"Sculpting Language:
A Review of the David McNeill Gesture Trilogy[*]"

Daniel L. Everett
Dean of Arts and Sciences
Bentley University
deverett@bentley.edu

[*] I would like to thank David McNeill, Gianluca Giorgolo, Sherman Wilcox, Eve Sweetser, Charles Goodwin, Wallace Chafe, Randy LaPolla, Sarah Grey Thomason, and Eitan Grossman for comments on earlier versions of this review. A special thanks to Stanley Dubinsky for asking me to consider writing this review article.

"Sculpting Language:
A Review of the David McNeill Gesture Trilogy "

1. Introduction

According to the works reviewed here, *dynamic* human gestures interact in complex ways with *static* human grammars to produce human language - neither grammar nor gesture is language apart from the other. It is underscored in this work, moreover, that human linguistic behavior is complex and therefore that gestures, as components of the multimodal behavior of language, are themselves intricate in structure, meaning, and use. The research of David McNeill and others on gestures make it clear that gesticulations are as analytically challenging and as intricate in design and function as syntax, phonology, morphology, discourse, and other aspects of language that some linguists consider "language proper." Gesture is not simply an add-on to language. There is no language without it.[1]

Kenneth Pike - the first person to ever discuss gestures in my linguistic training - saw gestures as evidence for his proposal that language should be studied in relation to a unified theory of human behavior (Pike 1967, 25):

"In a certain party game people start by singing a stanza which begins *Under the spreading chestnut tree*... Then they repeat the stanza to the same tune but replace the word *spreading* with a quick gesture in which the arms are extended rapidly outward, leaving a vocal silence during the length of time which would otherwise have been filled by singing. On the next repetition the word *spreading* gives place to the gesture, as before, and in addition the syllable *chest* is omitted and the gap is filled with a gesture of thumping the chest. On the succeeding repetition the head is slapped instead of the syllable *nut* being uttered... Finally, after further repetitions and replacements, there may be left only a few connecting words like *the,* and a sequence of gestures performed in unison to the original timing of the song."

From this example, Pike raises a question and draws a conclusion. The conclusion is that gestures can replace speech (in this case with "language-slotted gestures;" see below) in the story above because both speech and gestures are aspects not merely of language but of human behavior more

generally.[2] Human behavior, including communication, is the working out of intentions. Language is the best tool for communicating those intentions (Everett (2012: 160ff)). Or, as Giorgolo (2012: 8ff) puts it, "Communication has evolved into a cooperative behavior guided by principles of rational interaction."

The question Pike raises is an interesting one. I will call it "Pike's Problem," one that should be addressed by a theory and typology of gestures. The problem is just this: why is it that only sounds made by our mouth can be used in syllables and speech more generally? Why couldn't [sla#], where [#] is the sound of slapping your chest, not be a possible word or syllable in anyone's language?

As a student in Pike's introductory class, in which he asked this question, I thought it was interesting but did not appreciate adequately the degree to which the question impinges on the core of the linguistics enterprise.

Before proceeding to more detail on gestures, however, we might address the obvious question of whether speech itself is a gesture in McNeill's analysis. If so, then gesture research quickly begins to encompass all of linguistics. One reason to take speech to be a form of gesture is easy enough to comprehend - it involves movement of body parts (articulators to points of articulation) for the purpose of communication. On the other hand, there are two principal motives for excluding the movements of the articulatory apparatus from the gesture continuum. First speech gestures are generally unseen, happening behind the lips. Second articulatory movements and manual gestures contrast functionally. Manual gestures add information that syntax and speech alone do not carry. Thus there are both a functional and a formal distinction between speech and gesture.[3] These then exclude both movements of the vocal apparatus and the static grammar of hand movements in sign language (though sign language, like nonsigned speech, also has gestures) from the theory of gestures we discuss here.

Further, gestures aim towards perlocutionary effects, such that if some information is not part of speech proper, the speaker will try to express it in gestural terms in order to help the listener use or react to the information in the way intended.

Pike also asked his class to contemplate the following scene. Two men are watching other men move some heavy furniture down the stairs in their apartment building. The man passing them on the stairway landing is huffing and puffing and concentrating solely on his heavy load. His wallet is hanging loosely from his back pocket, about to fall out. He clearly wouldn't notice if someone relieved him of this burden. The first observer looks at the second observer with raised eyebrows, looking at the wallet. The second one sees him and simply shakes his head to indicate "No." What

kind of communicative act is this? Is it language? Is it parasitic on language? Is it a communicative act of a different nature? Should such acts fall under a theory of gestures? McNeill's theory provides detailed answers to these questions. A basic principle is that gestures enlarge to fill the communicative space shared with speech. As speech is diminished, as in Pike's example of the contemplation of robbing the overworked mover, or in the game that Pike mentions, gesture will take on a larger role to keep communication as effective as possible. This is further illustrated in Delayed Auditory Feedback experiments (see section 3 below) or the use of gestures to communicate in the absence of speech while hunting or fishing. Other examples are discussed below.

This paper overviews gesture research in general, focusing especially on the work of David McNeill, as developed in the three books listed above (hereafter referred to as the "trilogy," or by their individual dates). The review also adds my own perspectives on the significance of gesture for understanding the nature, origin, and use of language. However, the reader is especially urged to consult all sources cited in this text directly, especially the works by McNeill.

2. History of Gestural Studies

There is a broad, though perhaps vague, popular interest in gestures. For example, in a much commented-on article in the **New York Times**, Rachel Donadio wrote (June 30, 2013) that "When Italians chat, hands and fingers do the talking." She begins her entertaining piece with the line, "In the great open-air theater that is Rome, the characters talk with their hands as much as their mouths."[4]

But all languages have gestures. So is there any scientific sense in which Italian gestures are noteworthy? For example, can gestures vary significantly from language to language or from culture to culture? Or are they innately prespecified? Or are all gestures different in every utterance by every speaker? Italians' gestures are frequently caricatured on television and in movies. Are such caricatures based in fact? What appears to draw attention to Italians' gestures is their extravagance. Even in the 17th century, Northern European Protestants disapproved of Italians' "flamboyant" gestures (Kendon 2004, 21ff). The first person to study these gestures from a scientific perspective was David Efron, a student of Frans Boas, who wrote the earliest modern anthropological linguistic study on cultural differences in gestures (focusing on gestures of recent Italian and Jewish immigrants) more than 70 years ago.

Efron (1942/1972)'s study, **Gesture, Race, and Culture**, is at once a reaction against Nazi views of

the racial bases of cognitive processes, a development of a model for recording and discussing gestures, and an exploration of the effects of culture on gesture.

The core of Efron's contribution is his description of the gestures of unassimilated Southern Italians and Eastern European Jews ("traditional" Italians and Jews), recently emigrated to the United States and mainly living in New York City, though some of his subjects came from the Adirondacks, Saratoga, and the Catskills. According to Efron, (pp330ff), Italians use gestures to signal and support content (e.g. a "deep" valley, a "tall" man, "no way," etc.), of illustrating what was just said, with many *signs* (emblems in McNeill's sense) among the gestures. The Jewish immigrants of Efron's study, on the other hand, use gestures as logical connectives (to indicate the changes of scene, argumental divisions, and so on). These conclusions underscore and support the Boasian perspective that language is shaped by culture.

The questions that Efron set out to investigate were (i) whether there were standardized group differences in gestural behavior between the two different "racial" groups and, if so, (ii) to determine what becomes of these gestural patterns in members and descendants of the same groups under the impact of social assimilation. He found a strong cultural effect - "Americanization" of gestures in each group over time, minimizing their initially strong differences.

Working with artist Stuyvesant Van Veen, Efron also developed an effective methodology for studying and recording gestures, as well as a language for providing what Geertz (1973, 3-30) would have called "thick descriptions" of gestures. Abstracting away from the long anti-Nazi-science section, the book continues to be of interest for contemporary scientists.

Efron's need for the services of Van Veen was due, of course, to the lack of adequate technology for recording gestures. With the advent of certain types of new equipment, e.g. the cassette recorder, the laptop, and digital sound analysis, quality portable video equipment, etc. the course of linguistic history, as in any science (or any culture in fact), took a different course, directed by the objects the culture produces for specific purposes. We tend to do more of the things we have tools for. The better and more accessible the tools, the more the research they come to facilitate is undertaken Tools alter science. Efron's work was a pioneering work in many senses, but by no means the first work on gesture. Kendon (2004, chapters 3-5) provides a detailed overview of the history of gesture studies, going back to classical antiquity.

For example, Aristotle (384-322BC) discouraged the overuse of gestures in speech, as manipulative and unbecoming. Cicero (106-43BC), on the other hand, argued that the use of gestures was important in oratory and encouraged their education. In the first century, Marcus Fabius Quintlianus

(Quintilian, 35-100AD) received a government grant (perhaps a linguistic first) for a book-length study of gesture. For Quintilian and most of the other classical (and some later) writers, however, gesture was not limited to the hands but included general orientation of the body and facial expressions, so-called "body language."

As the Renaissance rediscovered the work of Cicero and other classical scholars, Europeans became interested in gesture and rhetoric, as in the 16th century work of Peter Ramus (1515-1572) of Paris on gesture in the style and delivery of oratory. Giovanni Bonifacio's (1547-1645) published the first book in Europe on gesture, **L'Arte de' Cenni** (1616). The first book in English on gesture was John Bulwer's (1606-1656) 1644 book **Chirologia: or the Naturall Language of the Hand**.

The 16th century studies often looked a good deal like actual science, not merely prescriptive offerings for improving public discourse delivery. In the 17th century some researchers moved on to question explicitly whether gesture might be developed to serve as a possible universal language. By the 18th century some began to wonder whether gesture might have served as the original source for language (some researchers, such as Arbib (2005; 2012) and Corballis (2002) still think so) though McNeill strongly discourages this view. Although these studies were to some disagree misguided (as McNeill points out in his work), they do fit with the idea, also pointed out in Everett (2012, 171ff), among other places, that speech as we know it almost certainly was evolutionarily secondary to the development of some sort of language, whether gestures and whistling or gestures and much less clear speech (lacking the vowels which are the hallmark of the speech of all humans, i, a, and u, for example). Gesture is fundamentally part of communication and, according to the logic and history of evolutionary development of language, communication precedes language.

By the medieval period (Kendon 2004, 328ff) thinkers from the early Renaissance were writing about gestural variation across European countries. In the counter-reformation period there were attempts to reform gestures (due to the perceived inappropriateness of some).

One study that merits particular mention was the work of Abbeí Charles-Michele de l'Epeé (1712-1789), who wrote an 18th century study of deaf education and sign language. L'Epeé was ahead of his time because he was not interested in getting deaf people to use spoken language, but rather in helping both deaf and non-deaf people to understand and appreciate sign language.

Although interest in gesture research declined markedly in the 19th century, according to Kendon (2004, 17ff) there were nevertheless two fascinating studies published. In 1832 Andrea de Forio published on gesture and cultural community in Naples while Garrick Mallery published in 1881, **Sign Language Among North American Indians Compared With That Among Other**

Peoples And Deaf-Mutes. This study was part of the **First Annual Report of the Bureau of Ethnology to the Secretary of the Smithsonian Institution**, 1879-1880, Government Printing Office, Washington, 1881, pages 263-552. As common in North America, linguistic understanding improved via studies of America's first nations.

After centuries of interest in gesture, from the 19th century, significant attention to gesture declined markedly. This was due in part, according to Kendon (2004, pp64ff) to the rise of Behaviorism and Psychoanalysis. Both of these movements were interested in those aspects of human behavior that were beyond conscious control. Since gestures are clearly in many of their forms consciously controlled these two fields would have had little interest in them.

Another factor in the decline of interest and study of gestures was the growing reification of linguistics (Kendon 2004, 65; Everett (2013a)). As methods developed for distributional analysis (e.g. Harris 1947, Wells (1947), Bloomfield (1933)) it became harder for some to see how to fit gesture into language, since their distributionalist methods could not readily account for gestures. Bloomfield (1933, vii) says little more of gesture than that it is "governed by social convention." Harris (1951), arguably the leading Bloomfieldian (until his student, Chomsky) was, according to Kendon (2004, 67), only interested in gesture to the degree that it was subject to the same kinds of distributional analysis as the rest of language. The contrast with Bloomfield, Harris, and others with the views of Sapir and Boas discussed in the next paragraph is striking. To my mind, for example this contrast shows that Bloomfield really did represent a distinctly non-Boasian, structuralist (vs. descriptivist) tradition that today continues in formal linguistics, which continues the structuralists' focus on form over function (Sakel and Everett 2012, 152ff).

Although superficially there are statements, e.g. Sapir (1951 [1927]: 556ff) that "the unwritten code of gestured messages and responses is the anonymous work of an elaborate social structure," which seem similar to the Bloomfieldian view, Sapir, Boas and their students understood language as one manifestation of human behavior, however specialized. Kenneth Pike, heavily influenced by Sapir (Pike 1978 and 1998, p.c.), was perhaps the most explicitly concerned of all linguists of the era in understanding language and gestures as components of human behavior more generally. Though Pike never developed a theory of gestures, he was a pioneer in prosodic studies, the functions of which can overlap with gesture's (Pike 1943, 1945).

Putting aside the work of Boas's student Efron, the marked return to and growth in gestures studies did not begin until the 1970s (when Efron's study was republished). At this time researchers focused primarily on understanding how speech and gesture were integrated. Are they two distinct systems

or parts of a single system? The interest of researchers in reviving gesture studies was sparked by a renewed interest in the evolution of language and the role of gestures in language development, the crucial role of visible body actions in interaction and communication, the universality of gestures, and understanding more precisely the relationship of gestures and speech as components of language more widely understood.

Progress was made, though slowly. Then in 1998, there was enough interest to hold the first international conference devoted entirely to gesture studies. Four years after this conference, in 2002, the journal **Gesture** and its sponsoring organization, The International Society for Gesture Studies were founded.

Throughout the latter part of the 20th century and now into the 21st century, two names stand out in the study of gestures, among a large number of researchers, Adam Kendon and David McNeill. Since McNeill's trilogy is our subject here, I will turn now to focus more on his work.

There is a long way to go before mainstream linguists afford gestural studies the prominence they deserve in the discipline. For example, I do not recall many instances where gesture researchers have been invited to offer keynote addresses to general linguistics conferences, not even to conferences dedicated to the evolution of language. Hopefully this review will help draw the attention of a wider body of linguists to the importance of gesture research for understanding syntax, speech, and language more generally.

3. McNeill's Theory

Gestures can be classified in various ways. McNeill finds it most illuminating to look at gestures in terms of their dimensions and their relationship to grammar and language. Relative to their relationship to language, gestures may be placed along the continuum in (1):

(1) The Gesture Continuum:

Gesticulation	Language-slotted	Pantomime	Sign languages

Gesticulation is the core of the theory of gestures. It involves gestures which intersect grammatical structures at the "Growth Point" (see below) and which are practiced by all humans so far as we know, even the blind and others with different types of cognitive disorders, such as proprioceptive deficiencies. Gesticulation is in fact what the theory of McNeill, Kendon, and others is mainly about. Gesticulations, like language-slotted gestures and pantomimes are not conventional - they may vary

widely and have no societally fixed form (though they are culturally influenced).

Language-slotted gestures are those which can actually replace a word, as for example, in the language game mentioned above with which Pike begins his 1967 book. Or imagine that you tell someone "He (use of foot in a kicking motion) the ball" where the gesture replaces the verb "kicked" or "She (use of open hand across your face) me (for "She slapped me")." These gestures occupy grammatical slots in an utterance and replace the grammatical units, usually words, that we otherwise expect there.

Pantomime: Pantomime gestures simulate an object or an action without speech. Like gesticulations and language-slotted gestures, they are also not conventionalized, meaning that their forms may very widely (though not utterly of course - see Everett 2015 for more information on tacit shaping of abilities and knowledge by culture and embodied cognition).

Emblems: These are conventionalized gestures that function as isolated "signs," such as the forefinger and the thumb forming the "OK" sign or "the bird," the upraised, solitary middle finger.

Sign-Languages are gesture-based languages that use gesture in a static, rather than dynamic, way. That is, there are gestural morphemes, sentences, and other grammatical units that are tightly conventionalized. Sign languages replace spoken language. They do not enhance it or interact with it - one reason that McNeill argues that spoken languages did not, and could not, have begun as sign languages. On the other hand, sign languages themselves make use of gesticulation, in addition to the conventional signs that form the lexical items of the grammar.

Part of what McNeill intends by the term "dynamic" is action-based, imagery-creating performances that are not conventionalized. Using the Vygotskian(/Marxist) terminology that influenced him (and me, among many others), he argues that language and gesture participate in a "dialectic interaction between unlike modes of thinking." (2004,4) This bimodal language-forming dialectic (but let us not forget other modes of language, such as affective prosody, register alternations, facial expressions, body language, and so on) is not an add-on to language. It is as much part of the multimodal whole of language as syntax.

McNeill claims that "dynamic cognition (2004,162) is environment-sensitive in its structure. Static cognition is synchronically insensitive to environment." I believe that he thus bridges a conceptual gap between different uses of the term "cognition" in modern linguistic theory. I have thought about this off an on for some time. Before reading McNeill (1992), I had thought that Everett (1994) was perhaps the first paper to draw an explicit distinction between "dynamic" and "static" cognition. In that paper I argued that sentential grammars are "static" (learned and fixed) objects whereas the use

of language in conversation and discourse was "dynamic (p131):" *"... I argue that discourse and sentence structures illustrate two types of cognition,* ***dynamic*** *vs.* ***static*** *and that these necessarily involve different theoretical constructs for their explanation. That is, they constitute distinct epistemological domains."*[5] In that paper, I further defined cognition as referring "*... to structures and processes which underlie reasoning, knowledge, and other higher-level, brain-caused behavior that cannot be accounted for in terms of neurons or arrangements of neurons in our present state of neurophysiological knowledge."* (p134)

I am pleased to learn that McNeill and I have been in agreement about the importance of these parameters for twenty years at least (though both of us could profitably have paid more attention to the dynamic aspect of even grammar and syntax that sociolinguists have discussed for so long). This seems quite an important distinction to draw. Much of the debate about what is "cognitive" in the cognitive sciences could be elucidated or eliminated if researchers were to recognize this distinction. Formal linguistics and functional linguistics, for example, both study cognition but from different perspectives if this distinction is correct. The sociological and scientific consequences of the failure to recognize this distinction have long been with us. One wonders how the history of linguistics and the cognitive sciences more generally might have been altered had gesture work impacted linguistic theories from the outset.

In the early rounds of funding for the cognitive sciences by the Alfred P. Sloan Foundation, grants were awarded to the University of California, San Diego; the University of Texas at Austin; MIT; Yale University; Brown University; and Stanford University. Later grants to further develop efforts were also awarded to Carnegie Mellon University; the University of Pennsylvania; the University of Chicago; the University of California, Berkeley; the University of Rochester; Carnegie Mellon University; and the Cognitive Neuroscience Institute. In these early awards and efforts by the different institutions (these grants coincided with my entry into linguistics as a student), there was conflict over definitions. Who was really studying cognition? (I recall participating in a LinguistList conversation about this subject before the era of web browsers, in the LL's earliest days.) Was discourse the proper linguistic domain of cognition or was sentential grammar? The dynamic-static distinction helps us to think through the implications of both perspectives without feeling the need to choose.

West coast cognitive scientists (such as Lakoff and Fillmore and Berkeley and Langacker at UCSD) often focused on discourse-construction as an active process that engaged concepts that included, inter alia, "live vs. dead metaphors," "framing," "newsworthy," "foregrounding," and "backgrounding." They emphasized those aspects of language that varied by specific speakers'

perceptions of context and communicative goals in real time. East coast cognitive scientists (clearly I am oversimplifying) at the time were concerned with greater knowledge of structure, formal semantic relationships, lexical meanings and structures, and so on. As McNeill (2004, 17) sums up this approach saying that "In this tradition, language is regarded as an object, not a process." He rightly links the concerns of the object-oriented vs. process-oriented perspectives to the early distinction between diachronic (process) vs. synchronic (static) studies introduced by Saussure. Another way of conceptualizing the static vs. the dynamic perspectives on cognition is in the importance of variation in each. Static, object-focused linguistics traditions think of idiolects or dialects as (at least for a moment) stable or discrete entities. It is possible in such a tradition to talk, for example, of someone's "I(nternalized/Intensional)-Language" as more than a convenient fiction, because a state of knowledge of grammar is seen as steady in some sense. But other linguists and cognitive scientists believe that the static view is a pernicious idealization. For these people, concerned with language as a process, variation is where the action is.

This aspect of McNeill's work is made all the more interesting to me because, as mentioned earlier, I have long advocated for the same distinction, arguing that sentence grammars (from a Chomskyan perspective at least) are formed from *static cognitive* capacities (e.g. Universal Grammar, learned constructions, templates, and so on), whereas discourse principles, e.g. emphasizing the newsworthy, emphasis, intonation, use of idioms, ideophones, and so forth present *dynamic cognition* (see Everett 1994 and 2015 for more discussion). At that time I argued for distinct theories for the analysis of each type of cognition, though I now believe that separate approaches may not be needed.

The "Growth Point" is the core unit of language in the theory of McNeill and it must be understood to understand the theory. The Growth Point is the moment of synchronization where gesture and speech coincide.[6] In McNeill (2012, 24) a growth point is described as that point where "... speech and gesture are (a) synchronized, (b) co-expressive, (c) jointly form a 'psychological predicate,' and (d) present the same idea in opposite semiotic modes." This description of the growth point contains several technical terms. By "co-expressive" McNeill means that (p12) for the symbols used simultaneously in gesture and speech "... each symbol, in its own way, expresses the same idea..." By "psychological predicate," terminology deriving from the work of Lev Vygotsky (1978), McNeill (2012, 33) intends the moment of the expression when "... newsworthy content is *differentiated from a context*." By "semiotic opposites" McNeill means that gestures are dynamic, created on the spur of the moment and, though influenced by culture and convention, are not themselves lexical or conventionalized units. Speech on the other hand contains grammar which is highly

conventionalized (grammatical rules, lexical items, etc) and is thus "static communication."

In short, gesture studies leave us no alternative but to see language in dynamic, process-oriented terms. Gestures are actions and processes par excellence. Nevertheless they do have object properties too, showing the ultimate weakness of dichotomies. McNeill provides some of the best analysis of the object properties of gestures. Thus he (2004, 31) defines a gesture unit as "... the interval between successive rests of the limbs." Like most units of human activity (Pike 1967, 82ff; 315ff), it is useful to break down gestures into margins (onset and coda) and nucleus. McNeill and others argue that gestures should likewise be analyzed in terms of *prestroke*, *stroke*, and *poststrokes*. And just as onsets, codas, and nuclei in syllables may be long or short, gestures may be lengthened in their different constituents, what McNeill calls "holds." In the prestroke, which like the other constituents of a gesture, may be "held" to better syncronize timing with the spoken speech, the hands move from their rest position in anticipation of the gesture. The stroke is the meaningful core movement of the hands. The poststroke is the beginning of the retraction of the gesture. The work on gesture is full of rich illustrations of these gestural constituents and how they at once synchronize with and add dynamically to speech.

These constituents and holds are strong indications that utterances are tacitly designed. As Kendon (2004, 5) says, "Gestures are part of the 'design' of an utterance." One of the clearest ways in which gestures show design are in the constituents of gestures (prestroke, stroke, and poststroke) and how they are often held to synchronize precisely with spoken speech.

Another crucial component of the dynamic theory of language and gestures that McNeill develops is the *catchment* (also referred to in places as a *cohesive*), which indicates that two temporally discontinuous portion of the discourse go together - repeating the same gesture indicates that the points with such gestures go together. In essence a catchment is a way to mark continuity in the discourse through gestures (Givon 1983) . McNeill (2004, 117) says that a

"... catchment is recognized when one or more gesture features occur in at least two (not necessarily consecutive) gestures. The logic is that recurrent images suggest a common discourse theme, and a discourse theme will produce gestures with recurring features... A catchment is a kind of thread of visuospatial imagery that runs through a discourse to reveal the larger discourse units that encompass the otherwise separate parts."

In the notion of catchment, gesture theory makes one of its most important contributions to linguistic theory in my opinion - it underscores that the constituents of sentences may have larger discourse functions that cannot be captured by sentence grammars. Thus catchments function simultaneously (each individual occurrence) at the level of the sentence and at the level of the

discourse (the shared features of the catchment gestures), illustrating that sentences and their constituents are themselves constituents of discourses, once again reinforcing the Pikean ideas on behavior, language, and "hierarchy" (where the apex of the grammatical hierarchy is not the sentence, but rather conversations, which may contain monologic discourses as constituents (see Reinbold (2004) for an excellent analysis of this in Banawá)).

On the other hand, gestures are not linked to sentences and discourses merely by timing and visual features. They are also connected semantically via "lexical affiliates." The lexical affiliate concept was first introduced by Schegloff (1984; McNeill 2004, 305) and it refers to "... the word or words deemed to correspond most closely to a gesture in meaning." Gestures generally precede the words that lexically correspond to them, thus marking the introduction of new meaning into the discourse. The close relationship, or "unbreakable bond," between gestures is demonstrated in several effects that result from real or imposed speech deficits: *Delayed Auditory Feedback* (DAF), blindness, and afferent disorder (proprioceptive deficits). In DAF experiments, the subject wears headphones and hears parts of their speech on roughly a .2 second delay, close to the length of a standard syllable in English. This produces an auditory stuttering effect. The speaker tries to adjust by slowing down (though this doesn't help because the feedback also slows down) and by simplifying their grammar. Interestingly, the gestures produced by the speaker become more robust, more frequent, in effect taking more of the communication task upon themselves in these situations. And yet the gestures "... do not lose synchrony with speech." (McNeill 1992, 273ff). In other words, gestures are tied to speech not by timing but by intent and meaning of the speaker and by their inextricable link to speech.

And this inextricability is quite special. For example, in McNeill (2004, 234ff), the case of the subject "IW" is discussed. At age nineteen, IW suddenly lost all sense of touch and proprioception below the neck due to an infection. The experiments conducted by McNeill and his colleagues show that IW is unable to control instrumental movements when he cannot see his hands (though when he can see his hands he has learned how to use this visual information to control them in a natural-appearing manner). What is fascinating is that IW, when speaking, uses a number of (which IW refers to as) "throw away gestures" that are well-coordinated, unplanned, and nonvisually-reliant speech-connected gestures. McNeill concludes that at a minimum this case provides evidence that speech gestures are different from other uses of the hands, even other gesturing uses of the hands. However, I am nonetheless unconvinced by McNeill's further conclusion that there is some innate thought-language-hand neural pathway in the brain. On the

other hand, I *am* convinced that such a pathway arises developmentally and that it is different from other pathways involving gestures and movement.

Finally, with regard to the special relationship between gestures and speech, McNeill observes that not only do sighted people employ gestures when talking on the phone, showing that gestures are not simply something added for the benefit of an interlocutor, but also that the blind use gestures when speaking (2012, 13), indicating that we use gestures even when we cannot see them and thus that they are a vital constituent of normal speech (whether this also means that we use gestures without learning them, as McNeill seems to suggest is a separate question). Thus DAF, the blind's use of gestures, and the "throw away" gestures emerging even in the context of IW's proprioceptive disorder suggest that the relationship between gestures and speech is in McNeill's words an "unbreakable bond."

Fascinatingly, however, though the bond may be unbreakable, it is culturally malleable. David Efron's work may have provided the first modern study to examine the link between culture and gesture. But it is not the only one. There is a now a sizeable literature on such effects. To take one example, de Ruiter & Wilkins (1998) and Wilkins (1999) discuss the case of Arrernte in which the connection or "binding" of speech and gesture is overridden by cultural forces (or, in the terminology of Everett (2012) cultural "values"). According to de Ruiter and Wilkins, the Arrernte regularly perform gestures *after* the co-expressive speech. The cultural reason the authors suggest is that the Arrernte make much larger gestures physically than are found in many other cultures, using movements of the entire arm in gesturing. Thus, as they interpret the phenomena, the larger gestures and space required by the Arrernte demand more planning time, favoring the performance of gestures following the relevant speech. A simpler analysis is suggested by McNeill (2004, 28ff), however, namely that the Arrernte simply prefer the gestures to follow the speech. The lack of binding and different timing would simply be a cultural choice. Gestures for the Arrernte could then be interpreted similarly to the Turkana people of Kenya, also discussed by McNeill, in which gestures function in part to echo and reinforce speech.

Whatever the analysis, one must appreciate the relevance and significance accorded to culture in McNeill's and other researchers modern analyses of gesture, following on in the Boasian tradition inaugurated by Efron.

A related area worth exploring in more detail, but which there is no space to do here, are the conventions for the annotation of gesture, especially with regard to variation of gestures cross-culturally. There are several sites (just enter "gesture lab" or "gesture annotations" in your

favorite search engine) that discuss annotation of gestures. One very good one is here: http://annotation.exmaralda.org/index.php/TASX, while cross-cultural variations in gesture and its annotation are discussed in Kita (2000).

4. Equiprimordiality and the Evolution of Gestures

By "equiprimordiality," McNeill claims that gestures and speech were equally and simultaneously implicated in the evolution of language. To understand this, we must ask how the Growth Point and the imagery-language "dialectic" evolved. Here McNeill (2012, 65ff) relies on the seminal work of George Herbert Mead's (1974) on the evolution of the mind as a social entity (see also Everett in preparation a), with special attention to language and gestures. Mead's claim on gestures (1974, p47ff) is that "Gestures become significant symbols when they implicitly arouse in an individual making them the same response which they explicitly arouse in other individuals" (this was probably written in the 1920s). McNeill's insight is to take Mead's conjecture and tie it in with Rizzolatti and Arbib's (1998) suggestion that mirror neurons underlie Case Grammar. What McNeill claims is that Rizzolatti and Arbib missed a crucial step which he refers to as "Mead's Loop," wherein one's own gestures are responded to by one's own mirror neurons in the same way that these neurons respond to the actions of others, thus bringing one's own actions into the realm of the social and contributing crucially to the development of a theory of mind - being able to interpret the actions of others under the assumption that they have minds like we do and think according to similar processes. Thus McNeill at once links his research program and the evolution of language more generally to the brain and society in an interesting and unique way.

According to McNeill (2012, 69), Mead's Loop entails that "speech and gesture had to evolve together." "There could not have been gesture-first *or* [emphasis McNeill's, DLE] speech-first.[7]" This follows, he claims, because Mead's loop creates a "dual semiotic." "To create the dual semiotic of Mead's Loop, they [speech and gesture, DLE] had to be equiprimordial." Mead's Loop made possible the dynamic aspects of speech as well as the analysis of otherwise holophrastic constructions into parts, such as words, phrases, sentences, morphemes, phonetic segments, and so on. McNeill explains this (2012, 67) by claiming that;

"Semiotically, it [Mead's Loop, DLE] *brought the gesture's meaning into the mirror neuron area. Mirror neurons no longer were confined to the semiosis of actions. One's own gestures ... entered, as if it were liberating action from action and opening it to imagery in gesture. Extended by metaphoricity, the significance of imagery is unlimited. So from this one change, the meaning potential of language moved away from only action and expanded vastly."*

I notice a couple of things in this quote. First, the language focuses on action and meaning rather than structure, which sets it off from, while allowing it complement, a great deal of linguistic analysis. Second, it places compositionality in a somewhat different light in the evolution of human language. Most linguists, myself included, would have likely answered, when asked what the great quantum leap in the evolution of language was, "compositionality." (This is in effect what Hauser, Chomsky, and Fitch (2002) intend by their obscure references to "recursion" in a famous paper.) But if McNeill is even half-right here, the Growth Point's evolution from Mead's Loop is even more important than compositionality. In fact, in Everett (2012a) I allude to the possibility that compositionality relies on non language-specific cognitive abilities. But the Growth Point is indisputably language-specific so one is tempted to agree with McNeill here that it is indeed the most important development in the evolution of language. Interestingly and unfortunately, this is completely ignored in most recent works on the evolution of language, e.g. Fitch (2010). But see Everett (in preparation b) for a full integration of the Growth Point into the understanding of language evolution.

Once we get past this initial hurdle of how gestures become meaningful for humans, other notions arise to fine-tune the evolutionary story of the gesture-speech nexus. McNeill's theory, (e.g. 1992, 311ff) takes a perspective similar to Construction Grammar (Goldberg 1995) in claiming that utterances - gesture/speech wholes - are initially "holophrastic," used as single words or unanalyzable wholes, and later, through reuse and gestural focusing on specific components of the holophrastic construction, analyzed in more detail, leading to grammatical rules in a way reminiscent of the discovery methods of Harris (1947, Longacre (1964), and others, i.e. distributional isolability and recombination).

As gestures and speech become signs in the social space, gestures take on one of two perspectives (McNeill 2005, p34). They either represent the viewpoint of the observer/speaker, *OVPT*, or the viewpoint of the person being talked about, or *Character View Point*, *CVPT*.

For example, McNeill gives an example of one person retelling what they say in a cartoon of Sylvester the cat and Tweetie-bird. When their hand movements are meant to duplicate or stand for Sylvester's movements, then their perspective is CVPT. But when their hand movements indicate their own perspective, then their perspective is OVPT.

Many researchers have speculated that gestures might have preceded speech in the evolution of human language. McNeill does not disagree entirely with this position. His reasons are similar to those suggested in Everett (2012a). Intentionality is a necessary prerequisite to language. And

intentionality is shown not only in speech but also in gestures and other actions and states, e.g. anxiety, tail-pointing in canines, focused attention in all species, etc. (see also Everett in preparation b). The implications of McNeill's analysis of Mead's Loop and the Growth Point are enormous. For one thing, if he is correct, then gesture could not have been the initial form of language. This is not to say that pre-linguistic creatures cannot express intentionality by pointing or gesturing in some way. It does mean that real linguistic communication must have always included *both* gestures and speech.

Another interesting component of McNeill's theory of language evolution concerns his own take on recursion. Recursion is (see Everett 2010) a tool for more tightly packing information into single utterances.[8] Thus he independently arrives at an important conclusion in recent debates on recursion by providing a model of language evolution and use in which recursion is useful but not essential, a very similar point to Everett (2005, 2008, 2009, 2010, 2012a, 2012b, and many others).

In recent years, Tomasello (1999, 2008), Corballis (2002), Hewes (1973), and Arbib (2005), among many others, have argued that "... language evolved, not from the vocal calls of our primate ancestors, but rather from their manual and facial *gestures*." (Corballis 2002, ix) However, there are two theory-busting problems with the "gesture-first" theory of language evolution, according to McNeill. First speech did not supplant gesture. Rather, as all the work of McNeill, his students, and many, many others show, the two form an integrated system. The gesture-first origin of language predicts asynchrony between gesture and speech, since they would be separate systems. But they are synchronous and parts of a single whole. Further, code-switching between gestures and speech is common. Why, if speech evolved from gestures, would the two still have this give-and-take relationship? Moreover, if the gesture-first hypothesis is correct, then why, aside from languages of the deaf, is gesture never the primary "channel" for any language in the world?

The second major problem with the gesture-first theory is that if gestures could be substituted by speech, they would not then be of the right type to form a language. This follows because in the absence of language, the available communicative gestures would have to be pantomimes. But, as McNeill makes clear throughout the trilogy, *pantomime repels speech*. Pantomime does not accompany speech - it fills in missing values or gaps in speech. It is used in lieu of speech.

Also, as McNeill makes clear throughout his trilogy, speech is built on a stable grammar. The only gestures which provide stability are the conventionalized and grammaticized gestures in sign languages. In this case, gestures are used instead of/supplant speech. Summing up, had sign language or other gestures, e.g. pantomimes or language-slotted gestures, preceded speech, then

there would have been no functional need for speech to develop. As McNeill (2012, 60ff) puts it "First, gesture-first must claim that speech, when it emerged, supplanted gesture; second, the gestures of gesture-first would [initially, DLE] be pantomimes, that is, gestures that simulate actions and events; such gestures do not combine with co-expressive speech as gesticulations but rather fall into other slots on the Gesture Continuum, the language-slotted and pantomime."

One might attempt to reply that Pike's example shows that gestures can substitute for speech. But the gestures Pike discusses are language-slotted gestures, parasitic on speech, not the type of gesture to function in place of speech. On the other hand, Pike's example suggests another question, namely, whether there could be "gesture-slotted speech" corresponding to speech-slotted gestures, i.e. an output in which speech substitutes for what would normally be expressed by gestures. If speech evolved from gestures, after all, this is how it would have come about. And gesture-slotted speech is not hard to imagine. For example, consider someone bilingual in American Sign Language and English substituting a spoken word for each sign, one-by-one, in front of an audience. Yet such an event would not really exemplify gesture-slotted language, since it would be a translation between two independent languages, not speech taking the place of gestures in a single language.

And as they stabilize by conventionalization, such gestures become sign-languages. But these are all gestures replacing speech functions and thus for speech to develop from these would make little sense either functionally or logically.

There does seem to be something missing in McNeill's reasoning about the absence of gesture-first, however, even though I agree with his proposals on this overall. If he were correct that two now-extinct species of hominin had used gesture-first/only language and that this is the first stage in the ontological development of modern language, then why would it be so surprising to think that Homo sapiens had also used gesture-first initially? I think his reasoning is solid relative to the nature of the gesture-speech relationship, but this is something I would have liked him to discuss.

Another question arising in connection to the equiprimordial relationship between speech and gesture is this: Is there a common, *specific* innate cerebral basis for language and gesture/gesticulation? McNeill (1992, 333ff) seems to think so if I read him correctly. My own opinion is a weaker "perhaps." McNeill reviews evidence to show that the cortical proximity of speech and gesture is directly proportionate to the distance leftwards the gesture is located in Kendon's gesture hierarchy. This means that the closer a type of gesture is tied to speech, the greater the proximity of that type to speech in the brain. Yet this is not support for any innate pathway. There is nothing in cortical proximity that could not be accounted for hypothesizing merely that the

two are learned together and initially experienced together. Somewhat stronger evidence for the neurological connectedness of speech and gesture emerges from studies of aphasia. In people with Broca's aphasia, meaningful gestures are produced in a choppy fashion. In Wernicke's aphasia, on the other hand, the gestures are meaningless but fluent - at least according to Hewes (1973). Yet this proximity again does not support a nativist "bioprogram" (a word McNeill occasionally uses) of any sort (moreover, Everett 2012a questions the very existence of these two types of aphasia, based on the simply fact that these correspond to no language-specific parts of the brain). It is explicable by the general principle of "adjacency" (sometimes also called "iconicity") noted across various studies in linguistics and other disciplines, namely, that the more two things affect one another, the closer they will be to one another. This idea has applications in the understanding of morphosyntactic constituents, vowel harmony, and neurological coordinates.

5. Other Types of Gestures

McNeill discusses a variety of different types of gestures. We have discussed catchments earlier, so now the other three - iconic gestures, metaphoric gestures, and beats, are introduced. Each reveals a distinct facet of the gesture-speech relationship. And each, like gestures and speech more generally, reveal a great deal of "tacit knowledge" (Polanyi (1974; 2009), Everett (in preparation a), among others).

McNeill (1992, 12ff) describes iconic gestures as bearing "... a close formal relationship to the semantic content of speech." Iconic gestures show that "... what is depicted through gesture should be incorporated into a complete picture of a person's thought processes." These gestures depict or represent concrete objects to flesh out the imagery and meaning of speech. For example, making the motion and appropriate hand shapes of pulling back a bow when discussing shooting an arrow, a common occurrence in Amazonian communication.

Alongside iconic gestures, speakers also use metaphoric gestures. Metaphoric gestures are simultaneously metalinguistic - representing discourses or discourse genres, etc. - and cultural (based on what counts as a metaphor). These gestures are abstractions. For example, McNeill (1992, 14) illustrates a speaker holding up both hands, palms facing inward, to represent a span of speech, e.g. a story told via a cartoon in the first example McNeill uses.

A third form of gesture is the beat. "Beats mark information that does not advance the plot line but provides the structure within which the plot line unfolds." (McNeill 1992, 15ff) These can signal departures from the main event line, such as a hand movement to accompany a "summing up" of

what has been said in a discourse to this point. Beats can also be used to segment discourses or to accompany phonological emphasis in the speech stream. In Everett (1988) I discussed the case of my language teacher, Kaioa, using gestures to indicate stressed syllables (seen at the webpage linked below engaged in a discourse and using hand gestures of different types: https://www.youtube.com/watch?v=SHv3-U9VPAs). Ladefoged, Ladefoged, and Everett (1997) also discuss training speakers to mark phonological beats in three different Amazonian languages. It turns out that this is not be as unusual as I thought it was at the time, since this is not an uncommon function of beat gestures.

6. The Growth and Spread of Language

McNeill's hypotheses on language evolution and spread are interesting and deserve a much more extensive and detailed treatment than I am able to offer here. But the themes of static vs. dynamic components of language are of foundational importance to all of his work in this regard. Mead's Loop (ML) and the Growth Point (GP) it produced are for McNeill the ur-moments of language evolution among hominins, as we have seen. But although the GP and ML are crucial, they do not get us to syntax and compositionality, perhaps the issues of greatest important to the vast majority of linguists (this is not to say that they should be the "issues of greatest importance").

When McNeill (2012,76ff) finally arrives at his narrative on the origin of syntax, he says "How, when, or even why syntax emerged is far from obvious, but here began the static dimension." The subsequent discussion is controversial, interesting, and very insightful for the most part. McNeill accepts, as I do and - as anyone must - that gesture and the gesture-language or dynamic-static interaction rest on a biological foundation: "I follow Lenneberg (1967) and claim that syntax rests on a biological foundation..." In this very general (and not particularly contentful sense), no one could possibly dispute the fact that language, like cooking and guitar-playing, etc. relies on human biology. The point is that there is nothing in McNeill's trilogy or other work in which specific biological mechanisms or genetics are causally implicated in his analyses. Thus "biology" and related terms could have been profitably omitted. [9]

More relevantly, however, McNeill (2012, 77) suggests that "an area of life where a syntactic ability could evolve is the cultural and social encounter." Here he cites the work of Freyd (1983) on "shareability" - the idea that structures and meanings must come to be shared among individuals if we are to say that they speak the same language, i.e. are utilizing the same outputs of conventionalization. In particular, McNeill appeals to Freyd's "discreteness filter," an idea akin to the

generative notion of "discrete" in the phrase "discrete infinity (for criticisms of the latter see Everett 2010)." The idea is that our utterances are initially holophrastic, noncompositional. Then as humans began to learn a repertory of such utterances, these would begin to change via the GP, such that gestures would highlight some portions of the previously unanalyzable whole, leading to an analysis of the holophrastic into component parts - top-down parsing that eventually results in compositionality.

This points to a stark difference between McNeill's theory and other theories of language evolution (such as Hauser, et. al. 2002). In McNeill's theory the compositionality of syntax arises from actual language use via GPs, not from a sudden mysterious appearance of compositionality via recursion. In fact, in McNeill's theory (and in Kinsella's (2009) and Everett (2012a), among many others), compositionality *precedes* recursion. The following quote expresses this well: "Contrary to traditions both philological and Biblical, language did not begin with a 'first word. Words emerged from GPs. There was an emerging ability to differentiate news worthy points in contexts; a first 'psychological predicate' perhaps but not a first word." Ironically, by demonstrating how compositionality could have come about by use and thus entered all human languages from early human interactions, McNeill undermines the need to appeal to genetics or biology to account for this. In the context of his discussion of compositionality, moreover, McNeill offers an extremely interesting discussion of how recursion itself might have entered grammar, one quite compatible with my own (Everett (2010), Everett (2012a), Everett (2012b)). The story begins with the analyzability of holophrastic utterance via Growth Points.

Recursion would not have begun with gestures themselves. This is because gestures, unlike the eventually compositional static outputs of grammar, are gestalt units (though not all gesture researchers accept this). This is a fundamental difference between these dynamic units vs. static syntax. Gestures are wholes without meaningful parts. And the meaning of the whole is not derived from the meaning of the parts. Thus although we can observe several submovements in the larger gesture, none of these smaller acts has any meaning apart from the gesture as a whole. Gestures are in this sense *anticompositional*.

But syntax became analyzable and *following this* recursion was able to play a role in the grammar. In this sense, recursion for McNeill, as for me (inter alia Everett 2005, 2008, 2009, 2010, 2012a, 2012b) is a nonessential, highly useful component of language evolution (contra Hauser, et. al.). Recursion is used to render the syntagmatic paradigmatic, enabling speakers to pack more information into single utterances and, as I point out in Everett (2012b), making it easier to follow complex events via oral

discourse. McNeill (2012, 223) cites Shelley's "The Masque of Anarchy" to illustrate the syntagmatic to paradigmatic shift:

"His big tears, for he wept full well,
Turned to millstones as they fell."

"The rhyming '-ell's, on the axis of combination, project a new semantic opposition..." This opposition is between paradigmatic parts of larger sentences (syntagmemes) that themselves derive from syntagmemes. Thus having provided a plausible for scenario for the evolution of syntax, McNeill turns to consider the resultant spread of static grammar.

7. Distance and Divergence in Language Evolution

McNeill (2012, 92) suggests that "If we take the Tower of Babel story as a parable of migration, it is not as far-fetched as one might suppose. The insight is that migration leads to encounters and breeds diversification; and the further the migration, the more the encounters, and the greater the diversification." Continuing with the Biblical metaphor, however, McNeill's speculation has to overcome the "Who did Cain marry?" problem. Many students of the Bible find it curious that Cain was the son of Adam, the first man, yet he was able to find a wife in a neighboring city (Genesis 4:1-5:5). Where did the inhabitants of that city come from? By the same token, one must legitimately inquire as to who encountered whom in the migrations from Africa. If the first encounter of human languages following the rising of Ur-Language produced linguistic changes because of distinct languages coming into contact, then how did the first language change? There must be change without contact. And in fact specialists in diachronic linguistics tell us that this is correct - that language change can be internal, e.g. sociolinguistic shift, or external, via language contact. But if change can occur without contact, does McNeill's hypothesis lose its force? No, because he also predicts that the trail of change should lead to greater and greater complexity the farther humans migrated from the geographical source of Ur-Language. The reasoning is that the earliest language would have had the simplest syntax, mapping meanings to temporal orders, i.e. iconically. But language contact would complicate that. To make this point, McNeill cites Sarah Thomason (see Thomason in press). However, here the reasoning weakens in my opinion. Thomason never argues that contact always makes languages more complex. She presents a few cases where it can. But folks who work in language contact know that there are too many variables to say that a given language contact situation will result in one of the languages becoming either more complex or more simple. In fact, theoretical linguists and typological linguists know that there

is almost no content to the claim that "Language A is more complex than Language B." About all we can say is that "Phenomenon A in language X is more complex than the same phenomenon in language Y." But even that is problematic because it assumes that we can say that this or that phenomenon in one language is the same as a phenomenon in another language, an idea I see little evidence for. That is, whether in smaller "societies of intimates" or larger "societies of strangers," diversity of pronunciation, grammar, and meaning all enter languages for any number of reasons that the diachronic linguistics literature is full of. So even if McNeill were correct (2012, 92) that initial syntax mapped meanings onto temporal orders of events - and, frankly, I see no evidence that this was ever the case - why would contact unidirectionally increase complexity? In fact, as Trudgill (2011) argues, contact can either simplify languages or complicate them. I am therefore unable to accept McNeill's conjecture that languages with a greater history of contact and thus farther from Africa should be more complex than languages with a lesser history of contact (and how could one even measure that in any case, apart from geography which need not necessarily entail greater contact). It would also be necessary, were we to take McNeill's proposal here seriously, to discuss the possible confounding factor of the "serial-founder effect" common to migrating populations and how this might also impact language (see, inter alia Slatkin & Excoffier, 2012).

Other (minor) criticisms of McNeill's work are also found even among gesture theorists. For example, Zlatev (forthcoming) and Andrén (2010) both develop alternative approaches to some aspects of the theory of gestures. Indeed the field of gesture studies has grown so quickly and has become so large, that it would be surprising if any single researcher's work were not reworked, as is standard in all fields. Nevertheless, the existence of such criticisms and the number of researchers with different emphases, research programs, and so on in the field should indicate to the reader that even though McNeill's trilogy is a splendid introduction and defense of gesture research and conclusions, it is by no means the final word on the subject.

Concluding this section on criticisms of McNeill's theory, I want to include a brief assessment of his speculations regarding the speech of Neanderthals and *Denisovan hominins*. McNeill (2012, 165) says this: "Gesture-first may have existed in the two now extinct human lines, Neanderthals and *Denisova hominin*. It could have existed in our line and extinguished as well, but we have survived to evolve a new form of language, Mead's Loop based on speech-gesture equiprimordiality. This new language, as we have seen could not have emerged from gesture-first and was a *second origin* [emphasis McNeill's, DLE]." From this speculation, McNeill speculates further that these two species lacked language and failed to survive in connection with that deficiency. To me (Everett in preparation b)

we know too little about these species to warrant such claims. One other speculation he offers, a variant of "ontogeny recapitulates phylogeny," is worth additional thought. It is that (2012, p165) while children apparently use gesture-first in their initial language acquisition, this dies out before two years of age. Then at three or four years of age GPs emerge, "... suggesting that gesture-first had existed once phylogentically but went extinct and was followed by a new form of language in which speech and gesture imagery merged into the unified packages inhabited by thought and being that we ourselves have now." Although, again, the significance of children's acquisitional stages is somewhat conjectural, it is interesting enough to merit further investigation. It is a very positive sign indeed that even where McNeill goes out on very thin empirical limbs, his suggestions are remarkably interesting.

8. Extensions of the theory

McNeill's work has been exceptionally fecund, inspiring brilliant work from a large number of scholars internationally. I will discuss only three here. The first is the work of Gianluca Giorgolo, of Carleton University and Oxford University. Giorgolo's work is sophisticated and formal. As he says in the abstract of Giorgolo (2010), "The paper presents a formal framework to model the fusion of gesture meaning with the meaning of the co-occurring verbal fragment. The framework is based on the formalization of two simple concepts, intersectivity and iconicity, which form the core of most descriptive accounts of the interaction of the two modalities. The formalization is presented as an extension of a well-known framework for the analysis of meaning in natural language. We claim that a proper formalization of these two concepts is sufficient to provide a principled explanation of gestures accompanying different types of linguistic expressions. The formalization also aims at providing a general mechanism (iconicity) by which the meaning of gestures is extracted from their formal appearance." Giorgolo's work is clearly within the general framework of formal linguistics, and thus ought to have a growing influence in the years to come on the integration of gesture into formal syntax and semantics studies. Many of the ways in which I would like to see mainstream linguistics incorporate gesture research are already being pioneered by Giorgolo.

Another prominent researcher whose work has been heavily influenced by McNeill is his former student, Justine Cassell, at Carnegie Mellon Universtiy. Cassell's work on computational communication is ground-breaking. This research is summarized as follows (taken from http://www.hcii.cmu.edu/people/faculty/justine-cassell):

"This research includes "developing the Embodied Conversational Agent (ECA), a virtual human capable of

interacting with humans using both language and nonverbal behavior. More recently Cassell has investigated the role that the ECA can play in children's lives, as a Story Listening System (SLS): peer support for learning language and literacy skills. And Cassell has also employed linguistic and psychological analyses to look at the effects of online conversation among a particularly diverse group of young people on their self-esteem, self-efficacy, and sense of community.

Once machines have human-like capabilities, can they be used to evoke the best communicative skills that humans are capable of, the richest learning? This is the goal of Cassell's research: to develop technologies that evoke from humans the most human and humane of our capabilities, and to study their effects on our evolving world."

I have often thought that one of the reasons that computers will never "have language" in the human sense is because they cannot gesture. The work of Cassell and her team remind me of the danger of using the word "never." On the other hand, what is most crucial to language is something that Searle (1983) calls the "background" and Everett (2012a) simply calls "culture." It is not clear to me how anyone can overcome the cultural acquisition problem for computers, especially if knowledge is largely embodied rather than represented (Everett in preparation a). This is one reason why fiction such as the movie *Her* seems so highly implausible to me. But Cassell's work and research like hers could, though perhaps not in my lifetime, show that this skepticism is also misplaced.

The final researcher I would like to mention is doing research that on the surface may appear orthogonal to McNeill's work. Yet this research further extends our understanding of the multimodal nature of human language that McNeill's work has pioneered. Dr. Jennifer Green, at The University of Melbourne, works on Aboriginal sand stories. Her summary (Green (2014, i and 1ff)) is worth citing at length because of the originality and significance of her focus of multimodal research for linguistics more broadly.

"Sand stories from Central Australia are a traditional form of Aboriginal women's verbal art that incorporates speech, song, sign, gesture and drawing. Small leaves and other objects may be used to represent story characters. This detailed study of Arandic sand stories takes a multimodal approach to the analysis of the stories and shows how the expressive elements used in the stories are orchestrated together.

Speakers of the Arandic languages of Central Australia have a range of semiotic resources or 'systems' in their communicative repertoire. These include everyday language, spoken auxiliary languages, such as those used to encode respect for certain kin, sign language, the esoteric language of songs, and symbolic or graphic conventions used in sand stories and in various forms of Aboriginal art. Spontaneous gesture is also part of this complexity. In everyday communication it is the norm for several of these systems to coexist and be interdependent. The performance of Arandic

sand stories (called tyepety in some Arandic languages) is a traditional form of visual storytelling in which co-speech graphics form an essential part. A skilled narrator of these stories incorporates multiple semiotic systems and uses the potentials within these systems to great creative effect. Speech, sign, gesture and drawing are employed, in sequence and in unison. As well as drawing on the ground, narrators may also use a variety of objects to establish a visual field in front of them, somewhat like a miniature stage-set. Leaves or sticks are used to represent story characters, and other small items which come to hand may be used to symbolize objects that are part of everyday life, such as shelters, shades, windbreaks and fire pits. The use of the ground for illustrative and explanatory purposes is pervasive in the environment of Central Australia where there is ample inscribable ground, and this attention to the surface of the ground arises partly from a cultural preoccupation with observing the information encoded on its surface. As a story begins, a space on the ground is cleared for inscription, and extraneous leaf litter, prickles and other debris are removed. If the ground is hard, it is broken up with a stick or crowbar to soften it. Some narrators prepare a screen of wet sand overlaid with fi ne dry sand, as this provides a solid base for drawing and makes for greater contrast of the drawn graphic symbols with the earth. A stick or a bent wire may be used to draw with, to provide a rhythmic complement to the verbal narration, and to augment deictic gestures as the narrator orchestrates the action in and around the space in front of them. A narrator builds up layers of real and imagined spaces, using drawing, signs and gesture and by moving objects around the story space. The accumulation of graphic elements is periodically erased as the narrative unfolds, yet the palimpsest of previous drawing may be referred to anaphorically even after it is no longer visible. Erasure of the drawing space marks the beginnings and ends of stories and signals changes in time and space within stories (Munn 1973a: 69–72; Wilkins 1997b: 144). This device enables sequences of visible narrative action to be superimposed on top of one another in the same spatial plane. The resultant drawings are both product and process and they involve a complex interplay between dynamic and static elements. They leave a mark or an artifactual trace that can be ' read' or observed by an interlocutor for a short amount of time, yet this semi-permanence is subservient to broader rhetorical aims as the story unfolds."

Giorgolo, Casell, and Green are on the cutting edge of gesture and multi-modal aspects of language research that one hopes will gain momentum. All of this research demonstrates as clearly as one could ever hope how misguided it was in the early days of generative linguistics, the foundation of the majority of formal linguistics today, to set the sentence as the "start symbol" of the grammar, rendering it in effect the exclusive empirical domain of linguistics research. In Everett (2013) I refer to this as the "reification" of linguistics and discuss many ways in which the otherwise innocuous idealization of restricting analytical focus to sentences has arguably retarded progress in the understanding of human languages.

9. Conclusion: Gesture and linguistic theory

One's overall assessment of McNeill's work and the vast amount of research that he - along with others, e.g. Adam Kendon - has inspired must be unqualifiedly positive. Linguists should be grateful for the careful, painstaking, long-term research into the multimodal nature and origin of human language - research that hopefully will not continue to be ignored in debates on the evolution, use, and structuring of human languages.

Perhaps the greatest lesson from gesture research is that it settles - to my mind at least - the debate of whether language evolved for mental life or for communication. If McNeill is correct about the role of the growth point in the evolution of syntax, for example, then the preeminence of communication as the key to language evolution (over expression of thought) is unavoidable. In this regard, his arguments that compositional meaning was not the "sufficient leap" necessary to provide hominins with language, that language is both static and dynamic, form and image are essential. His theoretical understanding of the role of the Growth Point, the theory of mind, and Mead's Loop in human language are or ought to be transformational for the discipline.

One can think of many ways in which gesture research should be incorporated into modern linguistic theory. First, gesture studies must come to play a significant role in field research, not least in field research of endangered languages. Everett (2004) and Everett & Sakel (2012), as well as most other fieldwork textbooks miss this point, to our loss. Second, linguists must acknowledge that there is simply no theory of language or language origins in the absence of a significant component dedicated to gestures.

The study of gestures is crucial to the understanding of the origin of language, the proper conduct of field research on language, the relationship between language and grammar (for any theory of grammar), the psycholinguistics of language acquisition and use, the relationship of language to culture, and the neurological bases of human language.

However one ultimately defines technical terms such as "core grammar," "I-language," and so on, gestures are a necessary component of our understanding. After reading McNeill's work I cannot in good faith allow myself to talk about a "theory of syntax," much less a "theory of language" that does not include gesture studies and theory. From McNeill's work, among others, we also learn that understanding gesture is important for theories of psycholinguistics, neurolinguistics, linguistic anthropology, and so on.

Another way in which gesture studies come to bear on fundamental issues of human language and cognition is their relevance to what has come to be known as "embodied cognition" (Gibbs, 2005;

Chemero 2011; Lakoff and Nuñez 2001; etc.) For example, in a recent report on research at the University of Chicago, McNeill's home institution, it is reported that the use of gestures, i.e. embodying cognition, can contribute to cognitive acquisition of concepts as difficult as mathematics (http://news.uchicago.edu/article/2014/03/10/gesturing-hands-powerful-tool-children-s-math-learning).[10]

The take-home from this review is simple. McNeill's trilogy and gesture studies should be part of the training of every linguist. I am very glad to have had the opportunity to read them and to reflect on them at some length.

References

Andrén, Mats. 2010. Children's Gestures from 18 to 30 Months, Centre for Languages and Literature; Centre for Cognitive Semiotics, Lunds Universitet, Sweden.

Arbib, Michael A. 2005. "From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics," *Behavioral and Brain Sciences*, 28:1-5-124.

Arbib, Michael A. 2012. How the Brain Got Language: The Mirror System Hypothesis, Oxford University Press.

Bloomfield, Leonard. 1933. Language, Holt, Rinehart, and Winston, New York.

Chemero, Anthony. 2011. Radical Embodied Cognitive Science, MIT Press, Cambridge, MA.

Corballis, Michael C. 2002. From Hand to Mouth, Princeton: Princeton University Press.

Duncan, Susan. 2002. "Gesture, verb aspect, and the nature of iconic imagery in natural discourse." Gesture 2:2, pp 183-2006.

Duncan, Susan. 2006. Co-expressivity of Speech and Gesture: Manner of Motion in Spanish, English, and Chinese. Berkeley: Proceedings of the 27th Berkeley Linguistic Society Annual Meeting.

Everett, Daniel L. 1988. 'On Metrical Constituent Structure in Pirahã Phonology,' Natural Language and Linguistic Theory 6, pp 207-246.

Everett, Daniel L. 1994. 'The Sentential Divide in Language and Cognition: Pragmatics of Word Order Flexibility and Related Issues,' The Journal of Pragmatics & Cognition, 2:1, pp 131-166.

Everett, Daniel L. 2004. "Coherent fieldwork." In **Linguistics today** (ed.) P. van Sterkenberg, 141-62. Amsterdam: John Benjamins.

Everett, Daniel L. 2005. "Cultural Constraints on Grammar and Cognition in Pirahã: Another Look at the Design Features of Human Language." Current Anthropology, 76: 621-646.

Everett, Daniel L. 2008. Don't sleep, there are snakes: life and language in the Amazonian jungle. New York: Pantheon.

Everett, Daniel L. 2009a. "Wari' Intentional State Construction Predicates," In: Robert Van Valin (ed.), Investigations of the Syntax-Semantics-Pragmatics Interface, John Benjamins Publishing Company, Amsterdam.

Everett, Daniel L. 2009b. "Pirahã Culture and Grammar: A Response to Some Criticisms." Language, 85:405-442.

Everett, Daniel L. 2010. "You drink. You drive. You go to jail. Where's recursion?" http://ling.auf.net/lingbuzz/@tCyuITrPicSKpoXb/wbOhgtZe?137.

Everett, Daniel L. 2012a. Language: The Cultural Tool, New York: Pantheon Books.

Everett, Daniel L. 2012b. 'What does Pirahã Have to Teach Us About Human Language and the Mind?' WIREs Cogn Sci 2012. doi: 10.1002/wcs.1195.

Everett, Daniel L. 2013. 'A Reconsideration of the Reification of Linguistics.' September 26, The Cognitive Revolution, 60 Years at the British Academy, London

Everett, Daniel L. 2014. "Concentric Circles of Attachment in Pirahã: A Brief Survey" In: Keller, Heidi and Hiltrud Otto, eds. Different Faces of Attachment: Cultural Variations of a Universal Human Need. Cambridge University Press.

Everett, Daniel L. 2014. 'The Role of Culture in Language Emergence' In: The Handbook of Language Emergence, Wiley-Blackwell, ed. by Brian MacWhinney and William O'Grady.

Everett, Daniel L. In preparation a. Dark Matter of the Mind: How Unseen Forces Shape our Words and

World, Chicago: University of Chicago Press.

Everett, Daniel L. In preparation b. The Origin of Language (working title), Liveright-W.W. Norton, New York.

Efron, David. 1972 [1942]. Gesture, Race and Culture a Tentative Study of the Spatio-temporal and "Linguistic" Aspects of the Gestural Behavior of Eastern Jews and Southern Italians in New York City, Living Under Similar as well as Different Environmental Conditions, DeGruyter Mouton, Berlin.

Fitch. W. Tecumseh. 2010. The Evolution of Language, Cambridge University Press, Cambridge.

Gibbs, Raymond G. 2005. Embodiment and Cognitive Science, Cambridge University Press, Cambridge.

Lakoff, George and Rafael Nuñez. 2001. Where Mathematics Come From: How The Embodied Mind Brings Mathematics Into Being, Basic Books, New York.

Shapiro, Lawrence. 2010. Embodied Cognition, Routledge, London.

Giorgolo, Gianluca. 2010. A Formal Semantics for Iconic Spatial Gestures, in M. Aloni, B. Harald, T. de Jager and K. Schulz 'Logic, Language and Meaning, Springer', p. 305-314, 2010.

Givon, Tom. 1983. Topic Continuity in Discourse, A quantitative cross-language study, John Benjamins.

Goldberg, Adele. 1995. *Constructions: A Construction Approach to Argument Structure.* University of Chicago Press.

Harris, Zellig. 1951. Methods in Structural Linguistics, University of Chicago Press, Chicago.

Hauser, Marc, Noam Chomsky, and Tecumseh Fitch. 2002. "The faculty of language: what is it, who has it, how did it evolve?" Science 298: 569-1579.

Hewes, Gordon W. 1973. "Primate communication and the gestural origins of language. Current Anthropology 14: 5-24.

Iverson, Jana M. & Goldin-Meadow, Susan. 1997. What's communication got to do with it? gesture in congenitally blind children. Developmental Psychology 33: 453-467.

Kendon, Adam. 2004. Gesture: Visible Action as Utterance, Cambridge: Cambridge University Press.

Kinsella, Anna R. 2009. Language Evolution and Syntactic Theory, Cambridge University Press, Cambridge.

Kita, Sotaro. 2000. How representational gestures help speaking. *Language and gesture*, 162-185.

Ladefoged, Peter, Jenny Ladefoged, and Daniel L. Everett. 1997. 'Phonetic Structures of Banawá, an Endangered Language,' Phonetica 54: 94-111.

Lenneberg, Eric. 1967. Biological Foundations of Language, John Wiley and Sons, New York.

Lieberman, Philip. 2007. The Evolution of Human Speech: Its Anatomical and Neural Bases, Current Anthropology Volume 48(1):39-66.

Longacre, Robert E. 1964. Grammar Discovery Procedures. The Hague: Mouton and Co.

Mead, George Herbert. 1974. Mind, Self, and Society from the Standpoint of a Social Behaviorist, University of Chicago Press. (A collection edited and with an introduction by C.W. Morris)

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

McNeill, David. 2005. Gesture and Thought. Chicago: University of Chicago Press.

McNeill, David (ed.). 2000. Language and Gesture. Cambridge: Cambridge University Press.

McNeill, David. 1992. Hand and Mouth: What Gestures Reveal About Thought. Chicago: University of Chicago Press.

Pike, Kenneth L. 1967. Language in Relation to a Unified Theory of the Structure of Human Behavior, Mouton.

Pike, Kenneth L. 1945. The intonation of American English. University of Michigan Publications in Linguistics 1. Ann Arbor: University of Michigan Press. xi, 200 pp.

Pike, Kenneth L. 1943. Tone languages: The nature of tonal systems, with a technique for the analysis of their significant pitch contrasts. Glendale, Calif.: Summer Institute of Linguistics, 121 pp. [Rev. ed. (1945), xi, 159 pp.]

Polanyi, Michael. 1974. Personal Knowledge: Towards a Post-Critical Philosophy, University of Chicago Press, Chicago.

Polanyi, Michael. 2009. The Tacit Dimension, University of Chicago Press, Chicago.

Reinbold, Julia Ulrike. 2004. Intonation and Information Structure in Banawá, unpublished M.A. Thesis, Manchester University, Linguistics.

Rizzolatti, Giacomo and Michael A. Arbib. 1998. Language within our grasp, Trends Neuroscience. (1998) 21, 188–194.

Sakel, Jeanette and Daniel L. Everett. 2012. Linguistic Field Work: A Student Guide, Cambridge University Press, Cambridge.

Schegloff, Emanuel A. 1984. On some gestures' relation to talk. In J.M. Atkinson and J. Heritage (eds.) *Structures of Social Action*, pp 266-298. Cambridge: Cambridge University Press.

Searle, John. 1983. Intentionality: An Essay in the Philosophy of Mind. Cambridge University Press,

Cambridge.

Slatkin, Montgomery and Laurent Excoffier. 2012. "Serial Founder Effects During Range Expansion: A Spatial Analog of Genetic Drift," *Genetics*. May 2012; 191(1): 171–181.

Skipper, Jeremy I, Susan Goldin-Meadow, Howard C. Nusbaum, and Steven L. Small. 2009. Gestures Orchestrate Brain Networks for Language Understanding. **Current Biology,** Volume 19, Issue 8, 661-667.

Thomason, Sarah Grey. in press. Does Language Contact Simplify Grammars?. In Gerd Hentschel and Gunter Spiess, eds., Pro- ceedings of the Arbeitstagung zum Thema Deutsch-slawischer Sprachkontakt, Entlehnungen, und sprachliche Identit¨at, Cottbus, Germany, September 2002.

Vygotsky, Lev S. (1978). Mind in society: The development of higher psychological processes. Cambridge: MA: Harvard University Press.

Wells, Rulon S. 1947. "Immediate Constituents." *Language*: 23. pp. 81–117.

Wilkins, David P. 1999. Spatial deixis in Arrernte Speech and Gesture: On the analysis of a species of composite signal as used by a Central Australian Aboriginal group. Paper 6, pp.31-45, in Proceedings of the Workshop on Deixis, Demonstration and Deictic Belief in Multimedia Contexts, held on occasion of ESSLI XI, ed. by Elisabeth André, Massimo Poesio & Hannes Rieser. Workshop held in the section 'Language and Computation' as part of the 'Eleventh European Summer School in Logic, Language and Information', August 9-20, 1999, Utrecht, The Netherlands.

Zlatev, Jordan. forthcoming. The Emergence of Gestures, In: Brian MacWhinney and William O'Grady, eds. The Handbook of Language Emergence, Wiley, Hoboken, NJ.

[1] This is one reason that computers cannot be said to have language, in my opinion. For others, see Everett (2012).

[2] Sign languages, as McNeill makes clear, represent a different case of combining static and dynamic gestures and diachronic as well as synchronic formation.

[3] Many researchers have noted, including McNeill (see below), that prosody is one part of speech that is gesture-like.

[4] http://www.nytimes.com/2013/07/01/world/europe/when-italians-chat-hands-and-fingers-do-the-talking.html

[5] I was perhaps incorrect in asserting that separate domains of study were called for to understand the dynamic vs. static. That remains to be seen.

[6] This term is nonintuitive to me however hard I try to absorb it and yet it is the term used by the field, so there is no way around it.

[7] As Sascha Griffiths pointed out to me (email May 04, 2014) the gesture/speech equiprimordiality hypothesis avoids the nagging question of how long after we had gesture languages the human larynx would have waited to descend (see Lieberman (2007)).

[8] It is also worth noting that if McNeill's story is on the right track, the contribution of Merge to the evolution of language was neither necessary nor sufficient.

[9] It is quite a different thing, however, to claim innocuously that language and gesture must find an explanation ultimately in evolution and to claim some much more specific biological underpinnings of linguistic structures - what these days is very loosely referred to as "biolinguistics," in which biology is never (rarely at best) causally implicated in any analysis, except in two ways. First, there is reasoning that usually goes like this: "Here is a fact that seems to be unlearnable. Therefore it must be innate." And second, "Only humans are capable of language so language must be an outgrowth of human biology in some interesting way." These arguments are discussed at length and rejected in, inter alia, Everett (2012). Nevertheless

language clear draws on physical and mental characteristics that have developed in hominins differently than in other animals.

[10] My one remark on the actual publishing mechanics of the books is that their indexes are all poor. The index to the final book, McNeill (2012) in the trilogy, is laughably bad. Many times in writing this review I searched for topics that I knew were in the book only to find no reference to them in the indices. This was very irritating and inconvenient. Overall, though, since the first two books are published by the University of Chicago Press and the third by Cambridge University Press, the printing, binding, etc. were of the highest quality. However, McNeill (2004) was to my mind the best-bound volume of the lot.