Gesture and Language:

Establishing Gesture as a Component of the Linguistic System

Ian White

Abstract

Gestures have been historically absent from formal language analyses. It was not until Kendon's work in gesture studies that the unification of these fields began. There is a clear continuum of degrees of impact that gesture can have on linguistic utterances and its endpoints are defined, ranging from simple emphasis to outright disambiguation through gestures such as pointing. A semiotic approach is taken to analyzing this interaction, taking insights from the semiotic work of De Saussure, Buhler, and Sebeok. Evidence from neuroscience about the relationship between language and gesture production is discussed. Additionally, the development of gestural systems into full languages is used as evidence for the implicit link between the two communication devices. This paper establishes that gesture is a necessary component of the linguistic system based on this evidence and details extensively the continuum upon which they interact; the linguistic-gestural system is a signaling device for communication and social cohesion.

Introduction

To what extent does gesture interact with the linguistic system? In other words, what are the
different ways a gesture can impact the meaning of given a linguistic symbol or utterance?
Formal linguistics, and linguistics more generally, tends to neglect gestures altogether, focusing
instead on the mental representations responsible for language production and understanding.
However, it is clear that this approach does not fully capture human communication. If it did,
then there would be no need for humans to use any form of gesture, yet they do – there must be
some utility to gesture if its usage is so prolific. This is not to say that linguistic studies into
human language have no place or that they are incorrect; great strides have been made in
establishing explanations for linguistic performance since Chomsky's seminal texts. This
definition of human communication is simply non-exhaustive. Thus, there is a need for a more
extensive definition that includes other features of communication. This paper argues that
gesture is an integral, if not inseparable, component of the linguistic system. Gesture and
language can be thought of as a unified signaling system for human communication, and there is
a continuum upon which gesture interacts with language to different degrees; this continuum
ranges from co-speech gestures which supply some sort of emphasis or specification to outright
disambiguation of linguistic utterances and the traceable development of gestural systems into
full languages.

There is a rich body of research in the field of "gesture studies" which has historically
existed in a separate realm from general linguistic research. However, it is only in recent decades
that the notion of gesture has gained more ground in terms of its own robustness and its
relevance to other fields. This includes establishing a concrete definition and taxonomy of
gestures in terms of what sorts of gestures there are and what sorts of information they can

encode, gesture and its relationship to culture, and gesture as a situated practice (Kendon, 2007). Crucially, though, gesture studies remains a separate discipline from linguistics, especially the formalist domain. Cognitive-functionalist approaches to linguistics are more inclusive of the notions of language as a situated practice. This certainly gives way to a conception of language, or at least communication, that could incorporate insights from gesture studies.

Many of the investigations in linguistics into the study of gesture lie in sign language research (ex: Liddell, 2003; Ruben, 2005). This connection seems obvious given the apparent similarity between manual gesture and manual languages. While there are gestures, especially in signed languages, that are not manual (non-manual gestures will be discussed later in this paper), this perceived connection is not unfounded. There is a traceable development of gestural systems into full languages: these are seen in both real-world examples such as the development of signed languages from more primitive home sign systems (Coppola & Newport, 2005) and the theoretical development in the so-called "gestural theory of language origin" or something similar to it (McNeill, 2012). This development suggests an implicit relationship between gesture and language insofar that language can arise from gesture. In spite of this clear connection, the separation of linguistics and gesture studies perhaps arises from the assumption that gesture and language at some point diverged. Instead of viewing these trajectories as separate, they should be considered to a certain extent as unified, thus establishing a singular communicative system that includes gesture and language.

That said, it is clear that gesture is not always necessary for linguistic communication. Consider a phone call, wherein two speakers cannot see each other but are able to effectively communicate with one another. Therefore, gesture cannot be a necessary component of linguistic communication. This does not immediately exclude gesture from the linguistic system, however.

As mentioned previously, gestures must emerge for some reason. Given this example of successful communication without gesture, there is an upper terminus of gesture's impact on language; it needn't always be present. What, then, are gestures for? There are myriad types of gestures, and they can accomplish different tasks (Caradec, 2018). There are gestures that are separate from language altogether: consider a student raising a hand in class to signal that they have a question. Gestures such as these are somewhat confined to the field of gesture studies (Kendon, 2007) and are not considered extensively in this paper. Instead, this paper is concerned with so-called "co-speech gestures": gestures which are produced either simultaneously or in close temporal proximity to language use (Liddell, 2003).

Co-speech gestures have a range of potential interactions with language, from some sort of emphasis to potentially disambiguating linguistic utterances. Emphasis might comprise of expanding the arms to suggest that something is important (Mendoza-Denton and Jannedy, 2011). Disambiguation refers to the capacity of gesture to explicitly define the meaning of some utterance within a discourse (Furuyama, 2000). Pointing might be a perfect example of this, as, within some discourse, a speaker is able to disambiguate a referent by means of indexing it with a pointing gesture (Fricke, 2007). Thus, there is a continuum upon which gesture impacts language to different degrees. Gesture may not be a necessary component of all linguistic signaling, but it is certainly an important one, and the limits of this interaction should be established. That is the purpose of this paper: to assert that there are co-speech gestures that exist at the lower end of this continuum, such as those that emphasize, and there is an upper bound comprised of disambiguation and the demonstrable implicit connection between language and gesture. Gesture and language constitute a unified system for communication and social cohesion.

The "Linguistic System"

The linguistic system is an arbitrary signaling device used for communication which consists of a grammar which is mutually understood by members of a linguistic community. Arbitrariness refers to the division between form and meaning; in other words, the qualities of a word in a language do not relate to the meaning of the word. While there are exceptions to this, variation between different languages' more iconic forms reveal that there is not an absolute connection between some concept in the world and the form of the word that refers to it (for example, consider that English and Japanese have different words for the sound a cat makes, or that American Sign Language and Chinese Sign Language both have signs for TREE which resemble an actual tree in two different ways). Signaling refers to the capacity to transmit concepts and relations between concepts from one individual to another. The grammar of a given linguistic system (including its lexicon, syntax, and semantics) must be understood by all members of that linguistic community so that effective communication can occur. This facet of language is known as conventionality.

      Crucially, this linguistic system can be expanded in a number of ways. Natural languages are productive, meaning that they allow for new combinations of words which have never necessarily been learned by the user. Additionally, there are so-called "suprasegmental" features of language which provide a layer of meaning to linguistic utterances, a domain which linguists refer to as pragmatics. In spoken languages, pragmatic function might manifest as a change in tone or loudness. In signed languages, larger production of a sign relative to the signing space might accomplish a similar function. All languages are, of course, reliant upon context for certain signaling; this is a fundamental feature of pragmatics. Where does the notion of pragmatics terminate? The set of suprasegmental markers are typically considered finite for a

given language. However, it is certainly possible that any additional signaling which a person engages in to convey a clear meaning constitutes pragmatic function. Perhaps the accessibility of more articulators than simply the vocal tract allows for greater emphasis or disambiguation of a given utterance. In this sense, linguistic cognition is an embodied practice. The notion of context, too, lends credence to a language user being situated in some environment which can impact linguistic production. It would be easier, for example, to describe an external room if one compared it to the room in which they were standing. It would seem, then, that pragmatic function is quite broad. With these insights in mind, it is not difficult to conceptualize a linguistic system which includes gesture as a means to facilitate linguistic signaling.

This expansion is symbolic of a departure from the somewhat contemporary linguistic tradition. Chomsky's Universal Grammar, while a useful tool for understanding the evolutionary underpinnings of human language, excludes the notion of situated linguistic cognition. This is not to say that Chomsky neglected the importance of context for linguistic performance; context was critical for his development of UG. Instead, it is to say that a linguistic system with finite boundaries within the mind is inadequate for capturing all aspects of meaningful human communication. In other words, the Chomskyan formalist language module, which is heavily concerned with the manipulation of syntactic structures in the mind, is too finite. Even a formalist linguistic approach which considers pragmatics (i.e. any formalist approach) is still unsatisfactory for considering the tools that humans employ for linguistic signaling.

This does not mean that no linguists discuss gesture and other pragmatic periphery. Cognitive linguistics is concerned with the interaction between language and cognition and denies that linguistic cognition is separate from the rest of cognition. Thus, it is concerned with

such notions as iconicity and situated cognition. Linguistics within this domain allows for consideration of a wider range of communicative activities than does formal linguistics. In fact, a growing interest in the world's signed languages partially gave way to this paradigm shift (Wilcox, 2005) (((PS5: N5))). Observations of the transitions between gestural systems to full languages is a large facet of this (McNeill, 2012), a phenomenon that is discussed later in the paper. Clearly, there is a link between language and gesture. The trouble, then, is operationalization of this connection such that linguistic analysis of the broader linguistic system can occur. Perhaps there is a "universal grammar" for this linguistic system, or perhaps not. Regardless, it is clear that the combination of language and gesture constitute a broader linguistic system.

De Saussure presented the first semiotic approach explicitly concerned with language. For De Saussure, a "sign" in a language consists of a signifier (a word) and a signified (the concept the word refers two). The signifier and signified are inseparable. Meaning arises across an utterance via the difference between two signs. Thus, a string of signs constitutes a phrase with a comprehensive meaning. While De Saussure was exclusively concerned with spoken language and thus the seriality that accompanies it, his approach might be helpful in analyzing the interaction between gesture and language. The notion of a sign as an inseparable conjunction of signifier and signified can be maintained (Innis, 1985). Perhaps a non-linguistic production might constitute a sign in this way as well; consider the act of pointing, in which the indexing body part (the signifier) refers to something in the environment (the signified). The notion of meaning arising across a combination of signs can be maintained as well. While a co-speech gesture, for example, occurs simultaneously with an utterance (i.e. not in a serial

manner), the gesture and the utterance both have a respective signified. The meaning of the utterance, then, might be reliant upon the interaction between these two signs.

It seems clear that gesture is an important part of understanding the linguistic communication, yet the field of "gesture studies" has been historically separate from linguistics. What is the reason for this? As has been briefly discussed, gesture is quite a broad domain and, consequently, is difficult to synthesize with existing linguistic theories. Gesture studies arose independently of linguistics and might be lumped into the broader domain of "cognitive semiotics" (Zlatev, 2015) (((SS4: N7))). Perhaps the hesitance to include gesture in a definitive account of language is the lack of clarity surrounding whether gestures are voluntary or involuntary. In other words, as Kendon (2007) remarks,

> …what level of control guides gesture? to what extent do they count in conveying what a person is saying to others? and when, as appears to be the case, gestures appear even when the addressee can't be seen by the speaker, do they nevertheless play a role for the gesturer himself? (((PS10: N5)))

While the field of gesture studies, detailed in more depth in the following section, does not necessarily provide the answers to these questions, they are critical considerations. Additionally, not all gesture research is concerned with gestures that co-occur with speech. Gestures, when used alone, tend to take a conventionalized form (Kendon, 2007). Such insights are not abundantly helpful for understanding a connection between gesture and language. Nonetheless, linguistics strives to "solve" the human faculty for language, a tool used for communication, and should incorporate non-ambiguous communication into its theories in order to provide a more comprehensive account for human language.

Thus, neglecting to expand this definition weakens an understanding of the human faculty for linguistic communication. This is not a controversial perspective; it is clear that there are gestures which co-occur with speech and, if these gestures are consciously produces (or

even if they are not) they contribute to linguistic signaling in a variety of ways. Therefore, the linguistic system must be extending to include gestures. Only the questions of how, and to what extent, remain.

A concept that has been mentioned and is considered further in the paper is embodied or situated cognition. Its relevance to a broadening of the linguistic system is clear, and insights from cognitive semiotics and linguistics aid in this conclusion. Wilson and Golonka (2013) note that an embodied approach to language to maintain that embodiment refers to the first-person experience of an organism (((PS17: N1))). In other words, from an embodied perspective, a linguistic agent cannot assume a third-person perspective to understanding the effectiveness of their communication. Their communicative signals are reliant on their body and the environment around them but, more importantly, they are reliant on these features. Thus, language as an embodied or situated practice makes use of the body, whether consciously or unconsciously, and the surrounding environment to do its best to convey some meaning. This gives way to co-speech gestures which might be as simple as raised hands or as complex as indexing something in the surrounding environment. A critical following of this approach is that not all gesturing, including co-speech gestures, is "done in the head." Gesture, if anything, is a testament to language as a situated practice by exposing the linguistic system's need to rely on the body and the environment.

What is a Gesture?

Kendon (2007) presents a comprehensive definition of a gesture, part of which is employed for the purposes of this paper. He describes a gesture as "a phrase of bodily action [with] characteristics that permit them to be 'recognized' as components of willing communicative action." They can manifest in three ways: (1) utterances on their own, (2) components of

utterances in alternation with speech, and (3) in conjunction with speech (((PS10: N2))).

Alternation refers to gesture as a separate utterance immediately following an utterance which

sometimes finish a phrase. Kendon (2007) also makes the distinction between gestures "co-

produced with speech" and gestures used "in conjunction with speech" (known as gesticulation)

For the purposes of this analysis, this dichotomy and alternating gestures are condensed into the

concept of "co-speech" gestures. While these distinctions are certainly useful, this paper is

concerned with the interaction of gesture and language and is thus concerned with all three

equally. Each co-occurs with speech and thus must be considered as interacting with language.

Crucially, though, Kendon does rely on the assertion that gesture and language are connected and

references the variance of degrees of this connection.

Kendon's definition of gesture is not the only one used within the domain of gesture

studies. For example, McNeill and Pedelty (1995) divide gestures (and, more specifically, co-

speech gestures), into four categories: iconic, metaphoric, beat, and abstract pointing or deictic

(indexical) gestures (((PS11: N1))). Such distinctions, though, do not weaken the notion of a co-

speech gesture, instead providing further grounds for this analysis. Iconics, for example, provide

concrete supplementation to the co-occurring speech while beats are "timed to occur with

thematic discontinuities in discourse" without being imagistic. Metaphorics are similar to

iconics, but are more abstract, and deictics are explicitly confined to the set of pointing gestures.

These definitions are useful in establishing a continuum for co-speech gestures more broadly and

are thus considered further in the paper. Co-speech gestures, which can manifest in a number of

forms, are the central focus of this paper due to their significance for the impact of gesture on

linguistic cognition.

The contemporary status of gesture studies can be considered within the domain of cognitive semiotics. Zlatev (2015) defines cognitive semiotics as a

transdisciplinary field focused on the multifaceted phenomenon of meaning, "integrating methods and theories developed in the disciplines of cognitive science with methods and theories developed in semiotics and the humanities, with the ultimate aim of providing new insights into the realm of human signification and its manifestation in cultural practices." (((SS4: N6)))

An important concept herein is how meaning arises via human signification embedded in cultural practices. This is a testament to the relevance of situated cognition and how gesture-language interaction gives rise to greater meaning. Cognitive semiotics is concerned with conceptual metaphor as well (((SS4: N4))). The notion of conceptual metaphor can be extended to the physical realm and consequently gestures whose articulation is consistent with some spatial metaphor (e.g. UP IS HAPPY). In general, a cognitive-semiotic lens is a useful tool for understanding not only gesture but how it might interact with language to give rise to (more explicit) meaning.

While there are gestures that are non-manual, this paper is concerned primarily with manual gestures in order to limit the scope of the analysis. Manual gestures are henceforth considered any produced as or more distal than the arms (i.e. the arms, the hands, and the fingers). Work concerning manual co-speech gestures is quite ubiquitous and thus provides ample evidence for this analysis. Non-manual gestures might be produced with (but are not limited to) the head or any articulators within it (such as the eyes), the shoulders, the torso, or the legs (ex: Caradec, 2018). Non-manual gestures are still considered to an extent, especially in regard to pointing, as articulators such as the head and shoulders tend to be used for pointing or for other purposes, especially when the hands are inaccessible (Kita, 2003, p. 111) (((SS3: N3))). That said, the mention of a co-speech gesture should be considered a manual gesture unless

otherwise noted. An inclusion of co-speech gestures which are non-manual would be an effective starting point for further analysis of the interaction between gesture and language.

An important aside is to explain the difference between gestures in spoken and signed languages. A sign (that is, an expression in a sign language rather than a semiotic sign) necessarily utilizes at least one hand or arm. There are no attested conventional signs which do not adhere to this restriction. Therefore, gestures in signed languages are typically restricted to non-manual articulators. These might include the non-manual gestures discussed above. Sign languages make particular usage of articulators on the face such as the eyebrows and the mouth. Liddell (2003) notes that these are suprasegmental markers which are innately tied to the grammar of a given sign language (((PS14: N2))). They are not "peripheral or paralinguistic," and Liddell eloquently notes that restricting the analysis of sign languages to the grammatical organization of symbolic units is "hopelessly inadequate," placing such languages at an interesting point at which gesture is uniquely tied to the language (((PS14: N3))). Thus, gestures in signed languages are seemingly inherently different from those in spoken languages. However, this perspective gives way to the conclusion that gesture is, indeed, a component of the linguistic system. Sign language gestures should then be considered, for the purposes of this analysis, as grounds for linking gesture and language more tightly.

McNeill, a prominent early gesture researcher and contemporary to Kendon, takes an explicitly psychological perspective which extends the Saussurian sign and is concerned with the iconicity of gesture. He argues for a broader conception of language why combining the structural aspects of Saussure's *language* with something more dynamic and imagistic, which is thus reliant on the prevalence of iconic gestures (Zlatev, 2015) (((SS4: N1))). McNeill (2000) gives the example of cracking a nut with empty hands. In this gesture, it is only necessary for the

hands to be what they would naturally do in performing such an action and thus the gesture is iconic (((SS1: N5))). It is clear that the iconicity is has immense significance. An iconic gesture, while usually reliant on some context delivered by an utterance or through the overarching gesture, can have a fairly explicit meaning. Many gestures are iconic; that is, their forms relate to their meaning in some capacity. The term "iconic" might be replaced with "imagistic" given this insight. Iconicity can be explicit (as in the nut example) or more abstract (as in McNeill's metaphoric gesture discussed above). Iconic gestures may be restricted to a discourse, but they may also not be (consider someone gesturing a wide path in front of their body as they say the word *continuum*). Therefore, iconic gestures might exist at a number of points on a continuum of co-speech gestures' interaction with language. They might be specifiers, like in the *continuum* example, but, in an established discourse, they could disambiguate an utterance. Consider the following conversation:

> SPEAKER A: I'm going to throw out this TV!
> SPEAKER B: You're going to throw it in the trash?
> SPEAKER A: No, I'm going to throw it out!
> > GESTURING HOLDING A TV AND THROWING IT TOWARD THE NEARBY WINDOW

In this interaction, Speaker A disambiguates the meaning of *throw it out* by including a gesture of throwing it out the window. Of course, Speaker A could have achieved the same effect by adding *the window* at the end of their second sentence, but the gesture achieved the same effect nonetheless. See the *Language Disambiguation* section for more information on this sort of gesture usage.

Thus, a brief taxonomy for co-speech gestures for the purposes of this analysis has been established. Co-speech gestures are not restricted to the hands (Caradec, 2018; Liddell, 2003), but this paper largely considers them as such in order to narrow the scope of this research.

A co-speech gesture be iconic or metaphoric, interacting with gesture in a number of different ways. It can be deictic, or indexical, disambiguating an utterance by means of a pointing gesture (Fricke, 2007) (((PS4: N4))). It can be, as McNeill says, *beat*, co-occuring with speech as a byproduct with no explicit meaning while potentially providing the speaker with greater emphasis. Finally, it can occur somewhere in the middle, specifying an utterance's meaning without necessarily disambiguating it. Any gesture discussed henceforth should broadly fit within one of these categories. Crucially, these categories aid in the establishment of a continuum upon which gesture impacts language or linguistic cognition to varying degrees, the ultimate goal of this analysis.

Semiotics and Gesture

Broadly, semiotics is an immensely useful tool for understanding gestures more generally, but also for understanding their impact on language. The lens of semiotics is wide and has quite a lengthy historic tradition. However, a number of semiotic thinkers present arguments which are necessary for analyzing the impact of gesture on language. In this section, two such arguments are discussed: Buhler's signal and social cohesion and Sebeok's biosemiotics. The inadequacies (and applications) of the Saussurian sign system have already been discussed. These perspectives lay the foundation explanations of the emergence of gestures and how they can interact with language.

Buhler's signal defines a sign that consists of a sender, a receiver, and a referent, a sign definition which is used henceforth in this analysis. This conception of a sign necessitates communication between two agents (Innis, 1985). This framework also provided the grounds to introduce the concept of deixis, or indexical signaling. Buhler's sign distinguishes between perceptual deixis and imagination-oriented deixis (Fricke, 2007) (((PS4: N2))). This distinction

is helpful in understanding how thinking agents use gestures to refer to points in space, for example; either by assuming their own person as the *origo* or some other point in space. Buhler's sign is concerned primarily with communicative signaling; that is, conveying some referent from one individual to another. This communicative sign not only suggests a relevance to a linguistic signaling device such as that considered in this paper but also the notion of social cohesion; thus, the relevance of Buhler's social cohesion to gesture and the linguistic system should be considered. Perhaps gesture arose as a means of social cohesion over time. It can be used to coordinate ideas by referring to objects in the world. Gestures interaction with language, then, is a further example of social cohesion, as the introduction of gesture to a linguistic utterance should clarify its meaning and facilitate cohesion between individuals in a community. This notion is important in understanding the development of gesture from an evolutionary perspective, but also why it is such a prevalent aspect of human linguistic communication.

  Sebeok introduces the field of "biosemiotics", a useful perspective for analyzing how humans might use linguistic signaling for interaction and cohesion. Biosemiotics, too, relies on evolutionary grounds to account for how living agents interact with one another. Human (species) survival is dependent on the survival of the individual and this survival might be brought about by some facet of cooperation (or lack thereof) in a society. Biosemiotics unifies ideas in biology in relation to semiotic concepts. Broadly, biosemiotics should be applied in the context of this paper by considering how gesture and language facilitate cooperation between individuals to perpetuate the survival of species. Thus, seemingly, biosemiotics builds upon Buhler's conception of a sign by considering how it might be applied for survival. A signaling device consisting of gesture and language is useful not just for individuals but for the survival of a society at large.

Thus, a comprehensive sign theory surrounding the utility of gesture and its impact on language can be established. Communicative signs are strengthened when they consist of both gesture and language. It is likely that such facets of human society arose in order to facilitate within group unity and cohesion of people and ideas. Language and gesture, then, can constitute a unified signaling device for human communication. The lens of semiotics is an effective perspective to take in analyzing this signaling system. Furthermore, a need to synthesize the historical separation of the trajectories of gesture of language arises. The manners in which language and gesture interact give way to a conclusion that is not consistent with this assertion. The section *The Development of (Sign) Languages from Gestural Systems*, in part, expands upon this notion.

Establishing the "Bounds" of Gesture and its Interaction with language

This paper intends to establish a continuum upon which gesture impacts language to different degrees. In establishing such a continuum, the terminal points (i.e. upper and lower bounds) of that continuum must be discussed. Toward the lower bound of gesture-language interaction lie emphasis and specification. While specification is likely more impactful on linguistic production, it is grouped with emphasis as both effects are essentially ambiguous; that is, neither facilitates outright disambiguation. At the upper bound lies said disambiguation of linguistic utterances. This is, seemingly, the upper bound upon which gesture can impact the meaning of a linguistic sign. Disambiguation gestures, like pointing, clarify meaning in an explicit manner. It is difficult to conceptualize a higher level of interaction (barring the relationship between gestural systems and language discussed later on). This continuum makes some suggestions about the impact of gesture. There are gestures which do not co-occur or alternate with speech, and these might constitute some peripheral set at the lower bound of the

continuum. In other words, the establishment of the continuum of gesture's interaction with language allows for the inclusion of linguistically isolated gestures, "gestures used alone" (Kendon, 2007), at the extreme lower bound. Any co-speech gesture (and, from a theoretical perspective, any gesture whatsoever) falls at some point on this continuum. This is not to say that all emphatic gestures are devoid of much meaning, or that pointing is always a successful disambiguation task. It is only to say that a continuum exists with a fairly nebulous lower bound and fairly definite upper bound.

A point of contention among gesture researchers and linguists concerned with language acquisition is the "ostensive gesture". Rodríguez et al. (2015) define ostensive gestures as gestures in which "the presence of an object which facilitates understanding in the absense of spoken language and pointing." They consider this act of "giving" or showing" a held object a deictic gesture. Rodríguez et al. stake the claim that ostensive gestures are acquired by children before pointing gestures (which are widely considered the gestural primitive from a developmental perspective) as comprehension of pointing requires "mastery of (1) the gesture itself (2) the understanding that it refers to *something* that is (3) located *in the distance*" (((PS7: N1))). In other words, a young child might look at the tip of the index finger instead of the referent in space when viewing a person pointing. They claim that, from a semiotic perspective, the sign and referent in a pointing gesture are *heteromateric*; that is, they do not coincide. Ostensive gestures, on the other hand, are *homomateric*. Thus, ostensive gestures are easier to understand (((PS7: N2))). Are ostensive gestures gestures at all? If so, where do they fall upon the continuum and what implications do they have? Clearly, they are not co-speech gestures as they are acquired pre-linguistically. Thus, they might fall at the very lower bound along with gestures produced alone. However, it is clear that are deictic. Consider if an adult were to

employ an ostensive gesture. Recall Speaker A who, instead of gesturing throwing the TV out the window, decides to pick up the entire TV and swing it in the direction of the window while saying *throw it out*. In this scenario, the ostensive gesture disambiguates. Thus, it would seem that the inclusion of held objects confounds the continuum. Therefore, they are best left out of this analysis. However, they need not be removed for the sake of convenience only; perhaps it is best to establish gestures as restricted to the body of an individual. This limits the scope of this analysis of the co-speech continuum.

It should be noted that there are more gestures (in terms of rote frequency) at the lower end of the continuum. With some brief consideration, this conclusion seems obvious. Kendon (2007) notes that most frequent environment in which gestures are observed are as components of spoken utterances (((PS10: N6))). Within this set of co-speech gestures, gestures for disambiguation are typically restricted to a given discourse. On the other hand, emphatic gestures tend to co-occur with many speech utterances; they are not restricted by discourse, only by physical barriers (e.g. inaccessibility to the hands). Thus, while disambiguating gestures constitute the upper bound of the continuum, they are subject to more restrictions and occur with less frequency. Gestures for emphasis and specification are not subject to the same higher-level restrictions and should be observed with greater frequency.

<div align="center">Language Emphasis and Specification</div>

The lower bound of the established gesture-language interaction continuum might consist of emphasis and specification. Gestures used for emphasis, emphatic gestures, could manifest as something as simple as raising the hands to exaggerate some notion (e.g. *he THINKS he know the answer*). Specifying gestures are those which are bound more tightly to some utterance but fall short of disambiguation. While they provide more explicit information than emphatic

gestures, they do not make explicit reference. Consider the following from Mendoza-Denton and Jannedy (2011):

> … while spoken discourse has a high referential resolution, that is, it is able to pick out referents with relatively little ambiguity, gesture has a low referential resolution, so most of the information presented gesturally is complementary to speech and not recoverable solely from the gestural channel… (((PS3: N1)))

With the exception of gestures used alone, gestural emphasis is the lowest level of interaction with language. Such emphatic gestures tend to co-occur with intonation changes (Mendoza-Denton & Jannedy, 2011) (((PS3: N2))). It is not difficult to conceive of a person who is upset about something to use large gestures while speaking emphatically, for example. These gestures are seemingly ubiquitous; spending a day observing the emphatic gestures of those around you would reveal this. This ubiquity is the most important consideration when discussing emphatic gestures. They co-occur so regularly with speech that they appear to be a part of language production more generally. While they lack explicit meaning, they seem to be inseparable from linguistic utterances. So, despite their status as low-level co-speech gesture, they are perhaps the most revelatory about embodied linguistic cognition.

Above gestural emphasis lies "specification" which is more impactful than emphasis but distinctly separate from it. Furuyama (2000) gives the example of someone saying "The cat goes straight into the building" while sweeping their hand horizontally to denote this movement. This gesture clearly refers to the cat's movement, but not explicitly. It must co-occur with speech in order to bear meaning. This example demonstrates a "tight synchronization tendency" between the signs (((PS1: N2))). Among these gestures, there is a greater connection between the gesture and the utterance with which it co-occurs. Not only does this ascribe greater linguistic impact to such gestures, it gives way to greater iconicity. Consider, for example, a person saying *Come hit me* while opening their arms. This gesture is not emphatic, it is clearly

opening up the body of the speaker to the person to whom they make the statement. Such a gesture is iconic in this sense of opening, but it is specifying because it does not disambiguate any meaning; the meaning of *come hit me* is fairly clear, and the accompanying gesture only bolsters it.

Gesture seems to regularly co-occur with language to some extent in unrestricted situations, and there must be an explanation for this. "Unrestricted situations" refer to situations in which a person has access to any number of gesture articulators. Consider a person who is cooking and has one hand on the pan and one hand on a spatula. While their hands are inaccessible for linguistic signaling, they may be able to use the head or a leg to perform some gesture. Situations in which gesture articulators are absolutely inaccessible are rare. Thus, the question of why co-speech gestures are so prevalent should be considered further. In a similar vein, disruption of gestural output can affect linguistic output. McNeill (2000) reports that inhibition of gesture affects speech fluency, including time spent pausing. Some accounts also note that inhibition of gesture can change the content of speech altogether (((SS1: N3))). People tend to use gestures even when no one is watching, such as when talking on the phone or practicing a speech in their living room (Mendoza-Denton & Jannedy, 2011) (((PS3: N3))). These gestures are likely emphatic or specifying. While they may be a mode of "self-expression", it is certainly possible that such gestures are byproducts of speech or a product of embodied linguistic cognition.

Language Disambiguation

As has been discussed, gestural disambiguation exists at the upper bound of the continuum. Deictics and iconic gestures used for explanation can disambiguate the meaning of a linguistic utterance in an explicit manner, and these seems to be the terminus for how gesture and language

interact. However, disambiguation may only be possible in a situated environment. It is more domain-specific than other co-speech gestures. Thus, despite having more power – Furuyama (2002) suggests that pointing "indexically entails" the demonstrative *that* and thus speech and gesture can be coreferential, a profound assertion indeed (((PS1: N1))) – these gestures are subject to more restrictions. Nonetheless, their place at the upper bound of the continuum warrants further discussion.

Pointing is perhaps the most salient act of disambiguation with co-speech gestures. While pointing is more frequently done with the hand, it can also be done with other articulators, as has been briefly discussed above. Note also that pointing is not necessarily a universal gesture (Kita, 2003) and it may not be a pre-linguistic primitive, though many do claim it to be (((SS3: N2))). Still, it is a prolific and robust gesture. Pointing gestures have a well-defined path that is typically linear. The articulator is usually held briefly at the end of this path (((SS3: N3))). It is also important to note that the referent of a pointing gesture can be ambiguous in many ways (((SS3: N1))). Thus, pointing is not always capable of explicit disambiguation, but it can be in certain contexts.

There are examples of disambiguation which are not a form of pointing. These, too, are seemingly restricted to a certain context or discourse and are otherwise meaningless. Furuyama (2000) provides the example of poetic structure, a discourse in which the location, handshape, or movement of a gesture can have explicit meaning therein (((PS1: N3))). Within such a context, says Furuyama, "communicational intent emerges thoughout the indexical relationship … between the linguistic and gestural signs" (((PS1: N5))). While this "catchment" could also include pointing, it does not need to (((PS1: N3))). Iconic gestures could also disambiguate. Consider an example of a person trying to explain the shape or size of an object to

another. If they trace the shape of the object in mid-air, or form it with the hands or arms, then they have utilized an iconic gesture which has disambiguated the meaning of an accompanying linguistic utterance.

Disambiguation lies at the upper end of the gesture-language interaction continuum because it is emblematic of the most profound impact that co-speech gestures can have. Still, in spite of the established continuum, language production *can* occur without gesture. Hearing humans are able to talk on the phone to one another and need not perceive the gestures of their conversant. There must be a reason for this. Evidence for it might lie in the neural substrates of gesture and language discussed in the next section. However, the existence of linguistic utterances without accompanying gestures does not disproof the hypothesis put forth in this analysis. From an embodiment perspective, the primary concern is the gestural behavior of the cognizing individual. From a communication perspective, linguistic meaning can only serve to be strengthened by the addition of gesture. Critically, in demonstrating a number of instances in which gesture can disambiguate, it is clear that gesture must be a component of the linguistic system even if it is not always present. Essentially, if gesture has the power to disambiguate, then it would be nonsensical to exclude it from a definition of a linguistic system because crucial signaling information would be excluded along with it. Thus, gesture is still a component of the linguistic system despite the terminal bounds of the interaction continuum.

Gesture and Language Processing Pathways

Evidence for the interaction between gesture and language might be found in evidence from neuroscience. Wolf et al. (2017) draw from Pierce's model of Universal Categories to stake the claim that semiotic categories describe cognitive categories and are thus able to be studied via "cognitive behavioral and neuroscientific testing" (((PS2: N1))). Their neuroimaging

research on the perception of co-speech gestures reveals that "conventional aspects of discourse-integrated gestures" are processed in areas of the brain responsible for language. In other words, the perceptual understanding of such gestures might rely on the neural substrates for language comprehension and, to a further extend, might be an integral part of the human language faculty (((PS2: N5). McNeill and Pedelty (1995) claim that, while gesture is not language, it is sometimes subject to the same constraints as language. They contend that co-speech gestures are controlled in a separate hemisphere from language production and thus a trans-hemisphere network is responsible for such productions (((PS11: N3))). However, when gestures must "bear the full burden of speech" (as is the case for deaf signers), the areas in the brain responsible for language production extend to gesture production (an obvious consequence of this is the development of signed languages, see the following section for further insight).

Cases of aphasia are particularly helpful, from a neurological perspective, in understanding the relationship between gesture and language. Nelissen et al. (2010) present a study which analyzes gesture use and semantic perception in patients with aphasia and dementia. They find that gesture imitation and discrimination and word repetition rely on the same perisylvian temporoparietal area (((PS9: N1))). They also note the left anterior temporal pole as an area responsible for both semantic processing and gesture discrimination (((PS9: N2))). These conclusions demonstrate shared nodes between language processing and gesture discrimination pathways. Marshall et al. (2004) present a case study of an aphasic Deaf signer. They conclude that aspects are language and gesture processing are shared insofar as language and gesture may emerge from the same conceptual system. In accordance with the assertion of McNeill and Pedelty above, they assert that "when gesture is synchronous with language" each system has the same semantic restraints and become co-expressive. However, they maintain that

gestures produced in isolation, and even in alternation with language, have a distinct semantic substrate (((PS12: N1))). Finally, Göksun et al. (2015) present a study of patients with lesions in areas associated with language production. Their findings suggest that there are relatively distinct regions governing naming and comprehending of motion events and, consequently, that patients with damaged linguistic areas are still capable of utilizing and understanding gestural patterns (((PS15: N2))). They do contend, however, that the networks governing these two systems do interact with one another.

Thus, two possibilities exist: either language and gesture share a neural substrate or are borne from different ones. Perhaps the answer to this might lie in some combination of the two. The neurological evidence discussed above reveals several crucial factors. The first is that there are aspects of linguistic and gestural competence that are governed by distinct areas of the brain. The second is that these such regions can interact and language networks in the brain can even utilize gesture in the situation of inaccessibility to spoken language. The third is that there are areas of the brain which govern both gesture and language. The latter two conclusions support the conclusion that, to some extent, language and gesture share neural substrates or can have overlapping neural substrates under certain circumstances. If language and gesture share neural substrates, even in restricted environments, then it is likely that gesture and language emerged from the same cognitive adaptations. Such a conclusion would lend credence to the relationship between gesture and language and provide evolutionary support for that relationship.

The Development of (Sign) Languages from Gestural Systems

The development of sign languages from gestural systems provides insights about how the mind uses gesture under different restrictions. One would be remiss not to discuss the emergence of signed languages without discusses the development of such sign languages as

Nicaraguan Sign Language (LSN) and Al-Sayyid Bedouin Sign Language. These sign languages are relatively young in comparison to older languages like American Sign Language and British Sign Language. Crucially, this recency has allowed for direct cross-generational observation by linguists. These sign languages emerged relatively quickly in relatively small domains. In the case of LSN, in the early 1980s, many deaf children were brought to one school in Nicaragua for the first time. These children had no language; at best they may have had a minimal understanding of mouthed or written Spanish and a set of home signs used with their families. When these students were brought together, they quickly developed a sign language that was passed through "generations" of students (new students entering the school). This situation provided a microcosm for linguists to analyze both the development of sign languages from non-linguistic systems and the rapid evolution of a language over condensed generations.

Such examples reveal how gestural systems can give way to full languages. Wilcox (2005) points out that there are at least two routes for the "non-linguistic" to be "codified into the linguistics": the development of words and the movement from the paralinguistic to the linguistic (((PS5: N1))). Wilcox notes how gestures like 'come here' can become lexicalized; that is, become conventional signs. One example of note is the use of COME-HERE in ASL for the concept of *I want more* (((PS5: N2))). Further examples are detailed below:

> The ASL evidential forms SEEM, FEEL, and CLEAR/OBVIOUS grammaticized from lexical morphemes MIRROR, FEEL, (used in the physical sense), and BRIGHT, respectively. Each of these lexical morphemes can be traced in turn to a gestural source. (((PS5: N3)))

Wilcox further claims that a developmental perspective such as this is necessary to understand the relationship between gestures and signed languages (((PS5: N2))). It has been previously discussed that the pathways in the brain responsible for language production utilize gesture when spoken language is unavailable (McNeill & Pedelty, 1995) and this, too, provides an explanation

for the transition gestural systems to languages and how the two are biologically linked. Ruben (2005), in his overview of the historical development of sign languages, notes the more recent ability to use neuroimaging for understanding the biological bases for sign languages and language in general. He notes that "Language is an intrinsic property of the nervous system which is dependent on a sensory input, but not on any particular sensory input." This conclusion suggests that language manifests in any form that it can, and therefore that sign languages arose based on the lack of auditory input (which, in turn, resulted in gestural-spatial output).

Home signs are another important consideration in understanding the connection between gesture and language. Home signs are essential for the emergence of many new signed languages (as in the case of LSN above). They typically emerge for non-linguistic communication between deaf and hearing members of a family or household. They do not have a fixed syntax and are usually restricted to content concepts. However, Coppola and Newport (2005) suggest that home signs bear more abstract linguistic structure than has been previously thought. The fundamental human language characteristic of grammatical SUBJECT is present in many home signs in their research. SUBJECTS do not have a specific semantic correlate but are usually the syntactic subject or object in an utterance (((PS13: N1))). Their findings suggest that fundamental language aspects appear in gestural communication even without linguistic input (((PS13: N2))). Therefore, even gestures which are seemingly primitive store semantic information without linguistic input. This provides support for the emergence of full languages from gestural systems as well.

In a somewhat related vein, the Gestural Theory of Language Origin (GTLO) is an important evolutionary consideration for understanding gesture and language as cognitive adaptations. The GTLO, in its most basic form, claims that human language emerged from early

gestures used for simple communication. It is possible that such gestures arose among early humans to be quiet while hunting. Given that infants can acquire gestures before spoken language, it is possible that gesture is the evolutionary precursor to spoken language and language more generally. Note, however, that the GTLO claims that gesture explicitly predates spoken language and full language systems. McNeill (2012) challenges this claim and suggests that language and gesture arose from a unified growth point, which can bring semiotically opposing modes of communication together at the same time (((PS6: N1))). He claims that gesture is likely generated as part of the process of speaking and was historically as well (((PS6: N2))). McNeill supports these claims with a fascinating theoretically adaptation known as Mead's Loop, which refers to the twisting of mirror neurons to respond to one's own gestures. Mead's Loop provides that (1) one's own gestures gained the sense of being social and public and

> (2) by mirroring one's own gestures and their significance, the new "twisted" mirror neurons made the gesture and its imagery available in Broca's area, the organ *par excellence* for complex action orchestrations; so vocal movements, originally for ingestion, could be orchestrated in new ways, by gesture imagery. (((PS6: N4)))

McNeill's modification of the GTLO presents a strong case for the unification of language and gesture into a single signaling system. It seems that the two modes emerged simultaneously and combined for certain tasks but remained separate for others. This conclusion is largely consisting with the neurological evidence discussed above. For nondeaf individuals, gesture and language exist separately but have a potent capacity to interact, a relationship which is likely backed by an evolutionary adaptation. If McNeill's theory holds true, than language and gesture have always been biological correlates and the historical separation of the trajectories of the two signaling modes is perhaps unjustified. Language and gesture constitute an a single

communicative device, a broad linguistic system, which has always been present in linguistic humans.

Concerning the gesture-language interaction continuum, the development of gestural systems into full languages from both the micro- and macro-perspective can be places at the upper bound. While this interaction does not constitute co-speech gesture, it reveals an inherent relationship between gesture and language more generally. These two examples are fairly concrete evidence for the implicit connection between human gesture and language. This analysis places this connection at the upper bound of its proposed continuum because it provides strong support for the claim that gesture can interact with language to the highest possible degree. In essence, evidence from the development of signed language and human language more general provides both the basis for the proposed gesture-language interaction continuum and the upper bound for it.

Conclusion

Language and gesture constitute a unified signaling system. This device is a tool for communication and human communication. This conclusion is supported by semiotic, neurological, and evolutionary evidence. While gesture and language can exist separately, they are unified upon a continuum which ranges from ostensive gestures and gestures which do not co-occur with speech to emphatic gestures to gestures with the capacity for disambiguation. There is a clear implicit biological connection between gesture and language which helps understand the points on this proposed continuum.

It is necessary to establish this system to better understand human communication. The field of linguistics has largely neglected gesture, but a field so concerned with human linguistic communication and cognition must consider it. While linguistics has an established

tradition that is not easily breached, the field should, at the very least, remain cognizant of the significance of gesture's interaction with language. Of course, there are areas of the field that are approaching this recognition, such as cognitive linguistics (which is closely tied to cognitive semiotics). In the future, the unification of gesture and language into a unified signaling system should allow for much better insights into the human mind and social interaction.

There are a number of further pursuits which could build off the discussions in this paper. One is the inclusion of more non-manual gestures and their potential similarities and differences to exclusively manual gestures. Another potential task is to devise a more comprehensive taxonomy for co-speech gestures under which any such gesture could be meaningfully categorized. It is also important to further consider signed languages in an analysis such as this, including how sign language gestures are different from those of spoken languages and where overlap, if any, occurs between gestures in spoken and signed language. If such steps were taken, it would only serve to bolster the continuum proposed in this paper and the other conclusions that it makes.

References

Caradec, F. (2018). Dictionary of Gestures: Expressive Comportments and Movements in Use around the World (Clarke, C., trans.). Boston: The MIT Press.

Coppola, M., & Newport, E. L. (2005). Grammatical Subjects in home sign: Abstract linguistic structure in adult primary gesture systems without linguistic input. *Proceedings of the National Academy of Sciences of the United States of America, 102*(52), 19249-19253.

Fricke, E. (2007). Where is here? The analysis of the German deictic hier with co-speech pointing gestures. In Mondada, L, (ed.), *Gestures and the Organization of Social Interaction: Ethnomethodological and Conversational Perspectives.* Proceedings of the 2nd Conference of the International Society for Gesture Studies (ISGS).

Furuyama, N. (2000). The Semiotic Making of Speech and Gesture and How It Contributes to the Structuring of Discourse. *Cognitive Studies, 7*(1), 71-77.

Göksun, T., Lehet, M., Malykhina, K., & Chatterjee, A. (2015). Spontaneous gesture and spatial language: Evidence from focal brain injury. *Brain & Language, 150*, 1-13.

Innis, R. E. (1985). Semiotics: An Introductory Anthology. Bloomington, IN: Indiana University Press.

Kendon, A. (2007). An Agenda for Gesture Studies. *Semiotic Review of Books, 7*(3).

Kita, S. (Ed.). (2003). Pointing: Where Language, Culture, and Cognition Meet. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Liddell, S. K. (2003). Grammar, Gesture, and Meaning. In Grammar, Gesture, and Meaning in American Sign Language, 355-362. Cambridge: Cambridge University Press.

Marshall, J., Atkinson, J., Smulovitch, E., Thacker, A., & Woll, B. (2004) Aphasia in a user of British Sign Language: Dissociation between sign and gesture. *Cognitive Neuropsychology, 21*(5), 537-554.

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

McNeill, D. & Pedelty, L. L. (1995). Right Brain and Gesture. In Emmorey, K., & Reilly, J. S. (eds.), *Language, Gesture, and Space*, 63-85. Hillsdale, NJ. Lawrence Erlbaum Associates, Publishers.

Mendoza-Denton, N., & Jannedy, S. (2011). Semiotic Layering through Gesture and Intonation: A Case Study of Complementary and Supplementary Multimodality in Political Speech. *Journal of English Linguistics, 39*(3), 265-299.

Nelissen, N., Pazzaglia, M., Vandenbulcke, M., Sunaert, S., Fannes, K., Dupont, P., Aglioti, S. M., Vandenberghe, R. (2010). Gesture Discrimination in Primary Progressive Aphasia: The Intersection between Gesture and Language Processing Pathways. *The Journal of Neuroscience, 30*(18), 6334-6341.

Rodríguez, C., Moreno-Núñez, A., Basilio, M., & Sosa, N. (2015). Ostensive gesture comes first: their role in the beginning of shared reference. *Cognitive Development, 36*, 142-149.

Ruben, R. J. (2005). Sign language: Its history and contribution to the understanding of the biological nature of language. *Acta Oro-Laryngologica, 125*, 464-467.

Wilcox, S. (2005). Routes from Gesture to Language. *Abralin, 4*,(1/2), 11-46.

Wilson, D. A., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4*, art. 58.

Wolf, D., Rekittke, L., Mittelberg, I., Klasen, M., & Mathiak, K. (2017). Perceived Conventionality in Co-speech Gestures Involves the Fronto-Temporal Language Network. *Frontiers in Human Neuroscience, 11*(573).

Zlatev, J. (2015). Cognitive Semiotics. In P. P. Trifonas (ed.), *International Handbook of Semiotics* (1043-1067). Dordrecht: Spring Science+Business Media.