



Published in final edited form as:

Biol Theory. 2014 September ; 9(3): 296–308. doi:10.1007/s13752-014-0186-7.

On Quantitative Comparative Research in Communication and Language Evolution

D. Kimbrough Oller and

School of Communication Sciences and Disorders, University of Memphis, Memphis, TN, USA, & Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

Ulrike Griebel

Institute for Intelligent Systems, The University of Memphis, Memphis, TN, USA

D. Kimbrough Oller: koller@memphis.edu

Abstract

Quantitative comparison of human language and natural animal communication requires improved conceptualizations. We argue that an infrastructural approach to development and evolution incorporating an extended interpretation of the distinctions among illocution, perlocution, and meaning (Austin 1962; Oller and Griebel 2008) can help place the issues relevant to quantitative comparison in perspective. The approach can illuminate the controversy revolving around the notion of functional referentiality as applied to alarm calls, for example in the vervet monkey. We argue that referentiality offers a poor point of quantitative comparison across language and animal communication in the wild. Evidence shows that even newborn human cry could be deemed to show functional referentiality according to the criteria typically invoked by advocates of referentiality in animal communication. Exploring the essence of the idea of illocution, we illustrate an important realm of commonality among animal communication systems and human language, a commonality that opens the door to more productive, quantifiable comparisons. Finally, we delineate two examples of infrastructural communicative capabilities that should be particularly amenable to direct quantitative comparison across humans and our closest relatives.

Keywords

Animal communication; Austin; Communication; Illocution; Language evolution; Perlocution; Referentiality; Semantics

In the discussion of human uniqueness, language has played a central role for centuries (e.g., Condillac 1756). In recent time, scholars have sought language origins both in development and in evolution, and have been led to consider not only our uniqueness but also our common roots with other animals. While language in a strict sense is only human, animal communication is thought to share characteristics with language. But precisely what *are* the

Dedication

This paper is dedicated to the memory of Michael Owren with whom the work was jointly conceived. He died before contributing to the writing, but his numerous insights influenced this work deeply and will remain for us an enduring legacy of thoughtful interpretation and scientific integrity.

characteristics? A conceptual framework is needed to determine and quantify the extent to which features of language are shared across species.

Instead of quantitative comparison, a checkmark approach has often been taken (e.g., Hockett 1960). For example, one might check off for language a characteristic of “syntax,” while natural ape communication systems would receive no check. But repeatedly, claims that humans uniquely possess a communicative characteristic are challenged empirically—on syntax, for example, various papers have argued that natural “call sequences” of non-human primates sometimes convey “meaningful information” beyond the sum of the functions of the individual calls (Snowdon 1982, 1990; Arnold and Zuberbühler 2006a; Clay and Zuberbühler 2009, 2011). This claim has recently been argued to reveal no insight about language origins since a bacterial species also shows signal sequences that “achieve an effect that is different to the sum of the effects of the component parts” (Scott-Phillips et al. 2014). But the issue of animal syntax seems certain to remain on the table given the momentum of animal communication research on call sequences, and more importantly, given that apes trained in artificial communication systems have been shown to produce more human-like sequences of communicative units than in their natural communication (e.g., Greenfield and Savage-Rumbaugh 1990). Yet, in general the extent of the resemblance across species is left entirely unquantified.

We seek more precise definitions for characteristics of communication systems, definitions that will afford measurement of the extent of command for each characteristic within each species and across levels of development. Our approach synthesizes and extends conceptualizations from prior efforts, and specifies communicative characteristics as communicative *capabilities*. The approach is infrastructural, specifying naturally logical sequences, where certain capabilities form necessary initial foundations (infrastructure), that appear early in human development, are shared to a greater extent across species than other capabilities that appear later in human development, and may not occur at all in other species in the wild (Oller 2005). By differentiating among capability types in terms of the infrastructural natural logic, we hope to provide a basis for more productive and quantitative comparison across species.

This article will (1) highlight the central current controversy on the relation between human and animal communication, a controversy revolving around the notion of “functional referentiality”; (2) delineate characteristics of newborn human cry that might be deemed to show functional referentiality according to the criteria typically invoked in animal communication literature; (3) clarify key concepts needed to illuminate the controversy, and point the way to more productive cross-species comparisons; (4) illustrate how the key concepts clarify relations among communicative events, and thus clarify interspecies and developmental differences; (5) describe a scenario in which “perlocutionary effects” of signals can lead to a false impression of referentiality; (6) explore the essence of the notion “illocution,” illustrating an important realm of commonality among animal communication systems and language; and (7) delineate two examples of infrastructural communicative capabilities that should be particularly amenable to direct quantitative comparison across humans and our closest relatives.

The Central Controversy About Animal Communication and Language: Functional Referentiality

In the most prominent recent attempt at cross-species communication comparison, animal communication is portrayed as related to human communication through “functional referentiality,” the postulated ability of some animals to produce signals in a way that allows other animals to react *as if* they have been informed of the presence, for example, of a particular type of predator or food (Macedonia and Evans 1993). Research from a variety of species has been reported to support this view (Struhsaker 1967; Seyfarth et al. 1980; Gozoules et al. 1995; Crockford and Boesch 2003; Zuberbühler 2003; Clay and Zuberbühler 2009). The work seeks to analogize animal signals (at least in some cases) to words with referential “meaning.”

In a countervailing trend, animal communication is portrayed as less like language, and the notion that animal signals transmit referential information is resisted (Owren, Rendall, & Ryan, 2010). Instead, the proponents favor a management- or influence-oriented approach where animal communication consists of signals that influence the behavior of others rather than informing them of events or states in the world (Owings and Morton 1998). This approach emphasizes communicative adaptations through affect induction, affect conditioning, and associative learning (Owren and Rendall 2001; Rendall et al. 2009). Instead of constituting encoded references (as words often do), animal signals are seen from this viewpoint as reflecting variable states of senders and as inducing variable states in receivers. Appearance of referentiality can then (at least sometimes, perhaps always) be viewed as being based on affective responses or conditioned reactions. In addition, it has been argued that the appearance of referentiality might sometimes occur as an *indirect* consequence of affect induction—an animal might for example hear a fear-inducing sound and then look for danger, reacting thereafter in a way that shows adaptation to the particular type of danger (e.g., predator). In this view, information about a particular type of food or predator is not encoded *in* the signal, but rather signaling initiates a process of reaction in the receiver that may be purely induced or may be adapted by the receiver to reflect current and/or prior correlations with affairs in the world.

A recent review of these contrasting views emphatically rejects the functional referentiality perspective, but maintains the contention that information may be derived by animal receivers from animal signals and that decisions of receivers may be guided by that information (Wheeler and Fischer 2012). Nonetheless, the review continues to utilize the term “meaning” to designate the information assumed to be derived by animal receivers. The review specifically seeks to rule out the possibility that conditioning might account for the complex reactions of receivers to varying signals and contexts of use, i.e., in deriving meaning. The authors conclude that “the balance of current evidence favors [the] informational perspective” (p. 202), but grants the need for empirical testing to determine its viability as opposed to the “affect-conditioning” perspective they attribute to Rendall, Owren and colleagues, where animal learning of how to respond to signals would presumably involve less explicit awareness of correlations on the part of learners, and would instead involve direct learning of reactions that work to learners’ benefit. The review

concludes that empirically choosing between these alternatives “will be no easy task” (p. 203). In our opinion, the differences between them may in fact be illusory, because conditioning need not be exclusively affective, but can also constitute learning of correlations among circumstances, such as a call and a type of disturbance—the key objection to the notion of referentiality in animal communication by Owren and Rendall is that it implies information encoded *in* signals, an objection Wheeler and Fischer appear to share.

An additional suggestion of Wheeler and Fischer is that improvements in how we view animal communication may come from better integration of insights from linguistic pragmatics (Scott-Phillips 2010). We agree. A productive line of reasoning based on extensions of linguistic pragmatics has existed for some time in work on human infant development (Dore 1975; Bates et al. 1979; Bretherton 1988; Oller and Eilers 1989). Adaptation of these insights for animal communication has also been pursued (Oller and Eilers 1992; Oller 2000; Griebel and Oller 2008). The potential clarifications that may be possible based on evaluation of human infant development can be illustrated in a familiar example—newborn infant cry, which has been portrayed as an “animal communication system” (Owings and Zeifman 2004).

Newborn Human Cry and the Notion of Functional Referentiality

The newborn human cry is a paradigm example of a naturally-selected signal, a distress or discomfort display that caregivers react to by trying to comfort or nurture infants in widely different socio-cultural circumstances (Green et al. 1987; Barr et al. 1991; Zeifman 2001). Both senders (infants) and receivers (nominally caregiving relatives) can be imagined to benefit from cry communication, infants obtaining necessary care and related caretakers protecting their investment, both ultimately promoting their genes. The term “intention” is generally not invoked with regard to newborn infant cry, and yet caregivers derive complex information in making decisions about what to do when infants cry, information that could be argued, using reasoning common in animal communication literature, to be functionally referential.

A cry may be more or less intense, and caregivers may treat it as more or less urgent, assuming the more urgent-sounding cry is associated with pain, whereas the less-urgent is associated with hunger or mere discomfort (Meadows et al. 2000; Soltis 2004). Caregivers often engage in verbalized reasoning about how they interpret cry as they react to it by picking infants up, by changing their diapers, or by feeding them. It appears to be important that caregivers react precisely, since infant well-being in the second year has been found to correlate positively with how accurately mothers recognize and respond to the quality of their infants’ crying in the first year (Lester et al. 1994). Thus there are high stakes to systematic correspondence between how parents react and how babies cry, just as there are high stakes in how animals react, for example, to alarm calls.

By 3–6 months, the human infant cry is more clearly differentiated, and caregivers recognize a distressful cry reflecting physical discomfort along with a more intentional form of cry that appears to express frustration or anger and often appears to be directed at getting social

attention, typically evidenced by eye contact and reaching toward the caregiver (Papoušek and Papoušek 1984; Gustafson and Green 1991; Lester and Boukydis 1992; Chen et al. 2009). Both in the newborn stage (where we assume no intentionality on the part of the infant) and in the middle of the first year (where the infant seems to have learned to use cry differentially, sometimes directed toward a caregiver instrumentally), the information received through cry events would not be thought of by researchers in child language as resembling a word or a sentence—specifically it would not be thought to manifest referentiality. The cry does not *refer* to the infant state or intention, but rather *expresses* it, i.e., provides a performance or display of that state or intention. Crying babies do not say anything, such as “I am hungry” or “there is a pin sticking in my elbow” nor do they give any command such as “change me” or “pick me up”. And yet parents may act as if there has been information transmitted by the intensity or abruptness of the cry or by some circumstance that accompanies it—a long period since changing or feeding, the fact that the infant has just fallen onto the floor, etc.

Here then is a challenge for the interpretation of animal signals as potentially referential or meaningful: Provide empirical justification for the claim that animal signals transmit more language-like information than that which is transmitted by the newborn human cry. The evidence does not appear yet to irrevocably reject the possibility that the same kinds of processes for derivation of information by receivers and the same kinds of limitations on information *in* the signal are present in natural animal communication events and in newborn human cry. In neither case, we shall argue, does it appear justified to assert on the basis of current evidence that referential meaning is present *in* the signal. The situation changes, of course, for human infants by early in the second year, when word usage and understanding suggests true referentiality, implying that information is indeed encoded in the child vocabulary (Bates et al. 1979). Such referentiality has also occurred in some cases of human-trained animals, especially apes that have learned to communicate with signs or other visual symbols (Premack 1971; Savage-Rumbaugh and MacDonald 1988; Gardner et al. 1989). Below we provide details about capabilities implied by true referentiality.

We think the difficulties of comparing communication across species are fundamentally definitional, and a new framework of comparison is needed that can simultaneously 1) portray the nature of communication that is possible for animals in the wild and for human newborns, and 2) show crisply how true language, even words spoken by children early in the second year, depend on capabilities not required in cry or in natural animal communication. In our proposal, the notions “meaning” and “reference” will be situated in the perspective of several foundational aspects of communication required prior to the possibility of encoding meaning/reference in a signal. In the context of the new framework, better comparison possibilities, including quantification, should emerge.

Key Concepts to Illuminate Quantifiable Characteristics of Language

What is Communication?

Animal researchers often portray a communicative event as an action of one organism that influences the behavior (or state) of another. Under this style of thinking, any action can result in communication. In a more targeted form of this definition, an action influencing

another's behavior is required to have been naturally selected or "specialized" (Hockett 1960; Hockett and Altmann 1968; Owren, Rendall, and Ryan, 2010) to produce an influence in order for it to be considered communication—in this case the action is termed a "signal" whereas in the absence of specialization for communication, the action can be termed a "cue" (Maynard Smith and Harper 2003).

In all approaches we know of, a communicative event is assumed to consist of a signal and a function. In machine communication, these can be related in a strict one-to-one fashion, and consequently the distinction between signal and function may be thought to be of no major importance. But animal communication is not so simple, even though the classical ethologists portrayed animal communication as consisting of "fixed signals" with each signal related one-to-one with a particular function—e.g., a threat call with an aggressive function (Lorenz 1951; Tinbergen 1951). Whereas this conception seems appropriate in many cases including newborn human cry, it remains necessary to account for the fact that observer reactions to a particular signal type can systematically differ on different occasions (a mother may offer care when an infant cries, but in different ways each time, and in some cases she may ignore the cry). Furthermore, some animal signals (and human cry) and their relations to functions are modulated by learning and development. Consequently, in animal communication (unlike in machine communication), a distinction between signal and function can always be empirically drawn and is not merely a philosophical nicety.

It has also long been recognized that some acts that modify behavior of others do so in some sense accidentally, not having been naturally selected *as* communicative acts (Otte 1974; Condillac 1756; Peirce 1934). For example, sneezing is assumed to have a basic role in airway function independent of communication. Yet a human caregiver, hearing an infant sneeze, may react by seeking to determine if the infant is well, and may follow up, for example, by trying to remove a possible allergen. Such transmission has been termed "indexical" (Peirce 1934). In recent animal communication literature, a sneeze or any other purely indexical act, would be deemed a "cue" rather than a "signal" (Maynard Smith and Harper 2003).

An infant cry would by any standard be deemed a "signal," because the cry is assumed to have been shaped by natural selection *as* a signal, that is, to have been "specialized" (Hockett 1960; Hockett and Altmann 1968) for the communication of distress. The aspect of any communication that results from reactions of a listener independent of naturally selected features of the signal and its typical function can then be said to be indexical. Notice that even acts that have been specialized by natural selection for communication still also bear indexical potential—a cry, for example, may *index* (or cue) the infant's identity in addition to *signaling* the infant's distress.

Adaptation of the Austinian Distinctions, Keys from Linguistic Pragmatics

Much recent discussion in the animal behavior literature has focused on the distinction between the potential content of animal signals themselves (implying that signals may have functions of their own at the level of the sender) and the reactions of other animals to those signals (implying that the effects of signals on receivers reflect a different kind of function, even if effects on the receiver are loosely predictable based on function at the level of the

sender). The study of human infant communication has also utilized a distinction between the function of the signal as produced by the sender and the function as interpreted by the receiver (e.g., Bates et al. 1979). Such literature typically references Austin's (1962) speech act theory, where the import of a signal can have sharply distinct levels. We propose that judicious adaptation of the Austinian distinctions can lay a fruitful foundation for the comparative endeavor, clarifying prior confusions, and providing much clearer concepts for cross-species comparison.

Consider a threefold distinction among Perlocutionary Effect, Illocutionary Force, and Meaning (hereafter capitalized), ideas modified significantly, but inspired by the insights of Austin. To be clear, Austin did not to our knowledge ever write about animal communication nor about communicative *development* in humans. This is one reason that modifications are necessary—to adapt the Austinian insights about mature language to a broader scope of communication circumstances and degrees of complexity. Austin *did* notice and specify certain structural characteristics of language that had (amazingly) eluded prior linguists and philosophers of language entirely. The follow-ups to publication of *How to Do Things with Words* (Austin 1962), especially by Searle and colleagues (Searle 1969; Searle and Vanderveken 1985), generated plenty of reaction among philosophers, although much of it has been negative, especially recently (e.g., Burkhardt 1990; Dörge 2004). But these philosophical controversies are not particularly relevant to the application and interpretation of Austin's triad as we shall propose it here. So allow us to boldly go where Austin appears to have pointed the way, but where formal philosophers have apparently not yet gone.

The terms Perlocutionary Effect, Illocutionary Force, and Meaning are intended to apply to communications that occur through (short-term) actions rather than through (long-term) states of being, and the focus here is primarily on *vocal* actions. Roughly, Perlocutionary Effects are reactions of receivers to signals, Illocutionary Forces are functions transmitted by senders in the act of signaling, and Meanings consist of truly referential information transmitted in signaling. It should be emphasized that these distinctions refer not to characteristics of signals as signals, but to types of functions that signals may serve.

A Perlocutionary Effect (or Perlocution) pertains to the reaction of receivers to a sender's signal, whether or not receivers interpret the signal as intended by the sender. Such reactions play a key role in natural selection of signals, where we include (1) the kind of selection that can occur across deep time, producing in senders the capability and/or the inclination to produce a particular signal in particular circumstances or states and thus to serve particular communicative functions, with the process of evolution yielding species-specific signals and accompanying functions, as well as (2) selection by learning, where a signal may be modified through experience of an individual, and this includes experience with usage of signals within a community, including the case of language, where signals can be "words" that are learned in each new generation of a community, again leading to the use of particular signals in order to serve particular functions. Perlocutionary Effects can be said to operate as the most important *selecting force* that shapes differentiated signals.

The notion of Perlocutionary Effect allows emphasis on the fact that receivers play a distinct and important role in communication—that we cannot assume we know what a communication constitutes merely by examining the actions of a sender. Receivers may react in different ways to an individual signal event, and this differentiation can provide a basis for natural selection of differentiated usage, that is, adaptation of particular signals for diverse usage. It is widely recognized in animal communication literature that receivers have much greater flexibility in their reactions to signals than senders have in either production or usage of signals (Seyfarth and Cheney 2010).

The term Perlocution was invented by Austin to highlight its distinction from Illocution, but prior literature in animal communication (without invoking the notion of Illocution) has also taken stock of a distinction between receiver (Perlocutionary) effects and sender effects (Owings and Morton 1998). Nonetheless, for the distinctions between Illocutionary Force and Meaning, it is fair to say Austin provided a sharply new idea that has been adapted and utilized to positive effect for more than a generation in human infant and child language literature (Dore 1975; Snyder et al. 1981; Ninio and Snow 1986; Tomasello and Brandt 2009). Yet the significance and potential value of the distinction has scarcely been recognized in animal communication literature.

An Illocution, as we use the term, is an act performed *in* producing a communicative signal, in the here-and-now. In the following examples, we underline the terms we use to refer to Illocutions. If Manfred is lecturing and says “It’s hot in the classroom,” looking at Joe and then at the window next to Joe, Manfred may have performed a request, and Joe may open the window in response (a Perlocutionary Effect of the Illocution). Note that when Manfred says “It’s hot in the classroom”, he also performs a statement whether or not the utterance is intended to perform a request as well. In Austin’s usage, the utterance can be said to have the Illocutionary Force of request if the producer intends it as such, and our adaptation of Austin adds that it may or may not be interpreted as having that same Illocutionary Force by the listener—we must remember that Joe may not recognize that Manfred wants him to open the window, because the utterance does not say so directly. Thus even at the level of Illocution, there is a sender/receiver distinction to draw.

In each angle on this example there is also a strict distinction, in accord with Austin’s insight, to be drawn between the Illocution performed in the utterances and their Meanings. In Austin’s work, Meaning is one of three aspects of the “locutionary” act produced in the performance of an Illocution—we will not be making use of the other aspects, but instead focus exclusively on Meaning. Whenever we speakers of English say “It’s hot in the classroom,” we perform a statement, and that statement expresses particular semantic content or Meaning, but we do not necessarily perform a request, a criticism, or an explanation, each of which is an additional possible Illocutionary Force of the utterance. If Joe presents this utterance in his lecture as an example to help illustrate the distinction between Illocutionary Force and Meaning, he performs a statement when he produces the utterance, but rather than performing a request or criticism, he could be said to use the statement to perform an exemplification as part of an explanation.

Yet notice that in all these cases there is consistent semantic content expressed in the English utterance “It’s hot in the classroom” that is presented along with *all* the different Illocutionary acts. This content is the Meaning of the sentence, which invokes on every occasion of its usage, the idea of high temperature, of the occurrence of such temperature in a kind of room, in this case the particular room the speaker is in, along with some additional syntactico-semantic material. Notice that the utterance by itself, even if it is intended as a request, does not include meaningful content that specifies that it is a request, although we speakers of English could easily elaborate the sentence to make it so: “Since it’s hot in the classroom, I request of you, Joe, to please open the window.” And yet even this sentence can be performed to be an example and not a request, as it is done here, in the presentation of the written sentence.

How the Key Concepts Shed New Light on the Idea of Referentiality in Animal Communication and in Human Infant Communication

When the Austinian distinctions were first applied to very early human infant communication, they cut a sort of Gordian knot. They made it possible for the first time to distinguish linguistic utterances that transmit both Meanings and Illocutionary Forces from utterances that transmit Illocutionary Forces without bearing Meaning. Consider again the newborn infant cry. It is a performance or display of distress, but it does not say anything—not “I am unhappy,” “I am wet,” “I am hurting,” or “I am hungry.” That a parent may react by changing or feeding the infant, for example, is a Perlocutionary matter, which varies according to what the parent decides the problem might be. Even if newborn cries tend to be more intense with pain than with hunger (Green et al. 1987, 1995), and the parent can recognize that fact, it is still not true that the infant “says” anything by crying more or less intensely. Parents derive information from their evaluation of the whole context of what occurs when the infant cries. Of course by the second year of life, children may be able to whine or cry *while* actually saying “I’m hungry,” and in such a case, they both perform a distress display and make a statement that may be intended as a request. Still, independent of the Illocutionary Force intended or interpreted, the statement always transmits content with Meaning, because it directly refers with words in all those cases to a state (hunger) that we as users of English recognize as such every time the word is presented and that specifies further by using additional English syntactico-semantic mechanisms that the state of hunger is one that the speaker is currently experiencing, and critically for the nature of language, the utterance transmits all of this, whether or not the speaker is in fact hungry.

For the comparative endeavor, this distinction between Illocutionary Force and Meaning is in our view decisively important, because as far as we can tell, natural animal communication is a great deal like human cry, and has not been shown unambiguously to involve Meaning as we have defined it any more than cry has. Instead natural animal communication appears to consist of signals that transmit Illocutionary Forces in much the same way that newborn infant cry transmits Illocutionary Force. In some cases animal communications may transmit somewhat flexible Illocutionary Forces under the influence of learning, just as human infants also learn to use cry to serve differentiated functions as they mature across the first months of life (Prescott 1975; Lester and Zeskind 1982; Gustafson

and Green 1991). Meaning occurs in human communication only after significant developmental steps are taken that make it possible for the content of certain signals to be detached from the here-and-now, and to stand thereafter in flexible relation with the possible Illocutionary Forces that are performed when those signals are produced in the here-and-now, and after that, the content deserves to be called Meaning.

Meaningful human utterances, in accord with Austinian usage, typically involve referential content, that is, they include words or morphemes that refer to classes of entities, states, or conditions. Human infants by early in the second year of life clearly show the capability for reference, using learned words (Bloom 1970). Let us return now to the notion of functional referentiality that has been proposed to facilitate the argument that in some cases animals (especially vervet monkeys and other alarm callers) can produce signals resulting in a limited kind of reference. Let us unpack this idea in the context of the Austinian triad, assuming for the sake of argument that there is indeed a sharp distinction in vervet monkeys between alarm calls for leopards, snakes, and eagles.

The claim that these are alarm calls proves problematical since it has been acknowledged from very early publications claiming that they refer to particular predators that the calls are sometimes used in non-alarm inter- and intra-group interactions (Seyfarth et al. 1980). More recently the functions of these calls has encountered even more severe problems because it appears, after more recent data collection and analysis, that there is *no* sharp distinction among the usages of the vervet calls in alarm to predators as opposed to in intraspecific aggression encounters (Price 2013). In fact all the calls appear to be graded for usage across aggression and alarm circumstances, and all appear to be differentiated in significant ways as a result of differential arousal. Playback studies do show a tendency for particular receiver responses to correspond to the three signal types, but the responses are probabilistic.

Let us first address these signals in their roles as alarm calls. In that context, they can be said to perform alarm displays (Illocutions), but none of the calls offers the possibility (as far as has ever been reported or claimed) to serve, for example, as a performance of *pure* reference. For example, no vervet has been reported to simply invoke the *idea* of a leopard by producing the leopard call or to invoke the *idea* of a snake by producing the snake call. Yet this sort of naming is a hallmark of human infancy by 12–18 months of age (Bates et al. 1979; Gershkoff-Stowe 2002), and a key feature of any referential term in human language. Infants name things, even pictures of things, merely for the purpose of naming them. Their awareness and usage of words comes to be *detached* from any particular circumstance of usage, and to be more precise, their words show vast flexibility of function. By 18 months human infants can name an object such as a doll, request the doll by naming it, correct someone who teasingly calls the doll a dog, and so on. All these different Illocutionary Functions can be produced with the same word on different occasions, even though on each occasion that word invokes the same Meaning or semantic content, viz., the class of objects that we call “dolls” in English.

Meaning thus stands outside the time boundaries that delimit Illocutions, which are always performed *in* the communicative present. Meanings are conventionally paired with signals (words or other linguistic units) by learning through cultural transmission, and maintain their

relations with signals over the extended lifetime of those words or other linguistic units within a particular language, which can be hundreds of years. The key point is that while Meaning *is* transmitted *in* acts of language, the Meaning of any utterance is not *bound* to the time frame of the act but rather is detached from and extends far beyond any moment in time, while an Illocution is precisely that act that is *performed in the moment* of a communication and is bound to that time frame.

Here we see a fundamental way that vervet alarm calls do not appear to transmit Meanings. The three calls are differentiated by the types of predators that tend to elicit them, but it appears they do not and presumably *cannot* constitute events of pure naming. Rather they reflect correlations of call features with predator types that may have arisen through long-term evolutionary selection (as suggested by the occurrence of similar vocal types and functions in different vervet communities), although it appears that the usage of the calls is at least tuned through learning (Seyfarth and Cheney 1999, 2010; Price 2013). In fact, the extent to which these correlations have been naturally selected over deep time is not clear. In the following section we explore a possibility that has not been clearly excluded by the empirical evidence—that the correlations emerge from different interactions of vervets with different predators.

How Perlocutionary Effects can Masquerade as Decoding of Meaning and Yield the Appearance of Referentiality

Remembering that human and nonhuman receivers are able to derive information and interpret events with substantial flexibility, one can imagine how vervet monkeys might react to alarm calls (and the cluster of events that may accompany them) *as if* the Signals were referential even if there is no Meaning to decode. The following outlines a scenario consistent in part with prior reasoning (Rendall et al. 2009) and with the tradition (since Darwin 1872) interpreting animal signals as being driven by emotion and arousal.

When an alarm call is emitted, let us assume for the sake of argument that it is induced by a potentially complex emotion involving surprise/disorientation, fear, and/or aggressiveness. That all these possible emotions may be involved is suggested by reactions reported to occur in vervet senders and receivers when any of these calls is emitted. The animals are often reported to perform predator-specific escape behaviors, suggesting fear. But escape is only one of several possibilities. Often senders call while looking towards the perceived danger, and receivers simply orient in the direction of the signaler and scan the environment, seeming to investigate (Seyfarth and Cheney 1980), suggesting surprise/disorientation.

The fact that vervets often use these calls in *intraspecific* agonistic encounters (Price 2013) shakes the claim that these are predator-specific calls to its foundations. Clearly, intraspecific aggression encounters can involve both fear *and* aggressiveness.

With these thoughts about possible emotions involved in presumed alarm calling in place, let us consider that the acoustic nature of calls may be affected by how the sender responds *while* signaling. At least four factors could affect signal characteristics: (a) The degree of each of the emotions experienced, (b) the abruptness of their onset, (c) the immediate nature

of the signaler's orientation, escape or aggression maneuvers or preparations for them, and (d) a combination of a, b, and/or c. If there is any correlation between the type of predator and any of these features or any combination of them, then the presumed alarm signals produced in response to different predators could be acoustically different. These differences could be learned by listeners. Thus avoidance or orientation behaviors of receivers could tend to distribute differentially, *as if* receivers thought senders intended to designate particular predator types on each occasion of alarm calling, while in fact, senders merely made alarm calls that tended to be shaped by the differentiated experiences of interaction with and consequent reactions to different predators.

Another key fact that seems inconsistent with the notion of semantic reference in the presumed predator-specific vervet alarm calls is that adult male versions of the signals are notably different from those of females and juveniles (Price 2013). On the other hand, this difference makes perfect sense if we take into consideration that adult males in many species act differently than females and juveniles when engaged in aggression and or when expressing fear.

Perlocutionary effect could masquerade as semantic reference in a variety of ways. One is that receivers have been reported to look toward the first alarm caller (Seyfarth et al. 1980), providing a basis for imitation of his predator-appropriate escape maneuver. Receivers could also use sound localization to determine where the caller is and thus see or guess what the predator may be if there is one, and thus with higher than chance likelihood orient appropriately or make the most likely effective escape. Or prior to the first alarm call, receivers may have already experienced subtle cues (for example, by noticing that some animals in the group have been looking up or into the grass as if they may have detected something) that a particular type of predation is likely at the time, and may tend to react based on that subtle awareness as soon as the initial alarm call is emitted. Some of the animals in a group may use the orienting or escape maneuvers of others in the group (including the initial caller) as a cue to initiate a particular type of escape. More experienced members of the group may have learned to use these correlated cues more effectively than others, and may lead the orienting or escape maneuvers, being followed by less experienced members.

Playback experiments (Seyfarth et al. 1980) have attempted to limit the possibilities suggested by the reasoning embodied in our scenario, but they do not rule out learning by correlation in the absence of true referentiality. They only prove that there is some real correlation between signal features and typical listener escape responses—our reasoning suggests this correlation could be the result of natural or learned response biases of senders to different predators and learning of the resulting correlations by receivers. The recent work of Price (2013) supports the ideas of correlational learning along with emotion induction and arousal as determining how vervet calls function.

Similarly, playback experiments with parents using human infant cry show that parents differentiate their responses based on cry features (Green et al. 1987; Lester and Boukydis 1992; Lester et al. 1994). But this ability does not prove that reference to differing states is encoded in infant cry. The differing degrees of intensity or suddenness of distress can be

interpreted by the intelligent parent as reflecting different possible types of distress. Perlocutionary Effects can then be different based on the different interpretations.

That alarm calls are induced by some combination of surprise, fear, and perhaps aggressiveness seems very likely. This interpretation is bolstered by the fact that alarm calling by many members of a group of vervets has been reported to continue long after the danger of predation is gone and the predator has left the scene (Seyfarth and Cheney 1980)—suggesting that emotions subside gradually and that alarm calls are performances (Illocutions) expressing alarm or surprise/fear/aggression.

The reasoning we have presented here has been tailored to the reports we know of on vervets, but it should be possible to adapt them to offer similarly skeptical treatment of other cases where functionally referential alarm calling has been reported (e.g., Macedonia and Evans 1993; Fichtel and Kappeler 2002; Suzuki 2012). Even researchers who have tended to support the possibility of functional referentiality in animal communication have acknowledged that proving that it occurs is difficult (e.g., Arnold and Zuberbühler 2006b). Our opinion is that the more traditional view of animal communication as expression of varying emotion and arousal remains viable and that the insights afforded by the Austinian concepts help clarify an alternative interpretation to the evidence that has been purported to show semantics or Meaning in natural animal communication.

The Essence of the Notion Illocution

Illocutionary Forces are in our view “natural kinds” (Quine 1969). They are *the class of possible interactive types that tend to emerge in social systems*, each illocutionary event constituting a performance of one of those types, and each type being specific to the social system in question, which is to say each is produced by senders and interpreted by receivers in that social system. Illocutionary Forces emerge when social systems come to be complex enough for them to benefit its members, assuming the members have both the cognitive capacities and the physical signal production capacities to support the selection process. Each Illocution can be thus naturally selected to be performed as a signal/function pair, where Perlocutionary Effects constitute the selection mechanism that shapes signals to serve Illocutionary Functions. This reasoning is compatible with the fact that some signals function even *between* species—barks, hisses, or growls of one species can often be understood as aggressive even by distantly related species.

It seems predictable that primitive vocally expressed Illocutions in non-human primates include displays of distress, threat, warning, affiliation, submission, courtship, exultation, and contact calling. All these primitive Illocutions can be treated as monadic or dyadic, occurring as pure expressions of senders or as interactions between senders and receivers. Meaning, on the other hand, cannot occur until there are triadic communications that occur between sender and receiver *with respect to* some third entity, a referent or referent class. Joint attention between caregivers and human infants with respect to possible referents is seen in the latter part of the first year, before infants produce words (Seibert et al. 1982; Butterworth 1996). After meaningful words are learned, it becomes possible to perform new Illocutions such as naming, requesting specific objects or foods, or correcting the user of a

false name (Dore 1975; Bates et al. 1979). As sentences come into the picture, many new more complex Illocutions (description, promise, criticism, stipulation, explanation, and so on) become possible and continue to advance with further increases in social complexity and sentence structure.

We must of course seek to determine functions of animal communication in part by evaluating external contexts that accompany signals. Yet, there is good reason to draw a strong distinction between any Illocutionary Function and the external context(s) that may tend to generate its performance. An external context yields particular perceptions, and these can induce (naturally selected) particular states of emotion, which at threshold levels may produce (naturally selected) particular signals with corresponding Illocutions. But flexibility in action is adaptive, and consequently it makes sense that different intensities of an emotion, different degrees of abruptness in onset of an emotion, and different styles of reaction to particular situations generating an emotion could yield different versions of sound types, each signaling somewhat different states. These differences in signal can then be thought to be emergent properties molded by external context.

Rather different external contexts may generate functionally *similar* emotional states. Fear, for example, can be generated in a context of potential predation or in a context of intra- or intergroup aggressive interaction. Thus the same kind of signal and corresponding Illocution (fear display) can occur in widely different external contexts. Furthermore, very similar external contexts can generate different emotions on different occasions depending for example on the mood of the animal, and consequently the kinds of Illocutions that are produced may differ on different occasions of a particular context. Thus, while external context clearly must play a key role in how and when Illocutions are expressed, external context is not equivalent to Illocution, and the scientific study of contexts in which particular signals occur is only a beginning in the attempt to determine Illocutionary Function.

The notion function in animal communication literature is often (perhaps usually) taken to refer to the reaction of receivers, the Perlocutionary Effect rather than the Illocutionary Force. But the distinction between the two remains critically important. Perlocutionary Effects are reactions that serve as selection pressures on signals that transmit Illocutionary Forces. Positive selection pressure of Perlocution can only occur if receivers recognize signals and their Illocutionary Forces. Similarly, any naturally selected signal/Illocutionary pair must, on balance, yield Perlocutionary Effects that maintain or promote selection for that signal/Illocutionary pair.

Infrastructural Capabilities Suggesting Quantifiable Characteristics in Cross-species and Developmental Comparisons

The attempt to compare non-human primate communication with language by focusing on the possibility of reference has been very widespread, and continues to be a primary focus of research (see the recent volume of papers edited by Stegmann 2013). This attempt has been criticized for reasons related to those expressed above (Owren, Rendall and Ryan 2010), although the notion of Illocutionary Force has scarcely been brought to bear on that critique. Once it is recognized that animal signals can best be treated as transmitting Illocutionary

Forces rather than Meanings, a key commonality of animal communication and human communication comes into focus. All human language involves Illocution. Both human infants and nonhumans of various taxa transmit communications that include Illocutions. But neither very young human infants nor nonhumans (in the absence of training by humans) appear to transmit Meanings.

The question then becomes: What additional capabilities beyond those required for transmission of relatively simple Illocutions would an animal have to acquire in order to make it possible for signals with Meanings to be acquired? Such infrastructural capabilities are developed in the human infant in an orderly fashion in accord with a natural logic (Oller 2000) starting from the first months of life and culminating early in the second year, when words with Meaning appear. The earliest appearing of these infrastructural capabilities are, we argue, measurable and should be possible to compare sensibly in quantitative terms across species. We delineate two examples.

Copious Spontaneous Vocalization

The human infant engages in considerable amounts of vocalization in the first months of life. Notably, much of this is not cry or laugh, sounds that express negative and positive affect respectively. These special “protophones” such as squeals, growls, vowel-like sounds (vocants), or raspberries (Oller 1980) bear no specific emotion, and show no obvious necessary relation with external context. And yet they constituted more than 90% of all vocalizations produced at 3–4 months in a longitudinal study, where infants were recorded playing with a parent or while the parent was talking to another adult (Oller et al. 2013).

Notable features of protophones are their copious production (several per minute), and the fact that they often occur “spontaneously,” i.e., in the absence of social interaction, and in the absence of any obvious intent to obtain social attention (Franklin et al. 2013). The fact that there are differentiable types of protophones this early in life suggests that categories of vocal type self-organize from spontaneous vocal exploration as well as from the exercise of producing protophones in extensive social interaction where parents and 2–4 month old infants engage in face-to-face vocal interactions often lasting several minutes (Stern 1974; Bakeman and Brown 1977; Tronick et al. 1977; Tronick and Cohn 1989; Feldman et al. 1999). We reason that protophones thus form the anchor for vocal interaction, providing a basis for sustained, comfortable vocal dyadic interplay between parent and infant, a sort of interaction that as far we know occurs in no other primate.

This copious spontaneous vocalization in the human infant appears to provide an infrastructural platform for all subsequent vocal learning and thus for language. Copious spontaneous vocalization occurs in the first months, long before other characteristics of language that are also often deemed to provide critical infrastructure for word learning, viz. canonical syllable production, joint reference (or joint attention), vocal imitation of canonical syllables, associative learning of pairings between canonical syllables and circumstances, and a few others. Consequently we reason a particularly important quantitative comparison between humans and non-humans (especially great apes) could be based on the relative amount of vocal production (“volubility”), especially spontaneous vocal production. Surprisingly there has been no prior direct comparison of volubility in

humans and non-humans as far as we know, although it is often assumed that humans, even in early infancy, are far more voluble than great apes. The speculation is consistent with the possibility that high volubility in human infants reflects the exercise of a capacity that must occur on the path to language.

In our own laboratory we are already involved in direct comparative research on volubility in humans and non-human primates. This work, especially to the extent that comparisons can be made at early ages of development, may provide significant new perspectives on how humans and non-humans may differ in infrastructure for vocal language. A significant challenge is to determine optimal points of volubility comparison. We presume one of the most important concerns spontaneous vocalizations, where there is no obvious social or external environmental stimulus that appears to elicit the sounds, and especially where the vocalizer is physically alone. In addition to comparisons across species of volubility within various circumstances (perhaps social interactive negative, social interactive positive or neutral, and non-interactive), it will be of considerable interest to determine the *relative* frequency of occurrence of vocalizations *across* those circumstances since the non-interactive circumstance may be particularly important for vocal language and may show particularly high relative frequency in animals with a high capability for developing flexible vocal communication.

Functional Flexibility of Vocalization

A second infrastructural capability that develops in the human infant by 3–4 months has been termed functional flexibility (Griebel and Oller 2008). Three protophone types (squeals, vocants, growls) have been shown in longitudinal audio-video recordings to occur with massive functional flexibility, while cry and laugh show a strongly biased pattern of function, where the former is overwhelmingly judged as negative and the latter positive, with either auditory or visual presentation of the recorded stimuli (Oller et al. 2013). All three protophones on the other hand, show primarily neutral facial affect (judged by video alone), while at the same time showing significant numbers of occurrences where each one is judged positive on some occasions and negative on others. The whole range from positive to neutral to negative facial affect with a single protophone has been found to occur often within a single 20-minute period from individual infants.

Functional flexibility of the protophones was also shown with Illocutionary judgments of the infant vocal acts. Again all three protophones showed positively valenced Illocution (such as exultation or comfortable social displays in face-to-face interaction, typically accompanied by smiling) and negatively valenced Illocution (such as complaint or plea, typically accompanied by frowning), as well as many vocal acts that occurred in what appeared to be solitary vocal play (usually accompanied by neutral facial affect).

And finally the functional flexibility of the protophones was shown in judgments of their Perlocutionary Effects. All three protophones showed Perlocutionary outcomes where adults attempted to initiate or continue comfortable vocal interactions in response to infant positive Illocutions and facial affect, while adults attempted to change the situation for the infant, or talked about the possibility of changing the situation in response to infant negative Illocutions and facial affect, and finally, where adults tended to produce a mixture of

positive, neutral and ambiguous actions when infants produced Illocutions that were interpreted as socially neutral. This pattern of functional flexibility was very robust even by 3–4 months of age.

Such functional flexibility can be viewed as an additional infrastructural platform for vocal language (built upon the prior platform established by copious spontaneous vocalization, which we presume yields the protophone categories), and it represents a capability that has never been illustrated in natural communication of non-human primates or for that matter any other animal. We harken back to the argument presented above that Illocutionary flexibility is required for all words of any mature language. Exemplifying again, if one says “leopard” in English, one must have the freedom to do so as a pure act of naming, as a correction, as a warning, as a request, and so on. If a speaker could produce the sounds of the word, but could only do so with a single Illocutionary Function, then the sounds would not constitute a word at all for that speaker, but something far more primitive. Infants at 3–4 months of age show with their Illocutionary flexibility in protophones that one of the most fundamental capacities that will be required when they are ready to start learning words is already in place.

A direct quantitative comparison on functional flexibility would appear to be possible now for at least human infants and chimpanzees. This possibility is supported in particular by the recent development of a facial affect coding system for chimpanzee (Parr et al. 2008) modeled after the Ekman scheme for humans (Ekman and Friesen 1978). The demonstration of functional flexibility in human infants (Oller et al. 2013) involved collapsing the various facial configurations of Ekman to three categories of affect—positive, neutral, and negative. Then five vocal types (three protophones plus cry and laugh) were categorized for facial affect. An important question remains about whether it will be possible to categorize recorded samples of five or so vocal categories from chimpanzees or bonobos and judge them from video as positive, neutral, or negative in affective valence in a satisfactorily comparable way to the human infant data. We anticipate beginning such a study in the near future with human infants compared to chimpanzee, bonobo or both, with the intent of quantifying the extent to which Illocutionary flexibility can be demonstrated in terms of this three-fold emotional valence scheme.

Conclusion

Quantitative comparisons of vocal foundations for complex communication are now possible across humans and other related species, especially the great apes. But the optimal points of comparison are not, we think, as has been supposed, based on the extent to which the great apes or monkeys may share Meaning with humans. Meaning as defined here requires a wide variety of infrastructural capacities that have not been demonstrated in non-human primates, and data that have been interpreted as showing “functional referentiality” should be reconsidered as likely being the result of intelligent Perlocutionary reactions by primate receivers to Illocutionary acts of senders, rather than as transmission of Meaning encoded in primate signals.

The proper place to begin quantitative comparison with non-human primate natural vocal communication, we contend, is with infrastructural features of language that occur very early in the human infant (e.g., copious spontaneous vocalization, functional flexibility of vocalization), and for which a basis for apples to apples comparison is present. On the other hand, if we ask how much Meaning primate calls transmit, we seemed doomed to a comparison that makes no sense, or if we make the comparison faithfully, we seem doomed to come up with zero for the primate calls, to compare with some huge number for even a two-year-old human—and no one should be surprised by such a result if they accept as reasonable the definitions we have supplied above.

In contrast, if we ask how relatively voluble human infants and non-human primates are, or to what degree each shows functional flexibility in vocalization, we face real empirical questions. The answers will quantitatively specify differences between humans and non-humans in foundations for language, while at the same time providing an assessment of the extent to which human language may be founded in certain capabilities shared in other species of the primate lineage.

Acknowledgments

The research for this paper was funded by Grants R01 DC006099 and DC011027 from the National Institute on Deafness and Other Communication Disorders and by the Plough Foundation. Thanks go to Drew Rendall for helpful comments.

References

- Arnold K, Zuberbühler K. Semantic combinations in primate calls. *Nature*. 2006a; 441:303. [PubMed: 16710411]
- Arnold K, Zuberbühler K. The alarm-calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans martini*. *Anim Behav*. 2006b; 72:643–653.
- Austin, JL. How to do things with words. Oxford University Press; Oxford: 1962.
- Bakeman R, Brown JV. Behavior dialogues: an approach to the assessment of mother-infant interaction. *Child Dev*. 1977; 49:195–203.
- Bard KA, Vaclair J. The communicative context of object manipulation in ape and human adult-infant pairs. *J Hum Evol*. 1984; 13:181–190.
- Barr RG, Konner M, Bakeman R, et al. Crying in pKung San infants: a test of the cultural specificity hypothesis. *Dev Med Child Neurol*. 1991; 33:601–610. [PubMed: 1879624]
- Bates, E.; Benigni, L.; Bretherton, I., et al. The emergence of symbols: cognition and communication in infancy. Academic Press; New York: 1979.
- Bertossa RC. Theme issue 'Evolutionary developmental biology (evo-devo) and behaviour.'. *Phil Trans R Soc B*. 2011; 366:2056–2180. [PubMed: 21690124]
- Bloom, L. Language development. MIT Press; Cambridge, MA: 1970.
- Bretherton, I. How to do things with one word: the ontogenesis of intentional message making. In: Smith, MD.; Locke, JL., editors. *The emergent lexicon: the child's development of a linguistic vocabulary*. Academic Press; San Diego, CA: 1988. p. 225-260.
- Burkhardt, A. *Speech acts, meaning and intentions: critical approaches to the philosophy of John R. Searle*. de Gruyter; Berlin: 1990.
- Butterworth, G. Species typical aspects of manual pointing and the emergence of language in human infancy. Waseda University International Conference Center; Tokyo. 1996.
- Chen X, Green JA, Gustafson GE. Development of vocal protests from 3 to 18 months. *Infancy*. 2009; 14:44–59.

- Clay Z, Zuberbühler K. Food-associated calling sequences in bonobos. *Anim Behav.* 2009; 77:1387–1396.
- Clay Z, Zuberbühler K. Bonobos extract meaning from call sequences. *PLoS One.* 2011; 6(4):1–10.
- Condillac, EB. An essay on the origin of human knowledge; being a supplement to Mr. Locke's Essay on the human understanding. (Translation of *Essai sur l'origine des connaissances humaines.*). J. Nourse; London: 1756.
- Crockford C, Boesch C. Context-specific calls in wild chimpanzees, *Pan troglodytes verus*: analysis of barks. *Anim Behav.* 2003; 66:115–125.
- Darwin, C. The expression of emotions in man and animals. London: John Murray; 1872.
- Dore J. Holophrases, speech acts and language universals. *J Child Lang.* 1975; 2:21–40.
- Dörge, FC. PhD Thesis. Universität Tübingen; Tübingen: 2004. Illocutionary acts: Austin's account and what Searle made out of it.
- Ekmann, P.; Friesen, W. The facial action coding system. Consulting Psychologists Press; Palo Alto, CA: 1978.
- Feldman R, Greenbaum CW, Yirmiya N. Mother-infant synchrony as an antecedent of the emergence of self-control. *Dev Psychol.* 1999; 35:223–231. [PubMed: 9923477]
- Fichtel C, Kappeler PM. Anti-predator behavior of group-living Malagasy primates: mixed evidence for a referential alarm call system. *Behav Ecol Sociobiol.* 2002; 51:262–275.
- Franklin, BS.; Warlaumont, AS.; Oller, DK. Infant volubility by random sampling from all-day recordings. Paper presented at the American Speech-Language-Hearing Association; Chicago. 2013.
- Gardner, RA.; Gardner, BT.; Van Cantfort, TE., editors. Teaching sign language to chimpanzees. SUNY Press; Albany, NY: 1989.
- Gershkoff-Stowe L. Object naming, vocabulary growth, and the development of word retrieval abilities. *J Mem Lang.* 2002; 46:665–687.
- Gozoules, H.; Gozoules, S.; Ashley, J. Representational signaling in non-human primate vocal communication. In: Zimmerman, E.; Newman, JD.; Jürgens, U., editors. Current topics in primate vocal communication. Plenum Press; New York: 1995.
- Green JA, Gustafson GE, Irwin JR, et al. Infant crying: acoustics, perception and communication. *Early Dev Parenting.* 1995; 4:161–175.
- Green JA, Jones LE, Gustafson GE. Perception of cries by parents and nonparents: relation to cry acoustics. *Dev Psychol.* 1987; 23:370–382.
- Greenfield, PM.; Savage-Rumbaugh, ES. Grammatical combination in *Pan Paniscus*. In: Parker, ST.; Gibson, KR., editors. "Language" and intelligence in monkeys and apes: comparative developmental perspectives. Cambridge University Press; New York: 1990. p. 540-578.
- Griebel, U.; Oller, DK. Evolutionary forces favoring contextual flexibility. In: Oller, DK.; Griebel, U., editors. Evolution of communicative flexibility: complexity, creativity, and adaptability in human and animal communication. MIT Press; Cambridge, MA: 2008. p. 9-40.
- Gustafson GE, Green JA. Developmental coordination of cry sounds with visual regard and gestures. *Infant Behav Dev.* 1991; 14:51–57.
- Hockett, CF. Readings from Scientific American. Freeman; San Francisco: 1960. The origin of speech. Human communication: language and its psychobiological bases.
- Hockett, CF.; Altmann, SA. A note on design features. In: Sebeok, TA., editor. Animal communication: techniques of study and results of research. Indiana University Press; Bloomington: 1968.
- Lester, BM.; Boukydis, CFZ. No language but a cry. In: Papoušek, H.; Jürgens, U.; Papoušek, M., editors. Nonverbal vocal communication. Cambridge University Press; New York: 1992. p. 145-173.
- Lester BM, Boukydis CZ, Garcia-Coll CT, et al. Developmental outcome as a function of goodness of fit between the infant's cry characteristics and the mother's perception of her infant's cry. *Pediatrics.* 1994; 95:516–521. [PubMed: 7700751]

- Lester, BM.; Zeskind, PS. A biobehavioral perspective on crying in early infancy. In: Fitzgerald, HEBML.; Yogman, MW., editors. *Theory and research in behavioral pediatrics*. Vol. 1. Plenum; New York: 1982. p. 133-180.
- Lorenz K. Ausdrucksbewegungen höherer Tiere. *Naturwissenschaften*. 1951; 38:113–116.
- Macedonia JM, Evans CS. Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology*. 1993; 93:177–197.
- Maynard Smith, J.; Harper, D. *Animal signals*. Oxford University Press; Oxford: 2003.
- Meadows D, Elias G, Bain J. Mothers' ability to identify infants' communicative acts consistently. *J Child Lang*. 2000; 27:393–406. [PubMed: 10967893]
- Ninio, A.; Snow, C. Language acquisition through language use: the functional sources of children's early utterances. In: Levi, I.; Braine, M.; Schlesinger, IM., editors. *Categories and processes in language acquisition*. Erlbaum; Hillsdale, NJ: 1986.
- Oller, DK. The emergence of the sounds of speech in infancy. In: Yeni-Komshian, G.; Kavanagh, J.; Ferguson, C., editors. *Child phonology*. Vol 1: Production. Academic Press; New York: 1980. p. 93-112.
- Oller, DK. The emergence of the speech capacity. Erlbaum; Mahwah, NJ: 2000.
- Oller, DK. The natural logic of communicative possibilities: modularity and presupposition. In: Callebaut, W.; Rasskin-Gutman, D., editors. *Modularity: understanding the development and evolution of natural complex systems*. MIT Press; Cambridge, MA: 2005. p. 409-434.
- Oller DK, Buder EH, Ramsdell HL, et al. Functional flexibility of infant vocalization and the emergence of language. *Proc Natl Acad Sci USA*. 2013; 110:6318–6323. [PubMed: 23550164]
- Oller DK, Eilers RE. A natural logic of speech and speech-like acts with developmental implications. *First Language*. 1989; 9:225–243.
- Oller, DK.; Eilers, RE. Development of vocal signaling in human infants: toward a methodology for cross-species vocalization comparisons. In: Papoušek, H.; Jürgens, U.; Papoušek, M., editors. *Nonverbal vocal communication*. Cambridge University Press; New York: 1992. p. 174-191.
- Otte D. Effects and Functions in the Evolution of Signaling Systems. *Annual Review of Ecology and Systematics*. 1974; 5:385–417.
- Owings, DH.; Morton, ES. *Animal vocal communication*. Cambridge University Press; Cambridge: 1998.
- Owings, DH.; Zeifman, DM. Human infant crying as an animal communication system. In: Oller, DK.; Griebel, U., editors. *The evolution of communication systems*. MIT Press; Cambridge, MA: 2004. p. 151-170.
- Owren MJ, Rendall D. Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evol Anthr*. 2001; 10:58–71.
- Owren MJ, Rendall D, Ryan MJ. Redefining animal signaling: influence versus information in communication. *Biological Philosophy*. 2010; 25:755–780.
- Papoušek, H.; Papoušek, M. Qualitative transitions during the first trimester of human postpartum life. In: Prechtl, HFR., editor. *Continuity of neural functions from prenatal to postnatal life*. Spastics International Medical Publications; London: 1984. p. 220-244.
- Parr L, Waller BM, Heintz M. Facial expression categorization by chimpanzees using standardized stimuli. *Emotion*. 2008; 8:216–231. [PubMed: 18410196]
- Peirce, CS. Pragmatism in retrospect: a reformulation. In: Hartshorne, C.; Weiss, P., editors. *The collected papers of Charles Sanders Peirce*. Vol. 5. Harvard University Press; Cambridge, MA: 1934.
- Premack D. Language in chimpanzee? *Science*. 1971; 172:808–822. [PubMed: 5572906]
- Prescott R. Infant cry sound: developmental features. *J Acoustic Soc Am*. 1975; 57:1186–1191.
- Price, T. PhD Dissertation. Georg-August-Universität; Göttingen, Germany: 2013. *Vocal communication within the genus Chlorocebus: insights into mechanisms of call production and call perception*. retrieved May 23 2014 at <http://d-nb.info/1044769017/34>
- Quine, WVO. *Ontological relativity and other essays*. Columbia University Press; New York: 1969.
- Rendall D, Owren MJ, Ryan MJ. What do animal signals mean? *Anim Behav*. 2009; 78:233–240.

- Savage-Rumbaugh, ES.; MacDonald, K. Deception and social manipulation in symbol-using apes. In: Byrne, RW.; Whiten, A., editors. Machiavellian intelligence. Clarendon Press; Oxford: 1988. p. 224-237.
- Scott-Phillips TC. Animal communication: insights from linguistic pragmatics. *Animal Behaviour*. 2010; 79:e1–e4.
- Scott-Phillips TC, Gurney J, Ivens A, et al. Combinatorial communication in bacteria: implications for the origins of linguistic generativity. *PLoS One*. 2014; 9(4):e95929. [PubMed: 24759740]
- Searle, JR. *Speech acts: an essay in the philosophy of language*. Cambridge University Press; Cambridge: 1969.
- Searle, JR.; Vanderveken, D. *Foundations of illocutionary logic*. Cambridge University Press; Cambridge: 1985.
- Seibert J, Hogan A, Mundy P. Assessing interactional competencies: the early social-communication scales. *Inf Mental Hlth J*. 1982; 3:244–245.
- Seyfarth, RM.; Cheney, DL. Production, usage, and response in nonhuman primate vocal development. In: Hauser, MD.; Konishi, M., editors. *The design of animal communication*. MIT Press; Cambridge, MA: 1999. p. 391-417.
- Seyfarth RM, Cheney DL, Marler P. Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim Behav*. 1980; 28:1070–1094.
- Seyfarth RM, Cheney DL. Production, usage, and comprehension in animal vocalizations. *Brain Lang*. 2010; 115:92–100. [PubMed: 19944456]
- Searle, JR. *Speech acts: an essay in the philosophy of language*. Cambridge University Press; Cambridge: 1969.
- Searle, JR.; Vanderveken, D. *Foundations of illocutionary logic*. Cambridge University Press; Cambridge: 1985.
- Snowdon, CT. Linguistic and psycholinguistic approaches to primate communication. In: Snowdon, CT.; Brown, CH.; Petersen, MR., editors. *Primate communication*. Cambridge University Press; Cambridge: 1982. p. 212-238.
- Snowdon CT. Language capacities of nonhuman animals. *Yearb Phys Anthropol*. 1990; 33:215–243.
- Snyder L, Bates E, Bretherton I. Content and context in early lexical development. *J Child Lang*. 1981; 8:565–582. [PubMed: 7309817]
- Soltis J. The signal functions of early infant crying. *Behav Brain Sci*. 2004; 27:443–490. [PubMed: 15773426]
- Sparrow, SS.; Balla, DA.; Cicchetti, DV. *Vineland adaptive behavior scales*. American Guidance Service; Circle Pines, MN: 1984.
- Stegmann, UE. *Animal communication theory: information and influence*. Cambridge University Press; Cambridge: 2013.
- Stern, DN. Mother and infant at play: the dyadic interaction involving facial, vocal, and gaze behaviors. In: Lewis, M.; Rosenblum, LA., editors. *The effect of the infant on its caregiver*. Wiley; New York: 1974. p. 187-213.
- Struhsaker, TT. Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In: Altmann, SA., editor. *Social communication among primates*. Chicago University Press; Chicago, IL: 1967. p. 281-324.
- Suzuki TN. Referential mobbing calls elicit different predator-searching behaviours in Japanese great tits. *Anim Behav*. 2012; 84:53–57.
- Terrace, HS. *Nim: a chimpanzee who learned sign language*. Columbia University Press; New York: 1979.
- Tinbergen, N. *The study of instinct*. Oxford University Press; Oxford: 1951.
- Tomasello M, Brandt S. Flexibility in the semantics and syntax of children's early verb use: a commentary on Naigle, Hoff, and Vear (2009). *Monog Soc Res Child Dev*. 2009; 74:113–132.
- Tronick EZ, Als H, Brazelton TB. Mutuality in mother-infant interaction. *J Comm*. 1977; 27:74–79.
- Tronick EZ, Cohn JF. Infant-mother face-to-face interaction: age and gender differences in coordination and the occurrence of miscoordination. *Child Dev*. 1989; 60:85–92. [PubMed: 2702877]

- Wheeler BC, Fischer J. Functionally referential signals: a promising paradigm whose time has passed. *Evol Anthr.* 2012; 21:195–205.
- Zeifman DM. An ethological analysis of human infant crying: answering Tinbergen’s four questions. *Dev Psychobiol.* 2001; 39:265–285. [PubMed: 11745323]
- Zuberbühler K. Referential signaling in non-human primates: cognitive precursors and limitations for the evolution of language. *Adv Stud Behav.* 2003; 33:265–307.