in our illustration, the first three are licensed and recognizable with the spellings "ask", "axe", and "sack". But the fourth, perhaps spellable as "cass", is not.

The fourth step is that of association which, as seen, is the combining of goal tier units from different concatenations. Here, a licensed morphemic shape becomes associated with a concept from the relevant clave of the semantic concatenation, and also with a set of grammatical properties from the relevant clave of the grammatical concatenation. The larger unit formed by the combination of these three smaller units is a "morpheme". In our illustration, the licensed morphemic shape /æks/, "axe", becomes associated with the concept 'tool with handle and sharp-edged metal head for chopping', and with a grammatical complex characterizable as "free (unbound) count noun taking third-person singular neuter agreements, and suffixable by the pluralizing bound morpheme *-s*".

Where a fifth step of **emergence** applies, a goal-tier unit acquires its own independent identity apart from the arranged source-tier units that compose it. In the present phonemes-to-morphemes clave, this step amounts to a type of symbol formation, one in which a particular selection and arrangement of phonemes arbitrarily comes to represent a particular concept and to have certain formal properties. Thus, in our example, the phoneme sequence /æks/ has no lawful relation to the concept and grammatical complex it is associated with—the association exists through convention, learned afresh by each new speaker, and so has a basically autonomous status.

The sixth step is that of **populating**, in which the progression of the first five steps repeats, forming new units that fill out (populate) the inventory of the goal tier. Some of the operations seen in the first step occur here as well. Thus, the operation of iteration repeats the processes that lead to the formation of a morpheme. Iteration here again acts in conjunction with the condition of alternativity, so that the new morphemes are distinct from each other. And the operation of limiting again ensures that the number of times that iteration occurs tends to fall within a certain approximate range, leading this time to the relative closedness of the inventory. In this case, though, the number is in the thousands. If the associations of a single morphemic shape with different homonymous or polysemous senses and with different grammatical properties are treated as distinct morphemes, the maximum will be substantially higher but not, say, in the millions. The relatively closed inventory that results from step six of the present clave is the morpheme lexicon.

2 Extent of Parameter Diversity

For any given communication channel, the term **parameter** is here used to refer to any aspect of a sender's actions, or of the stimuli they produce, that has the capacity to represent some conceptual content independently of other such aspects. That channel can then exhibit a particular **extent of parameter diversity**. This extent in turn rests on five **diversity factors**.

One factor is **number**: the sheer number of distinct independently variable parameters of this sort that are present. A second factor is **expressivity**: the degree to which these parameters not only have the capacity to convey conceptual content but in fact are in service doing so. A third factor is **specialization**: the degree to which different expressive parameters represent different categories of conceptual content, rather than the same overall body of content. A fourth factor is **cooccurrence**: the degree to which the different expressive parameters can be produced concurrently and perceived concurrently, rather than only sequentially.[3] And the fifth parameter is **balance**: the degree to which the different expressive parameters exhibit comparable amounts of use in actual practice, rather than some occurring often and others only seldom.

Greater parameter diversity would seem to have a certain communicative advantage over lesser diversity in a communication channel that has not incorporated combinance as a compensation. This apparent advantage is that more conceptual content can be transmitted more clearly in the same amount of time. The reasons are first that the greater the number of different parameters expressing conceptual content concurrently, the greater the amount of content that is expressed per unit of time. And the more specialized the parameters are for different categories of conceptual content, the less the parameters have the potential for cross-interference, hence the clearer the expression of the content.

Evidence will next be presented that the vocal-auditory channel has a much lesser extent of parameter diversity than the somatic-visual channel—lesser, in fact, for all five factors. Our proposal is that the incorporation of combinance into the VAC compensated for this limitation on parameter diversity and enabled the development of spoken language. By contrast, the greater extent of parameter diversity in the SVC enabled the development of a language based on signs that uses only a minimum of combinance in at least one of its major subsystems.

2.1 *The Greater Extent of Parameter Diversity in Signed Language*

We begin the contrast by looking at the greater extent of parameter diversity in the somatic-visual channel—specifically in signed language, which uses this channel—to demonstrate as a proof of principle that much conceptual content can be communicated without much combinance.[4]

As background, we note that apparently every signed language has two main subsystems for representing conceptual content (e.g., Emmorey 2003, Liddell 2003, Morgan and Woll 2007). These two systems are distinct from each other and can each make up the whole of any single expression, but they can also occur together within a single expression.

One is the lexical subsystem, which includes lexical signs (resembling the open-class, or lexical, forms of spoken language) and their modulations (resembling certain closed-class, or grammatical, forms of spoken language). The other is the so-called "classifier" subsystem, which schematically represents objects or substances moving, located, or oriented with respect to each other in space and time. We here designate this target of representation simply as "Entity Motion" (capitalized).

The lexical subsystem is closer to spoken language in the degree of combinance that it incorporates. But the classifier subsystem, which has no spoken-language counterpart, exhibits little combinance. We here focus on the classifier subsystem for its capacity to communicate much conceptual content with little combinance through its high ranking on the five factors of parameter diversity.[5]

The classifier subsystem of American Sign Language (ASL) is at the high end of all five factors for parameter diversity. With regard to the first factor of number, it has, by the analysis proposed here, as many as thirty independently varying parameters that are able to represent conceptual content. Regarding factor 2 for expressivity, all thirty of these parameters not only are able to represent conceptual content but actively do so within classifier expressions. Regarding factor 3 for specialization, these thirty different parameters do not overlap in what they communicate, but in fact represent thirty different categories of information that may be present in a particular physical event of Entity Motion.

Regarding factor 4 for cooccurrence, while certain constraints and simplifications do exist, a sizable number of the parameters can be represented concurrently in a single classifier expression. We can note that the different parameters are realized by distinct gestural complexes—that is, certain positions, configurations, and movements of particular parts of the body. These complexes do not interfere with each other when manifested—they at most adapt to each other—and this noninterference enables the parameters to cooccur. Finally, regarding the fifth factor of balance, it too appears to be at a high level. Although some parameters are always or often used across classifier expressions, the remaining parameters seemingly are not neglected and exhibit comparable levels of usage.

To expand on these factors, we will shortly list the thirty parameters, demonstrating their great number, expressivity, and specialization (factors 1, 2, and 3). But we begin with an illustration that demonstrates the great extent of their cooccurrence and balance (factors 4 and 5).

Thus, a spatiotemporal event that English might express as *The car drove past the tree* could be expressed in the ASL classifier subsystem as follows: The signer's dominant hand, used to represent the Figure object, here has a "3 handshape" (index and middle fingers extended forward, thumb up) to represent a land vehicle. The non-dominant hand, used to represent the Ground object, here involves an upright "5 handshape" (forearm held upright with the five fingers extended upward and spread apart) to represent a tree. The dominant hand is moved horizontally past the nondominant forearm.

Further though, this basic form could be modified or augmented to concurrently represent additional characteristics of the path and manner. For example, the dominant hand could move along a curved path to indicate that the road being followed was curved, it could slant upward to represent an uphill course, or both could be shown together. As for manner, as the dominant hand moves along, it could for example oscillate up and down to indicate a bumpy ride, or advance quickly to indicate a swift pace, or both might be shown together, as well as with the preceding two path properties.

Still further, the dominant hand can concurrently show additional relationships of the Figure to the Ground. For example, it could pass nearer to (or farther from) the nondominant hand to indicate the car's lesser (or greater) distance from the tree when passing it. Or it could make the approach toward the nondominant hand longer (or shorter) than the trailing portion of the path to represent the corresponding lengths of the car's path toward and away from the tree. Or it could show both of these together or, indeed, with all the preceding additional characteristics. The single expression resulting in this last case, itself wholly within the classifier subsystem, demonstrates the high degree of cooccurrence (factor 3) possible in that subsystem: it concurrently incorporates as many as ten distinct parameters, each representing a distinct category of conceptual content. For an immediate comparison (analyzed further several times below), the English sentence in (1) represents roughly the same total conceptual content as this full classifier expression.

(1) The car quickly bounced some distance uphill curving close past a tree
and a shorter distance beyond it.

With regard to the fifth factor of balance, it appears that all thirty of the parameters indicated in table 1 below for ASL are in regular active use for conveying concepts in classifier expressions. In the present illustration, For example, ten of the parameters from throughout the list are in use.

Thus, the configuration of the dominant hand indicates that the Figure's type is a land vehicle [parameter 1]. The configuration of the nondominant hand indicates that the Ground's type is a tree [2]. The translational movement of the dominant hand indicates that the motive state of the Figure is that of motion [10]. The curvature of the dominant hand's path indicates that the contour of the Figure's path is curved [15]. The upward slant of the dominant hand's path indicates that the elevation of the Figure's path is at an upward slant [21]. The up-down bobbing of the dominant hand indicates that the divertive manner is that of bouncing [24]. The swiftness of the dominant hand's advance indicates that the dynamic manner is that of great speed [25]. The dominant hand's passing to one side of the nondominant hand indicates that the spatial conformation is one of passing [26]. The longer and shorter distances traversed by the dominant hand before and after passing the nondominant hand indicates the relative distances traveled by the car before and after passing the tree [27]. And the dominant hand's closeness to the nondominant hand while passing it indicates that the Figure is proximal to the Ground [29].

Table 1 in (2), then, presents our provisional analysis of the ASL classifier subsystem as having thirty distinct parameters—in turn grouped into eight classes—demonstrating its high ranking on the first factor of number. Further, its parameters all actively serve to express substantive conceptual content, demonstrating its high ranking on the second factor of expressivity. And still further, its parameters all express different categories of conceptual content—specifically, distinct aspects of spatiotemporal information—demonstrating its high ranking on the fourth factor of specialization.

The guiding principle for this analysis has been that the different parameters must be represented by different gestural complexes, all independently variable. This principle thus leads to certain separations and clusterings of the conceptual categories that get represented.

With regard to separation, if two categories of spatiotemporal information are represented by different gestural complexes, they are here assigned to different parameters and can generally be independently expressed together in the same classifier expression. For example, although an analysis of English would generally class them together as "Manner", our analysis of ASL distinguishes the "divertive manner" of parameter 24 (e.g. moving along with an up-down bumping oscillation) from the "dynamic manner" of parameter 25 (e.g., moving along rapidly) because they can be independently indicated and often concurrently so.

But with regard to clustering, if two or more different categories of spatiotemporal information are represented by the same gestural complex, then they belong to the same parameter, are mutually exclusive, and cannot both be represented independently at the same location in a classifier expression. For example, the conceptual categories represented by a Figure, by an instrument that acts on a Figure, and by a manipulator (say, a person's hand controlling an instrument or a Figure) are distinct and can in fact vary independently of each other across different events. Nevertheless, all three of these categories—as well as that represented by the so-called "index classifier"—are represented solely by the configuration of the dominant hand, which thus cannot indicate the categories' full range of independent variations. Accordingly, all four of these categories are grouped together under parameter 1.

Note that the table does not specify the thirty gestural complexes per se, but rather indicates the corresponding thirty categories of spatiotemporal information that the complexes represent. However, most of the different gestural complexes can be inferred because they are iconic of the spatial category they represent (see section 3). Thus, parameter 6 for a Figure object's pitch is represented by the up-down angle at which the signer's hand is positioned relative to the forearm. Still, the gestural complexes that represent certain of the spatial categories do need to be separately specified because they are less iconic or noniconic of those categories. An example is parameter 2, in which different types of Ground objects are represented by the nondominant hand in configurations not necessarily iconic of them.

(2) Table 1. Proposed set of parameters in the classifier subsystem of signed language

A. entity properties

1. type of Figure/instrument/manipulator (handling classifier) / or pointer (index classifier)

2. type of Ground
3. magnitude of an entity in some dimension
   e.g., a 'pizzalike shape' can be shown as 'small' / 'medium' / 'large'
   by the amount of separation between the two hands
4. magnitude of an entity in a transverse dimension
   e.g., any of the preceding pizzalike shapes, made only with thumb and index finger,
   can show depth by adding a second finger or the remaining 3 fingers (as for a bowl)
5. number of entities
   e.g., 1, 2, 3, etc. raised fingers can be used to represent 1, 2, 3, etc. people arriving

B. orientation properties: the angle at which an entity is rotated relative to a canonical position
6. an entity's rotatedness about its left-right axis ("pitch")
7. an entity's rotatedness about its front-back axis ("roll")
8. a. an entity's rotatedness about its top-bottom axis ("yaw")
   b. an entity's rotatedness relative to its path of forward motion
      e.g., vehicle classifier moved forward with its "front" leading / trailing / to one side to represent, say, a car moving forwards / backwards / sideways

C. locus properties
9. locus within sign space

D. Motion properties
10. motive state: moving / stationary
11. internal motion
    Components of a Figure/instrument/manipulator are shown moving relative to each other, as in:
    dilation (expansion / contraction), change of form, twisting, wriggling, swirling
12. confined motion
    Figure/instrument/manipulator moves as intact whole within confined region without overall change of location, as with:
    straight oscillation, rotary oscillation, rotation, local wander
13. translational motion
    Figure/instrument/manipulator moves as intact whole through space with overall change of location

E. Path properties
14. state of path's continuity: continuous / saltatory
15. contour of path
16. state of boundedness: unbounded / bounded at start / bounded at stop / bounded at both ends
17. path length and whether a path's beginning/end is in/out of view
18. vertical height
19. horizontal distance in front of signer
20. left-right positioning

21. up-down angle ("elevation")
22. left-right angle ("direction")
23. transitions between motion and stationariness
    These distinctions can apparently be made: a Figure's—
    stopping normally / slowing to a stop / stopping abruptly, as from impact
    being placed at a point of support / being given into someone's grasp

F. viewpoint properties
An observer's perspective on the depicted scene is largely indicated by E. Path properties
e.g., a path directed forward vs. right-to-left suggests a view from behind vs. from the side.

G. Manner properties
24. divertive manner
    a movement the Figure makes during and in addition to a forward path motion, e.g., bouncing
25. dynamic manner: speed

H. relations of Figure or Path to Ground
26. path's conformation relative to Ground—e.g., past it, above it, into it
27. relative lengths of path before and after encounter with Ground
28. Figure's path relative to the path of a moving Ground
    Usually Ground object is stationary; but can also show Ground moving along a path
        e.g., for Figure pursuing / catching up with / passing it
29. Figure's proximity to Ground
30. Figure's orientation relative to Ground

2.2 *The Lesser Extent of Parameter Diversity in Spoken Language*

The vocal-auditory channel of spoken language exhibits a far lesser extent of parameter diversity than does the somatic-visual channel of signed language. And it does so for all five of the diversity factors, addressed next in order.

2.2.1 Lesser on Factor 1 for Number

First, the VAC appears to rank relatively low on factor 1 for number. The distinct independently variable parameters in this channel with the capacity to represent conceptual content would seem to be on the order of only eight in number. The proposed parameters in this comparatively small set are listed in table 2 in (3).[6]

(3) Table 2: Proposed set of parameters in spoken language
1. phonetic quality
2. pitch (from low to high)
3. volume (from soft to loud)
4. timbre
5. precision (from loose approximation to sharp clarity)
6. rate (from slow to fast
7. duration (from brief to lengthy, e.g., length of sound, word, interword spacing)
8. consecutivity (sequence of different elements or iteration of the same element)

2.2.2 Lesser on Factors 2 and 3 for Expressivity and Specialization

In principle, all these VAC parameters have the capacity to represent conceptual content (factor 2) and to specialize in different categories of such content (factor 3), without any contribution from combinance. But at least parameters 2 through 8 fall far short on that capacity.

To provide a reference point for how such expressivity and specialization could have been realized, we here devise two hypothetical situations.

First, different values for pitch, volume, timbre, precision, rate, duration, and iteration (consecutivity) might have respectively represented a referent's elevation, horizontal distance away, weight, hardness, speed, permanence, and width. However, for reasons not fully clear, such substantive and distinctive representations of conceptual content are difficult to find (though see section 3 regarding timbre and iteration).

While these VAC parameters apparently fail to represent the preceding type of concepts in spoken language, they do seem able to represent some conceptual content pertaining to the speaker's state of mind. But even here, the parameters rank low on factor 2 for expressivity since the content they represent is limited and vague, and they rank low on factor 3 for specialization since different parameters can represent similar content without differentiation. For a sense of how there could have been a higher ranking on these two factors, we present a second hypothetical situation. A rise in pitch [parameter 2] might have been interpreted as representing a decrease in the speaker's self-confidence; an increase in volume [3] as representing a shift from mildness to aggresivity; an increase in rate [5] as representing a shift from calmness to excitement; an increase in precision and duration [6, 7] as representing an increase in seriousness about the message; and an increase in iteration [8] as representing an increase in a sense of urgency. However, any actual representations by these parameters are neither as clear nor as differentiated as presented here.

We can now turn to the first parameter of phonetic quality. Within the vocal-auditory channel, at least as used by primates, this first parameter may well be the one out of the eight with the greatest expressivity (factor 2) and the most specialization (factor 3). But without combinance, its use may be limited to a comparatively small number of holophrastic forms. An example may be the roughly two dozen calls of vervet monkeys (Seyfarth et al., 1980). These represent portions of conceptual content such as warnings about a snake on the ground or a raptor in the air. One seeming property of all these calls is that the conceptual content they represent always includes an indication of proximal deixis—that is, they do not refer to an entity or event removed from the caller and call in space or time. Thus, the just cited warnings are of a snake or raptor present here and now.

Possibly analogous to such holophrastic calls in human spoken language are interjections, which share their limitation in number a proximal deixis. Among English speakers, for instance, examples of such interjections with their possible approximate conceptual content are: *ow* 'I'm feeling a sharp pain right now'; *oh* 'This is new and somewhat surprising to me' or 'Now I understand this matter; *hmm* 'That's an interesting idea, I'm thinking about it'; *brrr* 'I'm feeling cold now'; *whew* 'I'm feeling relief now that this danger or arduous task has just ended' *hlyup* [produced during inhalation] 'This seems delicious just thinking about it now'. But though comparatively expressive and specialized, the parameter of phonetic quality here yields comparatively few holophrastic interjections of this sort.

2.2.3 Lesser on Factors 4 and 5 for Cooccurrence and Balance

With regard to factor 4 for cooccurrence, it is true that all eight of the proposed VAC parameters can be manifested concurrently while retaining their independent variability. This can be straightforwardly observed in spoken language. Thus, as a speaker uses phonetic quality [parameter 1] to communicate the majority of some conceptual content, she can at the same time talk higher or lower [2], louder or softer [3], with a range of voice qualities [4], with greater or lesser precision [5], faster or slower [6], with different amounts of spacing between words [7], and with more or fewer repetitions [8].

Comparably with respect to the fifth factor of balance, the preceding depiction of a person speaking shows that all eight of the proposed parameters can be in operation. In fact, they *must* all be in operation since, given the way the parameters were characterized, some value on each of them must be manifest.

However, in the introduction to section 2, factors 4 and 5 for cooccurrence and balance were specifically formulated to apply only to expressive parameters, that is, to parameters ranking high on factor 2 for representing substantial conceptual content. But as argued above, all eight VAC parameters ran comparatively low on expressivity without combinance and so do not qualify for a high ranking on factors 4 and 5.

2.2.4 Relating Combinance to Parameter Diversity

As just seen, the vocal-auditory channel has a low extent of parameter diversity, ranking low on all five diversity factors, and thereby has a lesser capacity to represent conceptual content. By our proposal, though, this limitation was overcome by the incorporation of combinance Into the VAC as spoken language evolved.

Now, much as the whole vocal-auditory channel can be assessed with respect to the five factors of parameter diversity, the use of combinance within that channel can be so assessed as well. The central observation is that combinance ranks very high on factor 2 for expressivity since it enabled an enormous increase in the representation of conceptual content. But it ranks low on the other factors.

Lesser on Factors 1 and 5 for Number and Balance

Consider first its ranking on factor 1 for number. Combinance, a system in which smaller units combine to form larger units, has mainly become associated with just one of the eight VAC parameters, parameter 1 for phonetic quality, where it has given rise to phonemes, morphemes, words and expressions. Accordingly, the sheer number of parameters affected by it—its ranking on Factor 1—is effectively only one.

Correspondingly, the balance with which combinance exploits the full range of available parameters—its ranking on factor 5—is at a minimum.

To be sure, some combinance can be, and often is, incorporated into other parameters. For example, the parameter of pitch [2] can be used for discrete tones on a syllable or for discrete intonation contours over an expression. The parameter of volume [3] can be used for discrete levels of stress on a syllable or for discrete stress patterns over an expression. The parameter of duration [7] can be used for discrete degrees of segment length. And the parameter of consecutivity [8] in the form of iteration can be used as a grammatical device to represent the concept of multiplexing. But such use of combinance with other parameters is small by comparison with its use with the first parameter. This supplementary use only slightly increases the effective number of combinant parameters beyond one and the effective balance of combinant parameters beyond the minimum.

Lesser on Factors 3 and 4 for Specialization and Cooccurrence

We can next observe that combinance in spoken language also ranks very low on factors 3 and 4 for specialization and cooccurrence. Consider again the sentence in (1). The ASL classifier expression that represented the conceptual contents of that sentence showed great degrees of cooccurrence and specialization. It included ten different parameters manifested concurrently, each representing a different category of Entity Motion information. These were the Figure's and Ground's identities as land vehicle and as tree; the Figure's state of being in motion; the path's curve and incline; the bouncy and rapid manners; and the Figure's lengthier, close, and bypassing approach to the Ground.

By stark contrast, the English sentence represented most of the conceptual content of the same event using just parameter 1 for phonetic quality. Only a bit of the total content was concurrently represented by pitch [2] for intonation, and by volume [3] for stress. Thus, unlike the signed language case, most of the conceptual content was represented not concurrently by many parameters but sequentially by a single parameter. And unlike the signed language case, all ten different categories of Entity Motion information present in the event were represented not by distinct parameters, each specialized for one of those categories, but by the single parameter of phonetic quality functioning generically for all the categories. This pattern resulted from the action of combinance on the parameter of phonetic quality, turning it in turn, as just noted, into phonemes, morphemes, words, and the whole expression.

2.3 *Unavailable Parameters*

Tables 1 and 2 listed only parameters actually available to their respective channels. But to fill out the analytic framework, we should consider parameters that are unavailable to a channel, but that in some respect are close candidates. We next address in order unpaired and blocked parameters.

2.3.1 Unpaired Parameters

One basis for a parameter's unavailability to a joint production-perception channel is that it is available to one of that channel's poles but not to the other.

In the Somatic-Visual Channel

Consider first the somatic-visual channel. Vision by itself can discern two phenomena, colors and textures, which can there be considered to be two additional independently varying parameters. But gesturally produced configurations and movements have no direct counterparts for representing these phenomena. Accordingly, such phenomena are absent as parameters in the joint somatic-visual channel.[7]

In a complementary way, gestural activity by itself has certain additional parameters available to it. One of these is the degree of pressure that a body part exerts on a contacted object. This may in fact be a parameter actually in use in a third channel of communication, the somatic-somatic—mainly manual-manual—communication system of the deaf-blind. But vision cannot directly discern such degrees of pressure—it can at most discern only concomitants of such pressures. Accordingly, this parameter too is absent from the joint channel.

In the Vocal-Auditory Channel

We can comparably cite unmatched parameters in the vocal-auditory channel. For example, auditory perception by itself can determine the locations and paths of sound-emitting objects. But vocal production is fixed in place, and so cannot directly represent such phenomena. A speaker cannot genuinely “throw his voice”, in the manner often attributed to ventriloquists, so that the hearer can perceive a vocally projected point’s location or path in space. Accordingly, such a spatial parameter is absent from the joint vocal-auditory channel, and is unavailable for representing conceptual content there.

As another example at the auditory pole, audition by itself can identify a wide range of nonvocal sounds, including the rustle of leaves and thunder claps. But such sounds cannot be produced vocally—except perhaps by sound effects specialists. Accordingly, such sounds can also not be included as a parameter on the joint vocal-auditory list.

What about the complementary case of an unmatched vocal parameter? One might cite movements of the vocal tract that do not produce sounds (as when unaccompanied by breath) and hence cannot be perceived auditorily. Any parameter comprising such movements would also be off the joint vocal-auditory list, unavailable for communicating conceptual content.[8]

2.3.2 Blocked Parameters

Another possible basis for the unavailability of a parameter is that it is somehow blocked from occurring. Such blockage might result from constraints on neuromuscular control as it has evolved in humans. For example, the vocal-auditory channel lacks a parameter involving the direction of breath—exhaled vs. inhaled—during the production of a phonetic sequence. Though such a distinction could be heard and might be exploited to convey conceptual content, it is blocked from regular vocal use presumably for physiological reasons.

As for the somatic-visual channel, we suspect that sign languages universally lack an independent parameter based on the angle at which the signer’s elbows are held—from some degrees above horizontal, to horizontally out, to some degrees below horizontal, to straight down, to pinched inward. Apart from cases where a particular angle is required so as to represent a spatial factor iconically, such elbow orientations could have been exploited to represent a particular category of conceptual content. Such angles could certainly be seen, and it seems that they could be readily produced motorically. So the reason for their apparent blockage is unclear—but blocked they nevertheless appear to be.

2.4 *Summary and Interpretation*

The extent of parameter diversity appears to be substantially greater in the somatic-visual channel than in the vocal-auditory channel. And it is greater in the signed language system, especially the classifier subsystem, than in the spoken language system that respectively came to use those channels. This imbalance is evident in all five diversity factors.

First, the extent of parameter diversity is greater with respect to the sheer number of independently variable parameters available to represent conceptual content (factor 1)—some 30 as against some eight.

It is greater with respect to the expressivity of these parameters (factor 2). The SVC parameters are all in service representing substantial conceptual content. But the First VAC parameter of phonetic quality by itself provides the bulk of such expressivity while, perhaps puzzlingly, the remaining parameters generally represent little content.

It is greater with respect to the degree of specialization in the types of conceptual content represented by the expressive parameters (factor 3). All of the SVC’s parameters express different categories of information. But the first VAC parameter of phonetic quality can express all types of information, while the remaining parameters seemingly tend to represent mainly some vague and not clearly differentiated states of mind in the speaker.

It is greater with respect to the extent to which different parameters with substantial expressivity can be realized concurrently and so represent more information in parallel (factor 4). A sizable number of SVC parameters can cooccur in a single expression. But since the first VAC parameter of phonetic quality represents most of the information in a spoken communication, little additional information is represented concurrently, and this first-parameter information is mainly presented sequentially.

Finally, it is greater with respect to a balanced distribution of information load among the parameters (factor 5). All thirty SVC parameters have their share in representing conceptual content across signed communications. But once again, the first VAC parameter of phonetic quality is seemingly called on more than the remaining parameters combined to represent conceptual content across spoken communications.

Our proposal, then, is that this limitation on the vocal-auditory channel due to its lesser extent of parameter diversity was overcome by its incorporation of combinance as spoken language evolved. This combinance, though, became associated mainly with the first parameter of phonetic quality—yielding phonemes, morphemes, words, and expressions—and only modestly with some of the remaining parameters (again perhaps puzzlingly). If combinance is considered by itself with respect to extent of parameter diversity, it thus ranks high on factor 2 for expressivity, but low on the remaining parameters—even lower than the VAC as a whole does.

The classifier subsystem of signed language uses a minimum of combinance. But its great extent of parameter diversity enables it to communicate conceptual content of great complexity—in our terms, a more advanced level of conceptual content—both quickly and clearly. It thus demonstrates the cognitive feasibility of advanced communication with little combinance. If human language had evolved using the somatic-visual channel, it might not have needed much increase in combinance as compensation.

However, human language did evolve in the vocal-auditory channel with its basically low extent of parameter diversity. Some compensation for that low extent was necessary if spoken language was to be able to represent advanced conceptual content quickly and clearly, and our proposal is that the extensive incorporation of combinance served that function.

It is not clear why language evolved in a way that combinance became associated preponderantly with the parameter of phonetic quality and comparatively little with the other parameters as a system for representing conceptual content. The puzzle increases under the following consideration. In their possibly original and basic character, parameter 1 is greatly, and parameters 2 through 7 are wholly, gradient, together providing what might be called **vocal dynamics**. This gradience may have constituted a more ancient system and have been carried over as the core discrete combinant system of language evolved. But music, a cognitive system that coevolved with language in the human lineage, does extensively structure the parameters of pitch and duration [2, 7] under a discrete combinant system of organization (see Talmy 2015 for further discussion). Music thus shows the cognitive feasibility of combinance applying ramifiedly to parameters little affected by it in language. Nevertheless, for whatever reasons, as spoken language evolved, it settled on the first parameter of phonetic quality as the main substrate in which to establish combinant organization based on discrete units to represent conceptual content.

The diversity difference between signed and spoken language can in part be framed in terms of a combination of specialization and balance (factors 3 and 5). We might say that signed language employs **parameter spread**, while spoken language employs **parameter concentration**. Put another way, signed language uses a wide range of means to represent different categories of conceptual content, whereas spoken language uses a single means to represent almost all such categories of content.

The question arises why this difference in specialization and balance occurs. Perhaps the tendency of a channel is that it will distribute its profile of use over virtually all the parameters available to it. Evidence for such a tendency may lie in signed language itself, whose classifier subsystem exhibits just this sort of distribution. After all, this subsystem might in theory have been able to represent all the same information about Entity Motion with only a subset of the parameters available to it. The fact that it uses the full parameter complement then supports the notion of a distributional tendency.

Spoken language then breaks with this tendency. Despite the counterarguments presented above, possible reasons for this might be that the range of parameters in their original state was less suited to representing conceptual content, and that combinance, whose incorporation compensated for this insufficiency, could most readily function with just one of the parameters, that of phonetic quality.

3 Extent of Iconicity

A given communication channel can include **iconicity**, a certain type of representation of conceptual content that we characterize in section 3.1. That channel can then exhibit a particular **extent of iconicity**. This extent in turn rests on four **iconicity factors** that might be related to some of the five diversity factors (section 2) but that are here treated in their own right.

One factor is **number**, which can have two interpretations. By one interpretation, it is the number of different types of iconicity that can be distinguished in the channel. But since there may not be a definitive way to mark such distinctions, we here adopt a second interpretation by which this factor is the number of parameters in the channel that exhibit iconicity, even if some parameter is judged to exhibit more than one type. A second factor is **strength**. As will be seen in the next section, we analyze iconicity as having different degrees of strength, and greater strength contributes to a channel's having a greater extent of iconicity.

A third factor is **relevance** to communication. A form of iconicity is more relevant if it pertains to areas of conceptual content that occur in communication more frequently, more pervasively, or more ramifiedly (for reasons that themselves can be separately examined), and it is less relevant otherwise. For example, the motion or location of objects relative to each other in space can be taken to be generally more relevant in communication than, say, the temperature of those objects.[9] And the fourth factor is **use**: the degree to which a type of iconicity is actually in service representing conceptual content.

A channel's extent of iconicity can thus be more specifically characterized here as the number of parameters in the channel available for exploiting iconicity to represent conceptual content (Factor 1), to the degree that that iconicity is strong (factor 2), relevant (factor 3), and in actual use (factor 4).

Greater iconicity would seem to have certain communicative advantages over lesser iconicity in a communication channel that has not incorporated combinance as a compensation.

Though empirical determination will be needed, we here propose four such advantages that respectively involve accessibility, cognitive load, learnability, and universality.

First, the cognitive processes needed for representing conceptual content through iconicity may be more accessible or available in cognition—and may have been so for a longer evolutionary period—than those needed for representing conceptual content through a combinant system of arbitrary forms including symbols. Second, the cognitive load may be less for processing iconicity than for processing such arbitrary forms. The latter entails the long-term establishment of a system relating arbitrary forms to concepts; the immediate encoding of certain concepts into such forms by a speaker; and the immediate decoding of such forms into concepts by a hearer. Third, the use of a channel for communication may be more readily acquired by an individual—either as a first or as a secondary system—where its type of representation consists of iconicity rather than of arbitrary form-concept relations. And fourth, the representation of concepts through the use of iconicity may be more similar across different instantiations of a particular communication channel, i.e., across different "languages"—that is, it may be more universal—than that through the use of arbitrary forms.

Sections 3.2 and 3.3 present evidence respectively that signed language within the somatic-visual channel has a great extent of iconicity and that spoken language within the vocal-auditory channel has a comparatively far lesser extent. The greater iconicity in the SVC enabled the development of a language based on signs that uses only a minimum of combinance in at least one of its major subsystems. But the incorporation of combinance into the VAC, even at some cognitive cost, compensated for this limitation on iconicity and enabled the development of spoken language.

3.1 *Characterizing Iconicity*

We begin the analysis by characterizing iconicity. Iconicity can be generally understood as an asymmetric relationship in which one quantity represents another through similarity with it. In our analysis, such representation consists of a certain sequence: when the first quantity is perceived by an individual, it leads her to perceive or to conceptualize the second quantity. In the discussion that follows, we will term these two quantities respectively the "Form" and the "Phenomenon" (capitalized). Iconicity is analyzed here as being graduated in strength rather than as having an all-or-none character. In particular, we propose that iconicity rests on six different properties, itemized below, and that its strength in a given application correlates with the number of properties in effect there.

Work on iconicity in language has a long history, as in Peirce (1931/1958), Bühler (1934), and Haiman (1985), while work on it in signed language specifically is more recent, as in Taub (2001) and Perniss et al. (2010). But the present analysis is distinctive in certain respects. First, we treat iconic representation in cognitive terms as an effect that a form has on a perceiver's processing, instead of building it into the form itself as in such traditional formulations as that the form is similar to its own meaning. Second, we treat iconicity as graduated in strength based on how many of six properties it exhibits. Third, we focus on less addressed areas of iconicity, including categories of object motion, the clustering of such categories, and the object/action distinction.

3.1.1 Prorepresentation

We here apply the term **prorepresentation** to the one property seemingly criterial to the occurrence of iconicity. As usual here, a particular Form and a particular Phenomenon must be present. By this property, the Form is iconic of the Phenomenon if two conditions are met. First, the Form must be similar to the Phenomenon. We here interpret such similarity to mean that the Form is identical with the Phenomenon in some particular respect but not in the remaining respects. Second, the Form must represent the Phenomenon precisely by virtue of this particular respect of identity it has with it. The Form does not represent the Phenomenon because some convention has established the correlation by making the Form an arbitrary symbol of the Phenomenon.

As prelude to an illustration, we first note that the word *way*, as used in (4), represents the concept ‘at a great distance’. But it is not iconic of it.

(4) The cell phone tower is way over there.

But now notice that when the word is pronounced with a greater vowel duration—manifesting parameter 7 of table 2 in (3)—as in (5), it no longer represents the original concept but a related one: ‘at a very great distance’. The increase in vowel duration in this morpheme represents an increase in the greatness of distance in the concept that the morpheme expresses. Now on the one hand, these two increases do differ: the former increase is in the temporal duration of a produced sound while the latter is in the spatial length of a conceived distance. But on the other hand, the two are identical in a particular respect: they both constitute extra magnitude along a scalar dimension. Moreover, the former increase represents the latter by virtue of this identity. Accordingly, the former increase can be treated as the Form and the latter increase as the Phenomenon. And this Form is iconic of that Phenomenon through the property of prorepresentation—it is identical to it in a certain respect and represents it through this identity.[10]

(5) The cell phone tower is waaay over there.

3.1.2 Covariation

A Form is more iconic of a Phenomenon if, in addition to minimally exhibiting prorepresentation with it, it further exhibits the property of **covariation** with it. By this property, both the Form and the Phenomenon have two or more variants, and each Form variant can be paired with one of the Phenomenon variants.

To illustrate, consider the three alternatives in (6) in which the word *way* has three different increases in its vowel. These can be regarded as three variants of the same Form seen earlier, namely, increase in the temporal duration of a produced vowel. These variants of the Form can be paired up one-to-one with three variants of the same Phenomenon seen earlier, that is, with three different increases in the greatness of a conceptualized distance. Accordingly, the three Form variants are iconic of the three Phenomenon variants with respect both to prorepresentation and to covariation.

(6) The cell phone tower is waay / waaaay / waaaaaay over there.

3.1.3 Proportionality

A Form is still more iconic of a Phenomenon if, in addition to exhibiting prorepresentation and covariation with it, it further exhibits the property of **proportionality** with it. Proportionality is present if the variants of a Form that one-to-one represent the variants of a Phenomenon under the property of covariation are not just distinct from each other, but can be ranked along some quantitative dimension in a sequence of correspondences.

For example, the three different increases in vowel duration seen in the words *waay / waaaay / waaaaaay* are not just independently and qualitatively distinct variants of the Form consisting of increase in vowel duration. Rather, the three variants can be ordered—in the sequence they were just listed in—in accordance with the magnitude of their extra vowel duration. In the same way, the three variants of the conceptual Phenomenon that these Form variants represent can also be ordered in accordance with the magnitude of the extra greatness of distance that they constitute. Further, these two sets of ordered variants correspond to each other one-to-one. That is, they covary proportionally along the quantitative dimensions that they are ranked on.

3.1.4 Proportional Directness

A Form is even more iconic of a Phenomenon if, in addition to exhibiting prorepresentation, covariation, and proportionality with it, it further exhibits the property of **proportional directness** with it. This additional property is present if the ordering of the Form variants and the ordering of the Phenomenon variants—just discussed under proportionality—either both grow larger or both grow smaller along their respective rankings in correspondence with each other. That is, the covariation of the two ordered sets is here directly proportional, not inversely proportional. We deliberately worded our characterization of the property of proportionality in section 3.1.3 to be neutral to whether the direction of correspondence between the two sets was direct or inverse.

This new property of proportional directness in fact held for the example in section 3.1.3. Progressively greater vowel duration in the three Form variants—that is, a progressive augmentation in the increases of their vowel duration—corresponded directly to a progressively greater distance in the three Phenomenon variants. It did not correspond inversely to a progressively smaller distance.

3.1.5 Cogranularity

As just seen, the property of proportionality (section 3.1.3) can in turn be either direct or inverse (section 3.1.4). In this section, we note that that property can exhibit a further type of dichotomy. The variants that a Form displays can either be gradient—that is, fall along a continuum—or be discrete. And the Phenomenon can independently exhibit the same dichotomy. We can thus now add that where a Form exhibits at least prorepresentation, covariation, and proportionality—whether direct or inverse—with a Phenomenon, it is more iconic of that Phenomenon if it further exhibits the property of **cogranularity** with it. A Form and a Phenomenon exhibit this property if they are both gradient—hence, **cogradient**—or if they are both discrete—hence, **codiscrete**. They do not exhibit the property if one is discrete and the other is gradient.

The example in section 3.1.3 did in fact exhibit cogranularity, specifically of the cogradient type. The vowel duration of the word *way* could be increased in a continuous manner, and the greatness represented in the conceptual Phenomenon at a great distance' could also be increased in a continuous manner.

By contrast, the otherwise comparable example in (7) lacks cogranularity, though exhibiting all the same properties leading up to it. In this new example, the Form is the adjacent iteration of a word—manifesting parameter 8′ of table 2—in this case, that of the morpheme *up*. The Phenomenon that the Form represents is an increase in the concept of the length of an upward path. The Form here has three variants—a single, a double, and a triple iteration. And these respectively represent three variants of the Phenomenon—three progressively greater increases in the conception of upward path length—with prorepresentative, covariant, proportional, and proportionally direct forms of iconicity. But increase in the number of iterations is discrete in character, while increase in upward path length is continuous. Hence, this "up" example is not cogranular, whereas the "way" example was cogranular and thus exhibited a greater degree of iconicity.

(7) The bird flew up up / up up up / up up up up and away.

3.1.6 Codomainality

Finally, a Form can exhibit the property of **codomainality** with a Phenomenon and, if it also exhibits the other five properties, it is iconic of the Phenomenon to the strongest degree. This property is present if both the Form and the Phenomenon (as it is either perceived or conceptualized) are in the same qualitative domain.

The example in section 3.1.3 did not exhibit codomainality. The domains of the Form and of the Phenomenon differed in being respectively temporal and spatial, specifically involving vowel prolongation and the extension of great distance. Thus, the Form in that example was prorepresentative, covariant, proportional, proportionally direct, and cogranular (specifically, cogradient) with the Phenomenon, but not codomainal with it.

But consider a new example. Manifesting parameter 3 of table 2, the Form here consists of loudness in the pronunciation of a word like *shout* or *loud* itself, as in (8). The Phenomenon represented by this Form is then an increase in the loudness of another sonic event, the one being referred to. The Form is thus in the same domain as the Phenomenon, that is, it is codomainal with it. It may further be the case that a gradient increase in such a Form's loudness represents a proportional gradient increase in the Phenomenon's loudness. If so, this Form exhibits all six properties with respect to the Phenomenon, and is thus iconic of it to the strongest degree.[11]

(9) He really SHOUTED at us. / The music was just too LOUD.

3.2 ***The Greater Extent of Iconicity in Signed Language***

We look first at signed language—specifically, at the classifier subsystem of ASL. It will be seen to exhibit a great extent of iconicity that in fact covers three areas, treated next in order.

3.2.1 Iconic Representation of Entity Motion Categories

The first of the three areas of this iconicity is the main one. In it, the gestural complexes associated with the different parameters in table 1 represent correspondingly different categories of conceptual content by virtue of particular similarities that they have with them. As seen, these categories pertain to Entity Motion, that is, objects or substances moving, located, or oriented with respect to each other in space and time. In accord with our analysis, the gestural complex constitutes the Form, the category constitutes the Phenomenon, and the Form exhibits at least the property of prorepresentation with the Phenomenon.

American Sign Language can then be seen to exhibit a great extent of this area of iconicity, in fact ranking high on all four iconicity factors. First, in terms of sheer number (factor 1), of the thirty independent parameters listed in Table 1 for the ASL classifier subsystem, the gestural complexes for all but the first two (addressed below) are iconic of the conceptual content they represent. These are then 28 different types of iconicity—clearly a large number.

Just to provide some sense of this extent, we here cite some of the parameters in which a moving dominant hand can iconically represent a moving Figure. The up-down angle of the moving hand relative to its forearm is iconic of the Figure's up-down angle relative to its path [parameter 6]. The contour of the hand's path is iconic of the contour of the Figure's path [15]. The length of the hand's path is iconic of the length of the Figure's path [17]. The height of the hand's path is iconic of the height of the Figure's path [18]. The speed of the hand's path is iconic of the speed of the Figure's path [25]. And the proximity of the hand's path to the nondominant hand is iconic of the proximity of the Figure's path to the ground [29].

Next, regarding the strength of iconicity (factor 2), the 28 parameters largely exhibit iconicity in its greatest strength, that is, with all six properties of section 6.1 realized. We here illustrate this for parameter 17 for path length.

First, the criterial minimum of prorepresentation (property 1) is realized. Here, the Form is the length of the path that the hand moves along, while the Phenomenon is the length of the path that the Figure moves along. And the Form represents the Phenomenon precisely because of a particular respect that it shares identically with it: path length.

Covariation (property 2) is also realized. In any given context, the hand is capable of moving along paths of different lengths and thereby represents different Figure path lengths. But it is not the case that, say, different hand path lengths would all represent a single Figure path length.

Proportionality (property 3) is in effect as well. The different hand path lengths just cited would be understood to be ordered along a scalar dimension, as would the different Figure path lengths. If there were three different lengths, they would rank as small, medium and large. The small, medium, and large hand path lengths would then respectively represent the small or large, the medium, and the large or small Figure path lengths. But it is not the case that, say, the small and medium hand path lengths could respectively represent the medium and small Figure path lengths.

In addition, proportional directness (property 4) is realized. If the hand path length just cited is increased from small to large, it would correspondingly represent an increased Figure path length from small to large. It is not the case that an increasing hand path length represents a decreasing Figure path length. That is, this parameter is not inversely proportional.

Further, cogranularity (property 5), specifically of the cogradient type, is basically in effect. Down to some tolerance, a continuous increase in the length of the hand's path represents a continuous increase in the length of the Figure's path. It is not the case that, say, a continuous increase in hand path length represents the Figure's path increasing by discrete jumps.

Finally, codomainality (property 6) is also realized. The domain of the Form consists of a physical object moving along a path of a certain length through space and time, and the Form represents a Phenomenon of the same domain. Thus, greater length of the hand's path represents greater length of the Figure's path. It is not the case that it represents, say, greater circularity of the Figure's path nor, for that matter, greater beauty of the Figure.

With regard to the relevance of the iconicity that is available (factor 3), all 28 parameters would seem to rank high in that respect as well. They all represent different aspects of objects or substances located, moving, or oriented with respect to each other in space and time, a domain already noted as being of great relevance in human communication.

And finally, regarding the degree to which the available types of iconicity are in actual use (factor 4), recall the finding in section 2.1 that the ASL classifier subsystem ranks high on diversity factor 5 for balance. Since all thirty classifier parameters are accordingly in substantial use, the types of iconicity exhibited by the 28 parameters at issue here are also in substantial use and so rank high on that factor.

To return to parameters 1 and 2 of Table 1, they apply to the handshapes respectively representing the Figure/instrument/manipulator /index and representing the Ground. In some cases, such a handshape does exhibit the criterial iconic property of prorepresentation with the object it indicates. In such cases, the handshape is identical with the object in some respect and represents it thereby, so that they respectively constitute a Form and a Phenomenon. Such prorepresentation can be quite clear, as with a flat hand representing a planar object. Or it can be suggestive, as with the forearm, hand, and fingers held vertically upright, fingers spread, to represent a tree. Beyond this point, though, what handshapes exhibit starts shading off from iconic prorepresentation into arbitrary symbolism.

3.2.2 Iconic Representation of Category Clustering

The classifier subsystem of signed language exhibits iconicity not only in the area just addressed—different conceptual categories of Entity Motion—but also in a second area: how those conceptual categories are clustered.

Consider again the classifier expression described in section 2.1, referring to a car passing a tree, that concurrently realized some ten different parameters and roughly corresponded to the English sentence in (1). The nondominant hand there realized only parameter 2, representing just one conceptual category within the whole Entity Motion ensemble, namely the classified Ground object. But by concurrently manifesting a set of distinguishable gestural complexes, the dominant hand realized all nine remaining parameters, representing the remaining conceptual categories of the Entity Motion ensemble. In particular, the dominant hand represented the classified figure object, its state of motion, path properties, manners, and relations to the Ground object. Thus, the representations of these nine categories of conceptual content were clustered together—the dominant hand indicated them all—and they were separated from the representation of the Ground object by the nondominant hand.

But an observer of the scene would perceive exactly the same pattern of clustering. He would see the same single object presenting as a specific type of entity, enacting motion, describing a path with certain properties, exhibiting certain manners, and relating in certain ways to a Ground object. And he would perceive the Ground object as separate from the preceding features,

This classifier expression is thus identical to the Entity Motion ensemble with respect to the clustering of conceptual categories and can be regarded as a Form representing a Phenomenon through that aspect of identity. Hence, the classifier expression is iconic of the Entity Motion ensemble in this second area as well. The particular iconic properties present are prorepresentation; cogranularity, specifically of the codiscreet type; and codomainality, where the shared domain is that of one physical Entity manifesting a particular set of identity and Motion characteristics. Since the dominant hand would represent all the Figural characteristics chosen for indication in any classifier expression, no role is played by the property of covariation, and hence none by proportionality or proportional directness. Still, the number of Figure characteristics that the dominant hand can iconically represent is large. Thus, the overall strength of iconicity in this area of clustering can be regarded as substantial.

3.2.3 Iconic Representation of the Entity-Motion Distinction

The gestural complexes for all thirty parameters in table 1 are iconic of the conceptual categories they represent with respect to yet a third area. This area pertains to the distinction between an Entity and its Motion—that is between material and its behavior. As characterized in section 2.1, an "Entity" is an object or substance, while an Entity's "Motion" is its location, motion, or orientation in space and time. The observation, then, is that the Entity portion of every gestural complex—that is, a body part, typically a hand—represents just the Entity portion of the whole Entity Motion ensemble that is being referred to. At the same time, the Motion portion of the gestural complex—that is, the location, motion, or orientation of the body part—represents the corresponding Motion portion of the whole Entity Motion ensemble in reference. For one example, the extended forefinger of the dominant hand can be used to represent a pencil as a Figure, while the finger's location, motion, or orientation can respectively represent the location, motion, or orientation of the pencil.

By our analysis, a gestural complex is identical to an Entity Motion ensemble with respect to the Entity-Motion distinction, and can be regarded as a Form iconically representing a Phenomenon through that aspect of identity. The particular iconic properties present are prorepresentation; cogranularity, specifically of the codiscreet type, consisting of the correspondences between Entities, between Motions, and between the subdivisions of motion; and codomainality, where the shared domain is that of Entity Motion. As in the second area of iconicity, no role is played here by the iconicity properties of covariation, proportionality, or proportional directness. But the representation of the Entity-Motion distinction is so thoroughgoing that the strength of iconicity in this third area can also be regarded as quite great.

This thoroughgoingness can be highlighted by considering alternative arrangements that might have existed, such as one in which an Entity represents Motion and Motion represents an Entity. For example, a signer's hand through its static configuration might represent the site or path of a Figure—say, a fist shape to represent a stationary location, the outstretched fingers held flat together to represent a straight line path, the fingers in a curved plane for a curved path, and the fingers alternately forward and backward for a zigzag path. At the same time, the motion of the hand might represent the static shape of the Figure—say, the hand moving in a circle to represent a round Figure and in a straight line to represent a linear Figure. However, such alternatives do not occur in the classifier subsystem. [12]

3.3 *The Lesser Extent of Iconicity in Spoken Language*

Spoken language has a far lesser extent of iconicity than that just seen for signed language. First, as treated in section 3.3.1, it exhibits iconicity in none of the three areas of it illustrated for signed language. Second, as treated in section 3.3.2, its eight parameters have little relevant iconicity in actual use.

3.3.1 No Iconicity in the Three Cited Areas

We look first at the three areas of iconicity addressed in sections 3.2.1, 3.2.2, and 3.2.3 for signed language.

Entity Motion Categories

It was seen in section 3.2.1 that the ASL classifier subsystem exhibits extensive iconicity in its representation of Entity Motion categories. But spoken language exhibits iconicity for none of these. For illustration, consider again the English sentence in (1) that roughly corresponded to the classifier expression described nearby. There, say, the word *quickly* was not iconic of great speed (unlike the corresponding gestural complex of ASL parameter [25]); the word *uphill* was not iconic of an upward slant (unlike [21]); the word *curving* was not iconic of a curved contour (unlike [15]); the word *close* was not iconic of proximity (unlike [29]); and the word *past* was not iconic of one object passing another (unlike [26]).

To look more closely at the last instance, the ASL gestural complex for 'past' consisted of the dominant Figure hand progressing from the nearer side of the nondominant Ground arm to a point beside it and then on to its further side, much like the path progression one would see on viewing an actual car passing a tree. By contrast, the English word *past* is an indivisible lexical unit, a morpheme, whose form does not represent motion involving passing ahead in space.

Category Clustering

It was seen in section 3.2.2 that the ASL classifier subsystem also exhibits extensive iconicity in its representation of category clustering. Thus, it represents all the Motion and identity characteristics that a Figure object manifests with the same single articulator, the dominant hand, and all the Ground object characteristics with the nondominant hand, much as those characteristics are separately clustered in one's perception of the actual objects.

But all spoken languages diverge in a variety of respects from this iconicity. The English sentence in (1) can be used again for illustration. In it, for example, one syntactic constituent separated off by itself, the subject nominal *the car*, represented the Figure object and its identity. If the expression *quickly bounced some distance uphill* can be regarded as a single syntactic constituent, a verb complex, then it clustered together the representation of a dynamic manner in an adverb (*quickly*), a divertive manner in the verb (*bounced*), a path length in an adverbial noun phrase (*some distance*), and a path pitch in an adverb (*uphill*). And if the expression *close past a tree* can be regarded as a single syntactic constituent, a prepositional phrase, then it clustered together the representation of proximity in an adverb (*close*), a path conformation in a preposition (*past*), and the Ground object in a noun phrase (*a tree*). It in fact remains a mystery why all spoken languages using a preposition-like constituent to indicate a path conformation always conjoin it with the Ground nominal and basically never with the Figure nominal[13], even though the Figure is what executes the path, and is so represented in the classifier subsystem of signed language.

The Entity Motion Distinction

It was finally seen in section 3.2.3 that the ASL classifier subsystem also exhibits extensive iconicity in its representation of the Entity Motion distinction. That is, the distinction between a physical articulator and its location, motion, or orientation corresponds to the distinction in the situation being represented between a physical object or substance and its location, motion, or orientation.

To check for this distinction in spoken language, the counterpart of an Entity might be interpreted as being a morpheme, while that of its Motion might be interpreted liberally as the morpheme's location in a sentence or, more liberally, as its lexical category or syntactic function. But attempts to find iconicity under this interpretation fail both across and within languages. Thus, while material objects are prototypically represented by nouns in English, they are instead prototypically represented by verb roots in Atsugewi. And while path configurations are prototypically represented by prepositions and satellites in English, they are instead prototypically represented by verb roots in Spanish (see Talmy 2000b, chapter 1, for both these observations).

Even within a single language like English, no correspondence holds consistently. For example, the sentence in (9a) might suggest that nouns represent Entities—the Figure Entity toward the beginning of the sentence and the Ground Entity at the end—while a verb and a preposition represent aspects of Motion, respectively a manner and a path configuration. Now, in some circumstances, the sentence in (9b) provides the same information as (9a). Yet in it, the car as an Entity is represented by the verb *drive* toward the end of the sentence; the manner is represented by the adverb *bumpily*; and the path configuration is represented by the verb *pass* at the beginning of the sentence. This is quite a different representation of an Entity-Motion distinction than that in (9a). The conclusion once again is that, in this third area as well, iconicity is much lower in spoken than in signed language.

(9) a. My car bounced past a tree.
b. Passing a tree, I drove along bumpily.

3.3.2 Little Relevant Iconicity in Use among the Eight Parameters

While the eight parameters listed in table 2 for the vocal-auditory channel are not iconic of Entity Motion categories in spoken language, as just seen, they may well still manifest distinct types of iconicity, as itemized below. But that is still only a count of eight different types (iconicity factor 1 for number), far less than the 28 seen for signed language. Further, although some of these types of iconicity are great in strength (factor 2), most of them are of little communicative relevance (Factor 3), and two of them that are relevant are virtually unused (factor 4). Only parameter 8 is both relevant and in use.

Of Low Relevance

The types of iconicity offered by most of the parameters in table 2 are of low relevance to communication either because the Phenomenon that a parameter's Form represents is of limited occurrence or because speakers infrequently experience the need to express it.

Thus, one type of iconicity offered by phonetic quality [parameter 1] is onomatopoeia. The Form here is the particular selection and sequencing of the sounds in a morpheme, for example, those in the noun *moo*, and the Phenomenon is a sound in the external environment, in this case, one made by a cow. Such a Form exhibits the iconic property of prorepresentation insofar as it is taken to represent the Phenomenon through similarity with it, as well as the property of codomainality, since both the Form and the Phenomenon are sonic. But one basis for low relevance here is that the realm of sounds that are phonetically imitable is comparatively small.

The pitch [parameter 2] of a constituent in speech can be iconic of the expression's referent with the greatest strength. An example is the use of high pitch on the word *squeaked* in the sentence *He was so nervous, he squeaked during his answer*. The relevance of this type of iconicity, though, is again apparently low since the occasions on which a speaker might want to communicate about a pitch level would seem to be limited.

The situation is comparable for the parameter of volume [3]. For its main type of iconicity, seen in the "shout" example of (8), the Form is greater loudness of pronunciation, and the Phenomenon is greater loudness of a sound in the environment. This Form can exhibit the strongest degree of iconicity with the Phenomenon. But a main basis for low relevance here is that the realm of loud sounds that speakers frequently want to refer to is again comparatively small.

Next consider the parameter of timbre [4]. A speaker can use timbre iconically by adopting a particular timbre in her own voice to represent the timbre of another person's voice. This is again a strong form of iconicity. It is prorepresentative, covariant, directly proportional, cogradient, and codomainal. This type of iconicity is indeed used on occasion in the narrative style of some speakers or by storytellers to represent different characters. However, it is not of great relevance to communication—speakers generally do not feel a need to represent others' voices frequently, pervasively, and ramifiedly.

Perhaps the parameter of precision [5] can also exhibit iconicity. For example, a speaker might tend to enunciate a sentence like *All the steps must be followed in exact order*, while carelessly saying a sentence like *You can do anything you want*. If so, though, the effect would seem to be moderate and little called on.

Finally, as the *way* example in (6) showed, the property of duration [7], when applied to vowel prolongation, can be used as a Form to represent a Phenomenon of extra magnitude along certain scalar dimensions with all the properties of iconicity up to, but shy of, codomainality. Again, though, this type of iconicity seems of low relevance since speakers' need to express such extra magnitude appears limited.

Relevant but in Little Use

Two of the temporal parameters cited for the vocal-auditory channel of spoken language—rate [6] and duration [7]—are both available to serve as Forms that iconically represent the respective Phenomena of rate and duration, when these Phenomena are in reference. They would in fact be strongly iconic of these Phenomena. It could be further argued that they would be high in communicative relevance, since language users do seem to want to represent these temporal concepts extensively. Curiously, however, they are in little use. This dearth of use can be shown by attempted examples that do employ the parameters iconically but that do not represent normal speech.

The use of parameter 6 for rate is illustrated first in (10). The three phrases in the sentence refer to events that take place with successively higher speeds—from static to slow to fast. And these phrases are here meant to be uttered with successively greater speeds—from very slow as indicated by the spaced lettering, to moderate as indicated by the normal Roman font, to fast as indicated by italics. However, iconicity of this sort is in minimal use.

(10) attempted example of using vocal rate to represent rate of motion
T h e   p e n   l a y   o n   t h e   t a b l e, rolled to the edge,
and fell down.

Comparably, parameter 7 for duration can be attempted for iconic use as in (11). The three phrases in this sentence refer to successive events—entering, sitting, and pulling out a letter—with two particular intervals of time between them. Suppose that, in the situation being represented, each interval is longer than any of the three events, and that the second interval is longer than the first. Correspondingly in the sentence—as the strings of dots are intended to indicate—each of the two pauses lasts longer than the time needed to utter any of the three phrases, and the second pause is longer than the first. The attempted aim here is that the differences in duration across the pauses and utterances would be iconic of the differences in duration across the intervals and events of the represented situation. However, any such use of durations in speech seems to be minimal in practice.

(11) attempted example of using duration in speech to represent duration in occurrence:
He entered .... sat down .... and pulled out her letter.

We can conjecture why spoken language, as it evolved, did not take advantage of these two types of temporal iconicity, even though they were available and communicatively relevant. The reason might be that it concentrated its representation of conceptual content on a commitment to the single parameter of phonetic quality. Thus, the meanings of the attempted rate and duration iconicity in the preceding two examples are more likely to be expressed through the use of phonetic quality in the form of phonemes, morphemes, words, and expressions, as illustrated in (12).

(12) a. The pen lay still on the table, then rolled slowly to the edge,
and then fell down in a flash.
b. He entered, waited for a while, sat down, waited for a longer while,
and pulled out her letter.

Relevant and in Use

The parameter of consecutivity [8], specifically when realized as the sequencing of different elements, is perhaps the one parameter that exhibits significant iconicity with respect to factors 3 and 4 for relevance and use. In particular, the temporal order in which certain constituents occur in an expression can be iconic of the temporal order in which their referents occur in the world. This iconicity is in fact of the strongest degree, manifesting all six iconicity properties. Thus, it is covariant, proportional, and proportionally direct (properties 2, 3, and 4) in that different numbers of ordered constituents will represent the same number of referents in the same order (not, say, in reverse order); it is cogranular (property 5), specifically codiscrete, in that both sequences are made up of discrete elements; and it is codomainal in that both the Form and the Phenomenon consist of temporal ordering.

An example of this sequential type of iconicity was seen in the "letter" example in (11), where the sequence of the three verb phrases corresponded to the sequence of the events they represented. The same holds for the sequence of the three satellites in (13), which corresponds to the sequence in which my hand follows three different portions of a path.

(13) I reached down around behind the clothes hamper for the vacuum cleaner.

However, such sequential iconicity is often not required or in some cases not constructionally feasible. For example, the basic construction for representing a path sequence with prepositional phrases places the starting point at the end of the sentence, as in (14).

(14) The toilet is down the corridor around the corner from here.

Comparably, the basic form of a complex sentence with the subordinating conjunction *after* requires that the later event be represented before the earlier event, as in (15).

(15) I went home after I stopped at the store.

It can thus be seen that even in this one instance where a type of iconicity, the representation of sequence, is both relevant and in some use in spoken language, such use can be neglected or blocked.

3.4 *Summary and Interpretation*

We have argued that the extent of iconicity is far greater in signed language using the somatic-visual channel than in spoken language using the vocal-auditory channel. The classifier subsystem of signed language is in fact greater in three different areas of iconicity, those pertaining to the categories of Entity Motion, the clustering of such categories, and the division between Entity and Motion. And it ranks high on all four iconicity factors. Thus in the first area, ASL ranks high on factor 1 for number in that 28 of its 30 parameters represent different types of iconicity; on factor 2 for strength in that most of these parameters exhibit all six properties of iconicity; on factor 3 for relevance in that all 30 parameters represent aspects of the communicatively relevant domain of objects or substances located, moving, or oriented with respect to each other in space and time; and on factor 4 for use in that all 30 parameters are in substantial service. By contrast, spoken language exhibits iconicity in none of these three areas.

Spoken language does have the capacity to manifest different types of iconicity—ones unrelated to those seen in signed language for Entity Motion—in the eight parameters cited for the vocal-auditory channel. But these types of iconicity add up only to a low number of eight (factor 1); most of them are of little relevance (factor 3); and of the three with high relevance, two are virtually unused (factor 4).

The classifier subsystem of signed language provides an existence proof that a communication system within the somatic-visual channel can represent a great deal of advanced conceptual content through iconicity. But as spoken language evolved in the vocal-auditory channel, it could not take advantage of iconicity for this function. Here again, though, the incorporation of combinance enabled spoken language to overcome a major limitation. Even if at some cognitive cost, noniconic combinance—with its use of combinations of arbitrary forms to represent conceptual content—has performed the roles played by iconicity in a somatic-visual communication system.

4 Degree of Fidelity

Specifically directed research and theory will be needed for a notion of fidelity to be contrastively applied to different channels in use to communicate conceptual content. But we here suggest one such notion. Granting that a sender can produce a signal to a certain degree of fine-grainedness at a certain rate in a particular channel, the degree of fidelity in that channel is the degree of accuracy with which the receiver can perceive the signal. "Accuracy" here is understood to combine four factors. The receiver's accuracy is greater to the degree that she is able to identify more components of the signal at a finer granularity with greater correctness and with greater certainty. For both the sender and receiver, finer-grainedness is understood to consist of the production or perception of finer distinctions involving smaller components.

The advantage of greater fidelity is that it permits more conceptual content to be communicated per unit of time. The reason is that the greater the degree to which a sender's signal is fine-grained, reaches the receiver, and is accurately processed there, the greater the amount of information that can be transmitted per time unit.

We speculate that the vocal-auditory channel has a lower degree of such fidelity than the somatic-visual channel. The incorporation of combinance in the VAC would then have compensated for this proposed third limitation as well and further enabled the evolution of spoken language.

Though experimentation will be needed, our supposition is that the senders in both our channels of comparison—the signer in the SVC and the speaker in the VAC—have the cognitive processing capacity to controlledly produce signals at approximately the same degree of fine-grainedness at a comparable rate of production, and regularly do so. Comparably, the receivers in both channels—the sign viewer and the hearer—may have the cognitive processing capacity to perceive signals with approximately the same degree of accuracy. If so, the capacity of the sender to produce and the receiver to perceive a signal plays little role in the fidelity of the two channels.

We conjecture, though, that the properties of the signal and the properties of the medium in the VAC conduce to an overall lower degree of accuracy for the hearer, who thus lacks the circumstances in which to fully exercise her cognitive processing capacity for perception. If so, the VAC has a lower degree of fidelity than the SVC.

We consider first the possibility that properties of the signal in the VAC cause a hearer greater difficulty in determining signal components with accuracy. For a possible illustration, we consider a case from each channel, where a receiver tries to perceive in isolation a single signal component produced by a sender under optimal conditions of the medium. In the SVC, the viewer tries to discern the precise angle at which a signer is orienting an extended forefinger. In the VAC, the hearer tries to discern the particular nasal sound along the approximate phonetic continuum from [m] to [ŋ] that a speaker is uttering (without seeing the lips). Our conjecture is that the viewer would generally have greater accuracy (including greater certainty) in such a determination than the hearer would. Any multiplication of such less accurately perceived elements in a single VAC expression could result in a significant decrease of fidelity in the communication of the expression.

We consider next the possibility that properties of the medium in the VAC also cause a hearer greater difficulty in determining a signal with accuracy. In the SVC, the medium would seem to exhibit a high degree of fidelity if the sender's signaling body parts are sufficiently illuminated and close to the receiver's eyes without obstruction along a straight line. Fidelity would then generally seem to decrease to the degree that the illumination decreases, the distance increases, or opaque material intervenes. As for the VAC, its medium would seem to exhibit a high degree of fidelity if the sender's voice is sufficiently loud and close to the receiver's ears without much dampening or competition. Fidelity here then would generally seem to decrease to the degree that the loudness decreases, the distance increases, dampening material intervenes, or the voice is insufficiently distinct from or louder than surrounding sounds. We conjecture here that the circumstances for decreased fidelity—from a signal's being slightly suboptimal down to unusable—are more frequent in the VAC than in the SVC.

If as suggested the properties both of the signal and of the medium in the VAC conduce to lesser receiver accuracy in determining a signal and hence to lesser fidelity in that channel, this limitation may have been overcome by the incorporation of combinance as spoken language evolved. Certain features of combinance in particular would then have played a prominent role. These features are the relative closedness of the inventory, containing only a certain number of units, on a clave's source tier; the relatively fixed rules for combining units from the inventory there; the licensing of only some of those combinations; and the relative closedness of the inventory, containing only a certain number of units, on the clave's goal tier. Together, these features have the consequence that, of the larger units emerging at the end of the clave, most will be familiar to the hearer—she will expect to hear only units she knows are already members of the goal tier inventory.

To be more specific, consider any language at a particular point of its diachronic history with attention here only to the mainline concatenation. The relatively closed inventory of phonetic features undergoes fixed rules of combination and specific licensing, leading to a relatively closed inventory of a certain number of phonemes, ones not novel but already known to the hearer. This last inventory comparably leads to the inventory of known morphemes. This in turn leads to an inventory of mostly known multimorphemic words. The morpheme and multimorphemic word inventories lead to the inventory of known idioms. These last three inventories together—through the known rules of syntax—lead to expressions, though now these goal-tier units do not form a relatively closed inventory but in their basic character are nonce creations.

Thus, at almost every level at which a signal is formally organized, there is only a limited set of possible elements that the hearer can expect to perceive and, at some levels, these sets are relatively small. This arrangement constitutes an extensive system of constraints, a system inherent in combinance. Our proposal is that this system of constraints largely overcomes the posited lower fidelity in the VAC, enabling the hearer to process the signal with a greater degree of accuracy.[14]

The functioning of this system of constraints in spoken language might be evident under certain considerations. Take first the case of a speaker producing a sentence at a consistent rate of speed that, besides morphemes from the language's lexicon, includes an unknown proper name (cf. Burke et al. 2004). The hearer may well not make out the sound sequence in that name and ask for it to be repeated more slowly and distinctly, while at the same time experiencing the remaining words as having been clearly enunciated.

Since the forms all reached the hearer with the same degree of fidelity, the likely conclusion here is that this fidelity was relatively low and was in fact inadequate for the hearer to determine the sounds of the name, but was compensated for in the case of the other words by their having fallen within that language's system of combinant constraints—including their having belonged to relatively fixed and already known inventories.

In a related circumstance, a speaker about to utter a word he thinks might be rarer in occurrence or less familiar to the hearer might enunciate it more carefully (Gahl 2008, Bell et al. 2009). That is, he might pronounce it more slowly and with greater precision. A speaker might undertake this action where he sees a need to compensate for low fidelity because the combinant constraints cannot be relied on to do so.

A final consideration involves section 1.4's characterization of combinance within a clave, in particular the operation of limiting within the sixth step of populating an inventory. The largest number of goal-tier units that this operation allows occurs in the morpheme inventory, the morpheme lexicon, where it is now in the thousands. We might conjecture that the evolution of language included stages of increasingly larger lexicons. In one of the earlier stages, the lexicon may have consisted of just some dozens of units, themselves perhaps holophrastic calls. But any increase in its size might have caused a hearer difficulty in distinguishing accurately among the units—a problem of fidelity.

An incorporation of combinance may have resolved this problem through at least two of its aspects that we highlight here. First, holophrasis may have been replaced by a combinant system of component sound units, themselves fixed in character and limited to a relatively small number, hence more readily identifiable by a hearer. Second, the morphemes formed by particular selections and arrangements of these sound units were themselves fixed in their identities and limited to a specific set, again rendering them more readily identifiable by a hearer. Such increased identifiability, due to combinance, may have compensated for the fidelity problem, allowing a hearer to distinguish individual members of an increasingly large lexicon.

Endnotes

1. This article and the one planned to follow represent a substantial advance over and expansion of Talmy (2007).
2. The adjective "combinatorial" and noun "combinatoriality" may be the most familiar terms that come the closest to referring to the form of organization intended here. But the terms chosen instead, "combinant" and "combinance", have the advantages of being both shorter and freer from prior associations.
3. Factor 4 for cooccurrence plays a special role here and could have been treated as a phenomenon separate from the others, so that the subject of this section would have been better labeled as the "extent of parameter diversity and cooccurrence". But factor 4 has been included under the current aegis of the "extent of parameter diversity" because it is dependent on factor 1 for number. Factor 4 can exhibit a higher value only to the degree that factor 1 does.
4. For their personal help on a range of issues within signed language, my great thanks to Lucinda Batch, Paul Dudis, Karen Emmorey, Samuel Hawk, Nini Hoiting, Marlon Kuntze, Scott Liddell, Stephen McCullough, Wendy Sandler, Dan Slobin, Ted Suppala, Alyssa Wolf, and others—who are not responsible for my errors and oversights.
5. Besides the two cited here, signed languages have additional subsystems for representing conceptual content. One of these is the "SASS" subsystem of size and shape specifiers in which the signer schematically traces out the contours of objects (Aronoff et al., 2005). Another is a subsystem of facial expressions that largely represent the sender's attitude or modal stance toward the proposition being communicated (Emmorey 2009). Both of these subsystems, like the classifier subsystem, employ little combinance and lack spoken-language counterparts.
6. Although the first two parameters of phonetic quality [1] and pitch in table 2 may be specific to vocal production, the remaining parameters have counterparts for motoric production more generally, including that of manual signing. The original labels for parameters [3] and [4] might be changed as in (i) to represent the new generality. Those for [5] through [8] are already general enough.

   3´. intensity (generally correlating with energeticness of execution)
   4´. quality (the collective effect of slight differences in realization)

   As applied to co-speech gesturing, the parameter of precision [5] is elaborated in Talmy (2018, chapter 5).
7. With respect to texture, more precisely, what does not occur is a direct representation of static textures with static configurations of the hands. But the classifier subsystem does

have the capacity to represent some textures indirectly through fictive motion (see Talmy 2000a, ch. 2 for fictive motion). For example, a rough or stippled surface can be represented by many little pokes with the fingers (Karen Emmorey, personal communication).

8. A case comparable to this one is a receiver's purely visual perception of a sender's mouth configurations and movements—all the way to full lip reading. But to whatever degree this process is employed by receivers ranging from the hearing to the deaf, it manifests a fourth channel of communication, the **oral-visual channel**.
9. The issue of relevance also applied to parameter diversity in section 2. For example, in the ASL classifier subsystem, the representation of a Figure's path [parameter 15] might tend to be more relevant than its yaw [8]. Relevance there, though, was in effect folded in with factor 5 for balance, but is here treated in its own right.
10. Note that the Form consisting of an increase in vowel duration can also exhibit prorepresentative iconicity with respect to extra magnitude along a scalar dimension other than that of distance. For instance, this dimension can be that of size, as in *The rocket hangar is huuuge*.
11. We can point out two generalizations of the presentation so far. First, section 3.1 has illustrated iconicity using spoken language and VAC parameters from table 2 But the next section illustrates it using signed language and SVC parameters from table 1. Second, spoken language can manifest iconicity apart from the VAC parameters. One example is traditional linguistic markedness. In one realization of such markedness, the Form is the quantity of phonological segments needed to express a particular concept in a given language, from none to many. The Phenomenon represented by this Form is the degree of that concept's nonbasicness, from fully basic to peripheral (See Talmy 2000b, chapter 6, for an analysis of basicness and markedness).
12. The subsystem of size and shape specifiers (SASS's) in signed languages does permit movement of the hands to trace out an object's contours, but the hands cannot at the same time adopt a shape representing the object's path.
13. As the only apparent exception, a "demoted Figure" (see Talmy 2000b, chapter 1) can acquire either of two "demotion particles"—e.g., English *with* and *of*—that mark whether the Figure's path had a "TO" or a "FROM" vector, as seen in *The fuel tank slowly filled with gas / drained of its gas*.
14. The posited greater degree of fidelity in the SVC should, by this line of reasoning, enable greater latitude for novel gestural configurations, locations, motions, and orientations. This possibility seems immediately confirmed for the parameter of path contour [15 in table 1] but might need interpretation and testing for other parameters.

References

(All but the final Talmy references are accessible on his website: https://www.acsu.buffalo.edu/~talmy/talmy.html)

Aronoff, Mark, Irit Meir, and Wendy Sandler. 2005. The paradox of sign language morphology. *Language* 81(2): 301-344.

Babikova, Zdenka et al. 2013. Underground signals carried through common mycelial networks warn neighbouring plants of aphid attack. *Ecology Letters* 16: 835-43.

Bell, Alan, Jason M. Brenier, Michelle Gregory, Cynthia Girand, and Dan Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60 (1): 92-111.

Bühler, Karl. 1934. *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena, Germany: Gustav Fischer. In English: Theory of Language: The Representational Function of Language.

Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry pit primes Brad Pitt: Homophone priming effects on young and older adults' production of proper names. *Psychological Science* 15 (3): 164-170.

Emmorey, Karen. 2003. *Language, Cognition and the Brain: Insights from Sign Language Research*. Lawrence Erlbaum.

Emmorey, Karen. 2009. The bimodal bilingual brain: Effects of sign language experience. *Brain Language109* (2-3): 124-132.

Gahl, Susanne. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84 (3): 474-496.

Haiman, John. 1985. *Natural Syntax: Iconicity and Erosion*. Cambridge: Cambridge University Press.

Liddell, Scott. 2003. Sources of meaning in ASL classifier predicates. In K. Emmorey (ed.), *Perspectives on Classifier Constructions in Sign Language*. Mahwah, NJ: Lawrence Erlbaum.

Morgan, Gary and Bencie Woll. 2007. Understanding sign language classifiers through a polycomponential approach. *Lingua* 117 (7): 1159-1168.

Peirce, Charles Sanders. 1931-1958. Collected papers of Charles Sanders Peirce. 8 vols. Edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks. Cambridge, MA: Harvard Univ. Press.

Perniss, Pamela, Robin L. Thompson, and Gabriella Vigliocco. 2010. Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology* 1: 1-15. (https://doi.org/10.3389/fpsyg.2010.00227).

Seyfarth, Robert M., L.Cheney, Peter Marler. 1980. Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour* 28 (4): 1070-1094.

Talmy, Leonard. 2000a. *Toward a Cognitive Semantics*, volume I: *Concept Structuring Systems*. i-viii,1-565. Cambridge: MIT Press.

Talmy, Leonard. 2000b. *Toward a Cognitive Semantics*, volume II: *Typology and Process in Concept Structuring*. i-viii, 1-495. Cambridge: MIT Press.

Talmy, Leonard. 2003. The representation of spatial structure in spoken and signed language. In K. Emmorey (ed.), *Perspectives on Classifier Constructions in Sign Language*. Mahwah, NJ: Lawrence Erlbaum.

Talmy, Leonard. 2007. Recombinance in the evolution of language. In J. E. Cihlar, D. Kaiser, I. Kimbara, and A. Franklin (eds.), *Proceedings of the 39th Annual Meeting of the Chicago Linguistic Society: The Panels*. Chicago: Chicago Linguistic Society.

Talmy, Leonard. 2015. Relating language to other cognitive systems: An overview. *Cognitive Semantics* 1 (1): 1-44. Leiden-Boston: Brill.

Talmy, Leonard. 2018. *The Targeting System of Language*. Cambridge: MIT Press.

Taub, Sarah. 2001. *Language from the Body: Iconicity and Metaphor in American Sign Language*. Cambridge, UK: Cambridge University Press.