"Y/Our" Vocal Sounds
Toward an Ethological Classification of Human Vocalizing Behavior and Sounds

Jay R. Feierman

Department of Psychiatry, School of Medicine, University of New Mexico, Albuquerque, NM, USA. (ret.)

jay.feierman84@gmail.com

> *There have been many authorities who have asserted that the basis of science lies in counting or measuring, i.e., in the use of mathematics. Neither counting nor measuring can however be the most fundamental process in our study of the material universe - before you can do either to any purpose you must first select what you propose to count or measure, which presupposes a classification.*
>
> R.A. Crowson (1970, p. 2)

ABSTRACT

Human ethology has been accused of being the behavioral science that treats humans as though they don't make vocal sounds. Although heritable, coordinated patterns of movement or fixed action patterns (Erbkoordinationen), which are the fundamental building blocks of classical ethology, can generate "our" vocal sounds (e.g., laughing, crying, and screaming), they cannot generate "your" vocal sounds (e.g., spoken words in English, German, French, Spanish, etc.). These latter "your" vocal sounds are made by single muscle movements, which when volitionally repeated, combined or coordinated, can generate culture-specific, behaviorally-generated vocal sound symbols, such as the general names in each culture for persons who give birth to babies (e.g., mother, Mutter, mère, madre). This article is a first attempt at an ethological classification of human vocalizing behaviors and the sounds they make. Lastly, how human ethologists could study these two different types of human vocalizing behaviors, using the sounds they make as proxies, will be discussed.

Keywords: *Behavior, biology, classification, ethology, sound, speech, Type I behavior, Type II behavior, vocal.*

I. INTRODUCTION

Background

Ethology is the biology of behavior (Eibl-Eibesfeldt, 1975). One of the founders of classical ethology defined behavior as "the total of movements made by the intact animal" (Tinbergen, 1951, p.2). One function of some behaviors is to generate vocal sounds. Fitch (2000) makes a compelling argument that the evolution of behaviorally-generated human speech can be studied independently of the evolution of human language. This article goes one step farther. It proposes that human vocalizing behaviors and the sounds these behaviors make can be studied ethologically, somewhat independently of human speech (Fitch, 2000), language (Lieberman, 2006), and communication (Hauser, 1997). Because ethology is the biology of *behavior*, the emphasis is going to be on the vocalizing behaviors, even though vocalizing behaviors are usually recognized by the types of sound they make. In addition to the "physical" vocalizing behaviors, appreciate that vocal sounds are also "physical." When between mouths and ears the vocal sounds humans make are information-encoded, structured air waves; and as such, they have some ontological as well as evolutionary similarity to all other physical entities of biological interest.

Conceptual framework

This article is neither a research report nor a literature review. It is a conceptual framework by which human ethologists can begin to think about the differences between the vocal sounds made by all humans versus those particular to specific human cultures. Clearly, all human vocalizing behaviors and the vocal sounds they make do not fall within a homogeneous single category. As obvious evidence for this claim, whereas some human vocalizing sounds are understood by all humans, such as laughing and crying, we cannot understand the meaning of the most common vocal sounds made by all other humans. We are unique in this respect as a species. Also, another reason why our vocalizing behaviors and the sounds they make are not a homogeneous category is that only certain human vocalizing behaviors and sounds have potential phylogenies (evolutionary histories) and are amenable to all of Tinbergen's four questions by which one should attempt to fully understand any behavior (Tinbergen, 1963).

Although there is a phylogeny to the human vocalizing behaviors that produce the sounds of human crying (Zeifman, 2014), there is no phylogeny, in the traditional meaning of the term (i.e., that excludes "phylogenetic linguistics" within the human species, such as Longobardi et al., 2016), to the human vocalizing behaviors that produce the sounds "Oh, my God!" in the English language that often are uttered when something terrible is about to happen or just happened.

Therefore, for a human ethologist to investigate these two types of human vocal sound-generating behaviors through their sound proxies (i.e., human universal crying compared to the culture-specific, "Oh, my God!") as though they were one type of behavior would not generate as much useful information as if the two types of behavior were studied separately. Classification, which creates more homogeneous categories, is the first step in the pursuit of biological knowledge (Crowson, 1979; Panchen, 1992). To classify the sound-generating vocal behaviors that humans make, human ethologists also

must move beyond the tools that were developed initially by ethology's founders, who primarily studied non-vocalizing behaviors in fish and birds.

Perspectives and paradigms

Because this is a first attempt toward an ethology of human vocalizing behaviors and sounds, the article builds upon the methods and terminology of Nobel Laureate ethologists Konrad Lorenz and Niko Tinbergen as a foundation. That means other equally valid terminologies - e.g., "environmentally stable behavior" versus "environmentally labile behavior" - of other well respected authorities, such as Robert A. Hinde (1970), who takes more of a developmental approach, are not used, even though different terms from the two different schools (e.g., "environmentally stable" and "instinctual") as qualifiers of behavior mean almost the same thing.

In Part II of this article, "behavior" is going to be defined on the basis of what it is (i.e., movement), rather than on the basis of what it does (i.e., its function). That alone, i.e., defining "behavior" by what it is rather than what it does, is going to create cognitive dissonance in some, but not all, persons who read this article. In biology, form and function are equally important. It is just that in biology one almost always starts by figuring out what something is before one tries to figure out what it does. Most people trained in psychology are used to thinking about behavior primarily in terms of what the behavior does (i.e., its function), rather than first thinking about what the behavior is. That being said, this article is not a new paradigm. Rather, it just approaches human vocalizing behaviors and the sounds they make by building upon an old and somewhat neglected paradigm, that of classical ethology.

This article uses some of the long forgotten, systematic biological principles of classical ethology, which had its foundation in zoology (Tinbergen, 1963) and comparative anatomy (Lorenz, 1981). For many persons, who consider themselves human ethologists today, even though their object of study might be behavior (movement), their interpretations of what they are observing are often influenced to a large extent by newer, functionally-based psychological paradigms. It is worth quoting Lorenz's (1981, not numbered Preface) comments on this topic.

> In some respects the development of a science resembles that of a coral colony. The more it thrives and the faster it grows, the quicker its first beginnings - the vestiges of the founders and the contributions of the early discoverers - become overgrown and obscured by their own progeny. There is one drawback to the strategy of growth pursued by the coral tree. The polyps at the end of its branches have a much better chance of further development than those situated near the foundation. The ends go on growing faster and faster without considering the necessity for strengthening, in proportion, the base that must carry the weight of the whole structure. Unlike an oak tree, the coral colony does nothing to solidify its support. Consequently, there is a lot of coral rubble detached from points of departure, and this is either dead or, if still partly alive, growing in a indeterminate directions and getting nowhere.

This article will therefore be anchored in the systematic methods of zoology and comparative anatomy, which first uses classification to make one's object of study as homogeneous as possible. In doing so, the article systematically addresses what one's object of study is prior to trying to understand what it does. The object of study in this article is the behavior humans use to make vocal sounds, and then the sounds themselves, not so much for their semantic meaning but more so as proxies for the types of behavior that generate them. Vocal sounds can be thought of as functionally analogous to pheromones in that both are communicative media. Just like one can study the physiology of how different pheromones are produced, one can study the physiology of how different vocal sounds are produced.

The argument will be developed, hopefully persuasively, that the behaviors humans use to make vocal sounds are not a homogeneous "object of study." That is not so different as saying that white blood cells are not a homogeneous object of study, which was realized in the 19th century. In terms of vocalizing behaviors, what will result will be two mutually-exclusive types of vocalizing behavior, called Type I and II, that make two mutually-exclusive types of human vocal sounds, Type I and II.

To develop the argument that there are two, mutually exclusive types of human vocalizing behaviors and sounds requires reviewing some the terms, features, and principles that will be used.

Terminology

Volitional and spontaneous behavior and behavior-generated vocal sound

These terms will be used in the article as non-mutually-exclusive, dimensional features of behavior and sound. Volitional will mean self-initiated, somewhat unpredictable behavior or sound that always involves prefrontal and frontal, neocortical pyramidal neurons, including those in the frontal motor strip that create the monosynaptic, pyramidal tract pathway (Rubin and Safdieh, 2016). Spontaneous behavior and behavior-generated vocal sound is at the other end of the continuum, which for the purpose of this article can be thought of as behavior or sound that is simply less volitional and does not require (but also does not rule out) involvement of the neocortical prefrontal and frontal pyramidal neurons. These two features - volitional and spontaneous - are useful in characterizing vocalizing behaviors and sounds, some of which are more volitional and some of which are more spontaneous.

Intentional and expressive behavior and behavior-generated vocal sound

These terms will be used in the article also as non-mutually exclusive, dimensional features of both behavior and behavior-generated vocal sound. Intentional behaviors and vocal sounds are movements and sounds that primarily carry out utilitarian tasks, like walking, asking, and naming. As used by ethologists, "intentional" does not mean the same thing as "intentional" when used by philosophers of mind, such as Dennett's (1998) intentional stance. Intentional also does not mean the same thing as "volitional," as some intentional behaviors and sounds get executed more spontaneously. Expressive behaviors and sounds signal an individual's mood (specific readiness to act), often interpreted as a need, to a conspecific (member of the same species). They are

communicative signaling devices using many muscles of the body with a number of synonyms. In general they are called displays, but more specifically called affects (when on the face), prosodies (when modifying speech), and what is generally called "body language." Appreciate that a crying baby's expressive behavior can also carry out an intentional task, like getting milk, by using the parent's muscles as an extended phenotype.

Heritable

This term will mean that the construction information for a "structural design feature" is passed across generations through DNA, or sometimes by acquired, epigenetic modifications of DNA, rather than through cultural (social) learning mechanisms. A structural design feature is a very general term for any physical entity that has static or moving architectural mass by which it can be defined (Feierman, 2009a). The value in using the term structural design feature is that it precludes entities that can only be defined by their function and not by their form being heritable candidates. DNA only codes for form.

A structural design feature can be an inanimate material object like a flaked projectile point, the construction information for which is passed across generations by cultural social learning. Or, a structural design feature can be a heritable, animate object like a base pattern in DNA, a protein molecule, a neural network, a bone, or certain types of behavior and vocal sound.

For a behavior to be heritable it has to be able to be defined by its form (architectural pattern of movement in 3-D space) for a species-universal (and expected) function, which by definition qualifies the behavior as a structural design feature. Some vocalized sounds can also be heritable if they can be defined by their form for a species-universal function (which by definition also qualifies them as structural design features). This only occurs if the vocal sounds are produced by behavior that can be defined by its form for a species-universal function. Heritable does not mean "innate," a term that should be avoided (Lehrman, 1953).

Heritable, structural design features are phenotypes whose expressions almost always require interactions of the genotypes with other non-DNA environmental elements as proximate contributing causes to develop. Heritable phenotypes can be anatomical or some behaviors and behavior-produced vocalizations. Other, non-DNA elements include environmental Releasing Stimuli (including Social Releasers emanating from other human beings), and various stressful life events and environmental chemical exposures, both of which can influence the expression of DNA by epigenetic mechanisms (Allis, Caparros, Jenuwein, & Reinberg, 2015).

Ontogenic learning

This term is used in the article to distinguish learning during the lifetime of an individual from "phylogenetic learning" - i.e., by the so-called "innate schoolmarm" (Lorenz, 1981) - that occurs over long periods of evolutionary time at the level of a species and which results in the execution of species-typical instinctual motor behaviors upon encountering specific Releasing Stimuli.

Central pattern generator (CPG)

This term, which hereafter will be referred to by the acronym CPG, will mean "any network within the CNS [central nervous system] that coordinates a motor behavior or a part thereof" (Grillner, 2006). This includes the heritable motor pattern behaviors that produce certain types of vocal sound. As will be seen, some CPGs generate a sound-producing coordinated motor pattern once, like producing a single scream sound; other CPGs generate a rhythmic (repeated) coordinated motor pattern, like that which produces laughing sounds. Crying sounds have a frequently interrupted and more variable but still recognizable rhythm, which is not as regular in time period as that of laughing.

CPGs are species-typical, hardwired, heritable neural networks in the CNS, which operate through polysynaptic extrapyramidal circuits to influence certain cranial nerve and spinal cord lower motor neurons. They can be called into action for motor tasks when driven by other neurons with command functions (Grillner, 2006). By contrast, motor behaviors that are coordinated via ontogenic learning and practice, such as the breathing and articulation muscle movements used to say, "My name is Joe", or to play a song on the piano, are not coordinated by heritable CPGs. They have very different innervations and processes for being acquired and executed.

Value of the ethological perspective in studying vocalizing behaviors

Vocal sounds are produced by behavior. Ethology is the biology of behavior, which means human ethologists should be able to bring a unique and potentially useful perspective to this subject. Using the methods of ethology, two categorical classes or types of human vocalizing behavior are going to be developed. As will be shown, they have different biological features and are innervated, acquired, and executed by different biological mechanisms. Ethology is in a unique position to understand some of these biological mechanisms. As an example, if one takes the just mentioned English words, "Oh, my God!", there appears to be similarity by analogy (in biological function, which is more than semantic meaning) among different human cultural groups as to when different words meaning the same thing are uttered. "Oh, my God!" and "mon Dieu!" and "mein Gott!", when uttered by specific individuals, are instances of vocal sounds produced by vocalizing behaviors that for the entire human species are so variable in form at the species-wide class level for a particular function, that they can only be defined by their particular function. The instances are similar by analogy, a concept and way of thinking that can provide a source of knowledge (Lorenz, 1974). Similarity by analogy, which is a biological concept, is not how most linguists think. Many of the tools of linguists are not as well suited to understanding some of the analogous behaviors and sounds that humans make as are the tools of human ethologists.

Neuroethology

When aspects of the classification developed in this article can be related to Tinbergen's proximate cause question, which in some instances can now be matched with discoveries in neuroscience, data and references will be cited. Classes or types of behavior and behavior-generated sound can also be differentiated by different kinds of innervating neural pathways that generate them.

Quick preview

Prior to categorizing (i.e., creating homogeneous classes of) human vocalizing behaviors, let us take a moment to think about human vocal sounds themselves and not the behaviors (muscular movements) that create them. There are two obvious classes of human vocal sounds: (1) "our" vocal sounds that are species-universal in form and meaning across culture groups (e.g., the sounds of laughs, cries, screams, sighs, groans, moans, etc); and (2) "your" vocal sounds that are not species universal in form across cultural groups for the same meaning (e.g., the learned sounds of "I want a piece of cake" and "Du bist mein Mutter").

But, across at least two different cultural groups, "You are my mother" and "Du bist mein Mutter" are two different culturally acquired (learned) vocal sounds that have different forms for the same meaning. There is only one meaning to both of these sentences (in English and German) about the woman who gave birth to you, but when uttered by a child they could be uttered for many different functions (reasons), depending on the effect the child believes the vocal sounds will have on the woman who gave birth to her. Both classes of vocal sounds - the culture-universal in form and meaning "our" vocal sounds (which will be called Type I), and the culture-specific in form and meaning "your" vocal sounds (which will be called Type II) - will be discussed in reference to the two different categorical classes of behavior, Type I and II, that create them. The argument will be made that these two classes of behavior and sound are mutually exclusive biological categories (not just dimensions) that need to be separate objects of study when learning more about them.

Organization of the article

With that as background, let us now see in Part II what is meant by "behavior" and how different behaviors can be characterized. These behavioral characterizations in Part II of the article will then be used to create the Type I and II classes of behavior in general in Part V of the article. Then the Type I class of vocalizing behavior and sound will be explored in Part VI; and the Type II class of vocalizing behavior and sound will be explored in part VII. Finally, in Part VIII it will be shown how Type II intentional behaviors and sounds can acquire expression.

II. CHARACTERIZING HUMAN BEHAVIOR

What is meant by "behavior"

Although Tinbergen (1951, p. 2) defined behavior as "the total movements made by the intact animal", that definition can also apply to invertebrates, who often operate by different rules. By the use of the term "self-generated" as a qualifier of behavior, some movements in vertebrates will be excluded: passive movements (i.e., being pushed or pulled by another individual), and falling towards earth because of gravity. Using the term "skeletal/striated" as a qualifier of muscle excludes contractions of non-striated smooth and cardiac muscles, both of which are only innervated by the autonomic nervous system and which operate by different rules.

"Behavior" will also not mean the same thing in the article as an individual's "response," as a response can also include such things as autonomic nervous system mediated changes in heart rate, salivation, etc., which don't follow the same rules of engagement as skeletal muscle contractions. All behaviors are responses but not all responses are behaviors.

So, the definition of vertebrate behavior more comprehensive in the context of this article that can be applied to humans will be "the self-generated movement of an individual or a part of an individual as the result of the contraction of skeletal/striated muscles." Note that "behavior" is defined by what it is (movement) rather than what the movement, which would then be overlooked, does (i.e., its function).

Features, instances, classes, dimensions and categories of behavior

Ethology is an inductive behavioral science that ideally starts with observation, inference, and classification to create "objects of study" that are as homogeneous as possible. In this article, a feature is a component or aspect of behavior. It is something about a behavior, such as its rate or pattern. An instance is an example of one individual executing a behavior, where the behavior that is executed is a member of a more general class of behaviors. So John's walking is an instance of the class of behavior called "walking".

One should be able to assign class membership from a single instance but often not by knowledge of a single feature. As will be seen, features that are associated with a class can be dimensional or categorical. As a result, classes can be conceptualized categorically (e.g., mammal or bird) or dimensionally (spontaneous or volitional behavior).

Both dimensions and categories will be used in this article as components with which to build Type I and II categories of behavior. Because dimensions are features of things, they have to have at least some "thing" in common. Behaviors have movement in common. So one could put features of behaviors into dimensional classes of movement. Some dimensions of behavior can occur simultaneously, like moving up, sideways, slowly, and rhythmically all at the same time. Other dimensions of behavior occur along a continuum, such as from spontaneous to volitional. By contrast, categories, when formally defined, can create mutual exclusivity, which is why categories rather than dimensions are used in classification in biology. For example, lymphocytes and neutrophils are different categories of white blood cells, rather than different dimensions of the white blood cell. Dimensions have more use in the study of behavioral development or in characterizing certain features of behavior (e.g., spontaneous or volitional) or when classical ethology is synthesized with comparative psychology (Hinde, 1970).

If the categorical approach to grouping behaviors and behavior-generated vocal sounds in this article is challenged, which it should be, the two resultant, mutually-exclusive categories, which are going to be called Type I and II, can be refuted empirically by finding an exceptional human behavior or vocal sound not mentioned in the article that meets formal criteria for both Type I and II categories.

Ways of conceptualizing human behavior's form and function

Single versus coordinated pattern of human behavior.

Single human behaviors. One or more named muscles (or muscle groups) moving a part of the body once and only in one 3-D vector direction (e.g., up, down, in, out, away, towards, flexion, extension, pronation, supination, etc.) in reference to a fixed point either on the individual or in the environment. Volitional single behaviors can be repeated, combined, or coordinated, but these possibilities are facultative and not obligatory. Here are three examples of a single behavior.

i. *spontaneous orientation (i.e., taxis) reflex.* (e.g., automatically turning the trunk, head, eyes, or limbs towards certain objects or stimuli [often novel] or forces in the environment)

ii. *spontaneous, reflexive avoidance movement from a noxious stimulus.* (e.g., corneal blink reflex)

iii. *spontaneous, instinctual facial greeting display.* (e.g., human eyebrow flash)

Coordinated patterns of human behavior. One or more named muscles (or muscle groups) moving a part or parts of the body spontaneously or volitionally in more than one direction at the same time or sequentially, either once (i.e., non-rhythmically) or repeated (i.e., rhythmically). Appreciate that even spontaneously generated coordinated patterns of behavior are subject to volitional initiation, inhibition, and termination. Here are three examples.

i. *CPG-coordinated, spontaneous non-rhythmic pattern of behavior.* (e.g., non-rhythmic, spontaneous Duchene smiling)

ii. *CPG-coordinated, spontaneous rhythmic pattern of behavior.* (e.g., rhythmically walking, breathing, laughing)

iii. *non-CPG-coordinated, non-spontaneous (i.e., volitional) pattern of behavior acquired by ontogenic learning and practice.* (e.g., non-rhythmically taking a ball point pen apart or rhythmically playing a song on the piano)

Form versus function of human behavior.

"Form" (also called "structure") is what behavior is. It is a description or a definition of the pattern of movement of the whole body or parts of the body in 3-D space in time - i.e., "patterns in time" (Eibl-Eibesfeldt, 1975, p. 1) - either in reference to a point on the behaving individual or a point in the environment.

"Function" is what the behavior does, conceptually. It is a description or definition of the behavior's result or outcome (i.e., its consequence that is derived from

observation and inference). It is usually expressed by a present participle verb that ends in "ing", as in "calling" or "speaking". The conceptual function of what a behavior does has to be differentiated from the "physical" process of doing it, which is the answer to the proximate cause "How?" question (Tinbergen, 1963). There are three kinds of biological function, which will be explained. All of them require observation and inference to assign an instance of behavior to a class. Only the first one is commonly used in human ethology:

    i. proximate use function. This is what a behavior is doing (i.e., it's consequence) in reference to the next hierarchical functional level in which the index behavior is embedded (Tinbergen, 1951). For example, one proximate use function of "moving" the tongue could be "speaking". And one proximate use function of "speaking" could be "lecturing". Proximate use function will be the only meaning of "function" when the term is used in this article. Commonly used in human ethology.

    ii. selected effect function. This is what a behavior was doing (i.e., its functional outcome or consequence) historically when the behavior or the anatomical structures that generated the behavior were under positive selection pressure and increased in prevalence in the population. Less frequently used in human ethology.

    iii. activity function. This is what a behavior is doing (i.e., its functional outcome or consequence) in reference to a point on the behaving individual or a point in the environment but not in reference to the next hierarchical functional level in which the index behavior is embedded. For example, "flexion" of a joint is an activity function. Activity functions are used in orthopedic surgery, rehabilitation medicine, and physical therapy. Not used in human ethology.

Describing versus defining human behavior

Describing behavior.

When one describes behavior (by its form or function) one memorializes certain, not necessarily all, observable features of the movement into vocal speech or a written format, which also requires inference if the description is of the behavior's function. A description tells you how a behavior's form or function appears or what a behavior's form or function is like.

    The more extensive a description of a class of behavior, the more the listener or reader of the description will be able to assign a future instance of behavior to the previously described class. The communicative effectiveness of a description often depends on its length.

    Examples will be given of describing behavior by form and function using an extrapyramidally-mediated, heritable, CPG-generated coordinated pattern of

spontaneous behavior, the Duchenne human smile (Frank & Ekman, 1996). Single behaviors follow the same principles.

> *Describing behavior by form.* Her mouth opened slightly for a few seconds and exposed her teeth.
>
> *Describing behavior by function.* Over the period of an hour, she "countenanced" a number of men in a such a way that several of them walked over to her.

Defining behavior.

When one defines behavior by form or function one first puts the behavior into a familiar general category (based on its form or function) and then one indicates how the index behavior now in the familiar general category is different from other behaviors in the same familiar general category. This is called an intensional (or connotative) type of definition, which is the type used in all biological classification in which the familiar general category is the genus and the smaller category is the *differentia*, the part of the definition not provided by the genus. Differentia can also have major and minor components. If the definition is of the behavior's function, inference is required in addition to observation.

The effectiveness of a definition in assigning an instance of new behavior to a previously defined class need not depend on the length of the definition. Definitions create better categorical boundaries (borders) than descriptions, although both are methods for characterizing behavior. In classification from species to cell types, definitions trump (i.e., rank above) descriptions in assigning instances to classes.

Examples will be given of defining behavior by form and function using the same extrapyramidally-mediated, heritable, CPG-generated coordinated pattern of spontaneous behavior, the Duchenne human smile. Single behaviors follow the same principles.

> *Defining behavior by form.* She exhibited a facial coordinated motor pattern (the genus) lasting a few seconds and produced by the combined contraction of the zygomaticus major and the orbicularis oculi muscles (the differentia).
>
> *Defining behavior by function.* Over a period of an hour, she exhibited a communicative facial expressive behavior (the genus) that was followed by a series of different men casually walking over to her while exhibiting the same facial expressive behavior (the differentia).

Two ways of categorizing coordinated patterns of human behavior

Heritable and species-universal in form for a species-universal function.

There are some coordinated patterns of human behavior - e.g., facial expressive behaviors (such as those that produce spontaneous Duchenne smiling), the behaviors of walking, the behaviors that produce the sounds of laughing and crying, the behaviors that make-oneself-lower-or-smaller-or-more-vulnerable (LSV) that are used in submission, the behaviors of 3-5/second rhythmic pelvic-humping that are used in copulating - that are both heritable as well as species-universal in form for a species-universal function. The patterns of movement (non-rhythmic or rhythmic), which are spontaneously executed,

are coordinated sub-cortically or in the spinal cord by CPGs. They are initiated by command neurons (Grillner, 2006) and are not primarily caused by cranial nerve or spinal lower motor neurons being stimulated directly through the mono-synaptic, volitional pyramidal tract system (Arshavsky et al., 1997; Bucher et al., 2015; Dimitrijevic et al., 1998).

When initiated experimentally by either a very brief, sub-cortical electrical pulse in the brain or naturally by an environmental Releasing Stimulus, and unless volitionally inhibited or terminated, such CPG-generated coordinated patterns of movement almost always run to completion. This was first demonstrated many years ago by Lorenz and Tinbergen in the egg retrieval behavior in a graylag goose that was memorialized on a silent, grainy, 14 second, black and white film. As an example, when the central nucleus of the amygdale is stimulated experimentally by a very brief electrical pulse, it produces the more prolonged, species-universal-in-form, CPG-generated, ordinarily spontaneous, facial coordinated motor pattern of fear (LeDoux, 1996).

Also, direct sensory input is not needed for CPG patterns or rhythms to be continued. Sensory afferents to muscles can be cut and the rhythmic pattern continues, as first demonstrated in the classic experiments of von Holst (1973). However, sometimes indirect sensory input refines the pattern in response to external events, such as modifying one's walking on uneven terrain (Grillner, 1985). Because CPG-generated behaviors, once initiated, can occur without volitional input for their maintenance, a human person's attention can be directed elsewhere, thereby allowing us to think about other things or even have a complex conversation with someone while walking (Llinas, 2001).

Heritable and species-universal-in-form behaviors for species-universal functions are familiar to all vertebrate ethologists as fixed action patterns or coordinated motor patterns (Erbkoordinationen). They are used in both the appetitive (searching or vocal calling) and in the consummatory end act aspects of instincts. When used as consummatory end acts, they are disinhibited by environmental Releasing Stimuli acting on Innate Releasing Mechanisms in the central nervous system (Ewert, 1980; Immelmann & Beer, 1989; Lorenz, 1981).

In Part V of this article, these heritable, species-universal-in-form for species-universal function, non-pyramidal-tract-mediated, CPG-generated spontaneous behaviors will be called Type I behaviors. Natural selection treats them no differently from anatomical structures. They are as unique in form as organs and can potentially have a phylogeny (evolutionary history). Historically they were even used in taxonomy and phylogeny, which according to Konrad Lorenz (1981), was one of the foundations of classical ethology.

Non-heritable and non-species-universal in form for a particular function.

Not all coordinated patterns of human behavior are heritable and are species-universal-in-form for a species-universal function. In Part V of this article it will also be shown how non-heritable, volitional single behaviors can be repeated, combined, and coordinated through ontogenic learning and practice into non-species-universal-in-form behaviors for particular functions. Note that a "particular function" is different from a "species-universal function." Making a grocery list in longhand in cursive writing is a particular

function. These behaviors will be called Type II behaviors. They are mediated via the monosynaptic pyramidal tracts. Some are rhythmic and some are not.

As a non-rhythmic example, consider the volitional single, hand and finger movements that are repeated, combined, and coordinated in manipulating almost any small object in the hand, like the fine movements needed in taking a ball point pen apart. One sees almost identical repeated, combined, and coordinated volitional patterns of single hand and finger movements in the great apes when they manipulate small objects in their hands.

As a rhythmic example, consider the volitional, single individual finger movements that are repeated, combined, and coordinated to play a particular song on the piano. Even if all people who play the particular song on the piano do so with volitional finger behaviors with the same form, only certain members of the human species can play the piano, which is why the movements are not species-universal in form for a particular function. In other words, they are ontogenically-learned, coordinated patterns of behavior.

III. ONTOLOGICAL CATEGORIES

Form and function, where function is just the conceptual answer to the "What does the behavior do?" question, are in two different ontological realms within ontological pluralism (Turner, 2012). Forms are in the ontological realm of mass, energy, force, space, time and information. Functions are in the non "physical" conceptual ontological realm, where "physical" (in scare quotes) stands for mass, energy, force, space, time and information. In all of the life sciences forms do not change into functions like matter can change into energy in physics. Neither is function a feature (i.e., an inherent part) of form. Behavioral forms - patterns of movement in time - can and often do change proximate use function through ontogeny (development over the lifespan of an individual) and phylogeny (evolutionary history).

IV. CAUSAL CHAIN CATEGORIES

In the life sciences proximate, contributing, behavioral causation chains do not cross between the two different ontological realms of form and function. Only combinations of mass, energy, force, space, time and information, inside or outside of the body, can cause "physical" movement, such as movement in the inter-costal muscles, diaphragm, laryngeal muscles, lips and tongue that produce behavior-generated vocal sounds. Inferred "mental" brain functions in humans (e.g., thinking, remembering, believing, etc.), which are not the same as the "physical" processes (i.e., the answer to the proximate cause "How?" questions) of doing these things, can only cause functionally characterized concepts like calling or speaking. A functional, proximate, contributing causation chain in a professor would be something like thinking > speaking > lecturing > teaching, etc.

V. TYPE I AND II HUMAN BEHAVIORS

Classification issue

What has been covered in Part II in "Characterizing Human Behavior" will now be used in Part V as the components with which to make the bipartite, Type I and II, general categorization of human behavior. The terms "Type I" and "Type II" as categories of human behavior are consistent, with a slight modification in their application to making a classification of vocal-producing behaviors, with the terms originally developed by Feierman (2009a, 2009b) to understand the evolutionary history of the non-vocal aspect of human religious behavior.

As an aid in remembering what the terms Type I and II mean, "Type I" is the 1st or older type of behavior found in all vertebrates, similar to what classical ethologists called fixed action patterns. Type II behaviors are evolutionarily newer, have different neuroanatomical innervations and different methods of developmental acquisition and expression; and among other things (such as the hand and finger movements in humans and great apes), they are used to produce human vocalized symbolic languages.

In biology, classifications of entities from molecules (hemoglobin versus albumen) to cells (neurons versus astrocytes) to tissues (mucous membranes versus epidermis) to organs (heart versus liver), are always based on definitions, not simple descriptions, of what entities are, i.e. their form or structure, rather than on definitions of what the forms or structures do, i.e. their proximate use function. The familiar ethogram in ethology, which is arranging names of different fixed action patterns into functional groups, is not a formal biological classification. It is just a useful way ethologists group together behaviors with different forms that happen to have similar functions.

In biology - and ethology is the biology of behavior - we first need to define what an entity is before we try to understand what the entity does. Therefore, the definitions of Type I and II behaviors that are going to be used in this article have to be slightly modified for defining categories of vocalizing behaviors from what was first proposed by Feierman (2009a, 2009b) for understanding the evolutionary history of non-vocal-producing religious behaviors. Feierman's two categories of behavior, Type I and II, were not part of a formal, biological classification because one of the categories, Type II behavior, could only be defined by its function, not by its form, in differentiating it from Type I behavior.

After defining "behavior" simply as "movement," Feierman (2009b, p. 72) defined Type I behavior as behavior that is "*definable* by form and function in a natural environment and species-universal in form." He defined Type II behavior as behavior that is "*describable* by form and *definable* by function in a natural environment and not species-universal in form". [Italics in the original.] In this article, because vertebrate behavior is defined more elaborately and because Type I's behavior is slightly modified, both Type I and Type II behaviors will now both be able to be defined by their form, making Type I and II a formal, biological classification of human vocalizing behaviors and the sounds they make.

In Feierman (2009a, 2009b), Type I behavior did not have the qualifier "coordinated pattern of movement," which is a general form or structure of the behavior. As such, single behaviors that are traditionally considered "simple reflexes" and which are heritable and definable by form for a species-universal function (e.g., orienting reflex, the

corneal [as distinct from the visual optical] blink reflex, the eyebrow flash, and some others like the patella stretch reflex) all qualified as Type I behaviors for Feierman. However, in this article these single behavior "simple reflexes" no longer qualify as Type I behaviors because they don't involve CPG-generated coordinated patterns of movements, which has now been added as a Type I criterion.

Type I and Type II behavior as classes are now both in the same behavioral genus, "movement." The main differentia for the class Type I behavior is "coordinated pattern". The main differentia for the class Type II behavior is "single". Beyond the common genus and main differentia of the two classes are minor differentia for each behavioral class. To assign instances of behavior or sound to the appropriate class takes further observation, including further observation of other individuals in the same species in other cultural groups. Investigation may also be required to know if the instance of behavior that is observed is heritable or non-heritable, or if it is species-universal in form versus non-species-universal in form for a species-universal or a particular function. And, inference is always required in addition to observation when attributing function to an instance of behavior.

For the time being the single behavior "simple reflexes", which have now been excluded from Type I Behavior, are being considered members of an atypical category called Type Ia behavior. In many ways Type Ia single behaviors have more categorical similarity (i.e., spontaneous, heritable, species-universal in form for a species-universal function, and not pyramidal tract innervated), to the coordinated patterns of Type I behaviors than they do to the single, volitional, pyramidal-tract-innervated Type II behaviors.

All heritable, spontaneous, species-universal in form for a species-universal function, Type I human vocalizing behaviors are pre-packaged as coordinated movements of breathing muscles, laryngeal muscles, lips and tongue in which the single movements cannot be separated out and executed individually, like they can be in Type II behaviors. Also, the old idea that the coordinated pattern instinctual behaviors, another synonym for some uses of Type I Behaviors, are just strung together simple reflexes has long been discredited (Lorenz, 1981; von Holst 1973). In addition, Type Ia behavior reflexes are never used for vocal communication.

Type I human behavior

A class of spontaneous behaviors characterized by "heritable, coordinated patterns of movement, which in a natural or semi-natural environment, can be defined by their form for species-universal functions." Being definable by form for species-universal functions makes Type I behaviors species-universal in form for species-universal functions.

Type I behavior is just a more precise formal definition for classes of behaviors called fixed/coordinated action patterns (Erbkoordinationen) of classical ethology that were previously discussed under heritable, coordinated patterns of behavior that are species universal in form for a species-universal function. A natural or semi-natural environment is one that is at least similar to the one in which the behavior is ordinarily executed by members of the species. In non-human animals, some instances of Type I behaviors might not express in rather sterile, non-natural type environments, such as a Skinner

Box. Many instances of Type I human behaviors would also not express if a human were, for example, alone in a prison cell.

Type II human behavior

A class of volitional behaviors characterized by "non-heritable single movements that move a part of the body only in one direction and that for particular functions can only be described by their non-species-universal forms." Type II behaviors can be volitionally (i.e., facultatively and non-obligatorily) repeated, combined, or coordinated through ontogenic learning and practice

Excluding the heritable, Type Ia single-behavior simple reflexes just discussed and the orienting (taxis) and corneal blink reflexes, and the eyebrow flash, virtually all other single human behaviors are Type II behaviors. These behaviors are restricted to only a few vertebrate taxa other than human beings. The combined movements of Type II behaviors can be tactically modified in form through ontogenic learning. They can only be executed when there is a direct (i.e., monosynaptic) neural pathway through the pyramidal tracts between the upper motor pyramidal cell neurons in the neocortical motor strip and either cranial nerve lower motor neurons in the brain stem or anterior horn cells in the spinal column. Because of this monosynaptic upper to lower motor neuron innervation, these behaviors can be "volitionally" (i.e., other than spontaneously) executed.

The single movements of volitional Type II behaviors, especially when they are volitionally combined, are more functionally diverse (i.e., in their consequences or outcomes) than the heritable, spontaneous Type Ia single-behavior simple reflexes previously discussed. Great ape Type II finger, hand and arm movements, because of how they are innervated, can also communicate more nuanced information and acquire new information through learning better than the great apes can with their vocalizing muscle movements. Through observation one can see that finger movements in monkeys, lemurs, raccoons, sea otters, and some other movements in a limited number of other vertebrate taxa also have rudimentary Type II behavior features but not to the same degree of fine motor control as in the volitional finger and hand movements of the great apes and humans. Extant old world monkeys and apes also have volitional, Type II behavioral flexibility in their lips, jaw, tongue, and lower facial muscles, which they primarily use for feeding.

To repeat for emphasis, for the same definable function (e.g., saying the general name for people who give birth to babies), instances of Type II behaviors, which produce such sounds, do not have phylogenies (evolutionary histories). One cannot ask Tinbergen's, "What is the phylogeny of an instance of a coordinated pattern of behavior that makes the vocal sound "mother", even though to an outside observer the instance of Type II behaviors and the sounds they make (i.e., "mother") seem to be just as "physical" as bones, which do have phylogenies. The reasons why are subtle and take a certain degree of sophistication to understand.

VI. TYPE I HUMAN ("OUR") EXPRESSIVE VOCALIZING BEHAVIORS AND SOUNDS

Although not all Type I behaviors are expressive behaviors, all Type I human vocalizing behaviors and vocal sounds are expressive. They communicate one's mood (specific readiness to act), often interpreted as a need, to a conspecific (member of the same species).

Type I human expressive vocalizing behaviors

Ploog (2002) reviewed the neuroanatomical bases for two different neuroanatomical systems in the human brain that produce vocalizing behaviors, which correspond to the systems that produce Type I and II vocalizing behaviors, although Ploog did not use these new terms. The first system, which in this new terminology produces Type I expressive vocalizing behaviors, is very old. It is found in non-human primates (and other vocalizing vertebrates) as well as in humans. It involves limbic-structure CPGs whose neural pathways funnel into the peri-aqueductal grey of the midbrain and innervate the vocalizing muscles multi-synaptically through the extra-pyramidal system to produce Type I vocal sounds. Ploog points out that if this matrix is destroyed, all land living vertebrates except human beings become mute. Some Type I expressive vocalizing behaviors, like those that produce human laughing, also are associated with sub-cortical, rhythmic-generating CPGs.

Type I human expressive vocal sounds

The Type I human expressive vocal sounds themselves

Just like Type I behavior, Type I vocal sound, when it comes out of a human mouth, is "a heritable, coordinated pattern of air vibration, which in a natural or semi-natural environment, can be defined by its form for a species-universal function." Being definable by form for a species-universal function, Type I vocal sounds are species-universal in form for that species-universal function.

Function of Type I human expressive vocal sounds

Type I vocal sounds, which Ackermann et al. (2014) call "nonverbal affective vocalizations", communicate one's specific mood, often interpreted as a need, to conspecifics. Instead of the expressive signal being a particular muscular configuration, such as an expressive behavior facial affect, the signal is a readily distinguishable (i.e., definable by form) vocal sound with species-universal meaning. Certain types of strong emotion, negative or positive, can grade the vocal signal by changing the sound's acoustic structure in humans as well as in the bonobo (Clay, Archbold & Zuberbühler, 2015).

The vocal signal is graded by modulating volume, pitch and frequency of repetition, thereby being able to communicate more nuanced information than expressive muscle movements (i.e., affects) of the face. Type I human vocal sounds attempt to get conspecifics to do what is at least in the best survival and reproductive interest of the signal's sender (Hauser, 1997). All the human Type I expressive vocal sounds have functions that are almost always synonymous with their species-universal and species-

recognizable meanings. The meanings of human species-universal, Type I vocal sounds (e.g., laughs, screams, moans, groans, etc.) are quite interesting but are beyond the scope of this article, as are more volitional production of these vocal sounds, such as intentional laughing, which humans can distinguish from spontaneous laughing (Bryant and Aktipis, 2014).

Display rules for Type I human expressive vocal sounds

Although the CPGs of Type I vocal sounds are sub-neocortical, in humans the frontal and pre-frontal neocortices influence their spontaneity through culture-specific display rules similar to how culture-specific display rules control the execution of Type I non-vocal facial expressive behaviors (Eibl-Eibesfeldt, 1989; Ekman & Friesen, 1975). Persons with damage to their frontal and prefrontal neocortices, as well as persons under the influence of large amounts of alcohol, can become disinhibited and have culturally inappropriate laughing, crying and grunting-type anger outbursts (Ahmed & Simmons, 2013; Rinn, 1984).

The Type I human expressive vocal sounds as stimuli

Type I vocal sounds are in the "physical" ontological realm of mass, energy, force, space, time and information. They can be identified individually on an oscilloscope, where they can be defined by their form. As an example, human screams occupy a privileged, attention getting "rough niche" (30-150 Hz modulation rates) and also "light up" specific subcortical structures on fMRI in the human brain when heard (Amal et al., 2015). Type I vocal sounds, like crying and screaming, are classical ethological Releasing Stimuli and usually elicit a species-universal appetitive approach behavior response in an adult recipient, especially if the sounds are coming from a juvenile.

Heritability and ultimate causation of Type I human expressive vocal sounds

Type I vocal sounds themselves can be direct objects of natural selection, just like an anatomical structure, such as a bone. Type I behaviors and bones are both structural design features. Both can be adaptations. The heritable Type I vocal sounds act as proxies for the dedicated, heritable, coordinated patterns of Type I vocalizing behaviors that produce them; and the heritable, Type I vocalizing behaviors that produce them act as proxies for the heritable CPGs that produce them; and the heritable CPGs act as proxies for DNA that produces them. Within ontological pluralism, this is a "physical" causal chain.

Because, as previously explained, in behavioral biology causation does not cross the form/function ontological realm in ontological pluralism, Type I vocal expressive sounds can only be proximately caused by other "physical" entities (including conditioned physical entities) in this same physical ontological realm. These physical entities also include classical ethological Releasing Stimuli coming in by any of the senses. Some Type I vocal sounds can also be internally generated without external Releasing Stimuli by other "physical" elements within the body, such as when a hungry or colicky infant cries.

Development of Type I human expressive vocal sounds

Owren et al. (2010) proposed two different organizing principles for the development of vocal sounds in both non-human primates and human beings. They called them the production-first and the reception first types of development. Here we will deal with the production-first type. In Part VII we will deal with the reception-first type.

The difference between how these two types of vocal sounds develop has to do with the role of auditory-motor experience in each case. Essentially, Type I vocal sounds are spontaneous, affectively driven, species-universal sounds that don't require reception of sensory auditory input from conspecifics for their development. With the exception of occasional simple novel sounds in wild orangutans (Hardus et al., 2009; van Schaik et al., 2003) and the recent finding of a captive adolescent orangutan demonstrating his ability to learn how to control his vocal fold musculature in the laboratory in a Type II way (Lameira et al., 2016), all non-human primate and human Type I vocal sounds develop as production first sounds.

VII. TYPE II HUMAN ("YOUR") INTENTIONAL VOCALIZING BEHAVIORS AND SOUNDS

Type II human intentional vocalizing behaviors

Breath control

To produce any type of Type II vocalizing behavior, one needs a sufficient degree of neocortical (i.e., volitional or non-spontaneous) control over breathing, not just in the ability to subtly move muscles but also the ability of these muscles to subtly respond to ontogenic learning. All land mammals have an automatic, reflexive response to not breathe when their heads are held underwater, which is a different type of breath control from the nuanced, ontogenically learned breath control required to speak a human symbolic language.

Compared to australopithecines and Homo ergaster, Homo sapiens and Homo neanderthalensis both have expanded thoracic vertebral canals, where nerves innervating the breathing muscles transverse the spinal cord. This anatomical finding suggests that H. sapiens and H. neanderthalensis are and were more adapted to volitionally generated vocal behaviors that potentially were capable of vocal ontogenic learning compared to earlier hominids (MacLarnon & Hewitt, 1999).

It used to be thought that the great apes' lack of neocortical (i.e., volitional) control of their breathing was a major reason they did not have vocal symbolic speech. However, Koko, the human fostered gorilla, showed breath control playing wind musical instruments (Perlman et al., 2012).

Innervation

The specific details of the neural pathways underlying human Type II vocalizing behaviors are complex (Jürgens, 2002; Ludlow, 2006). Ploog (2002) simplified things by calling the Type II vocalizing behavior innervations "the neocortical voice pathway." It travels from upper to lower motor neurons in the monosynaptic pyramidal tract and is

indispensible for voluntary control of voice. It also involves cortical-striatal-cortical circuits (Lieberman, 2002). Destruction of it in humans makes speech impossible.

Owren et al. (2010), makes the point that when a species with an existing vocal system evolves a new functionally distinct vocalization capability, it occurs through the emergence of a second parallel neural pathway rather than through expansion of the existing circuitry, which would have been the extra-pyramidal pathway.

In the case of human Type II intentional vocalizing behavior, such a "new" pathway already existed in the monosynaptic, pyramidal tract innervations of the fingers, hands, lips, jaw, lower face, and tongue of the last common ancestor between human beings and the apes. All that needed to happen for human Type II intentional vocalizing behavior to have evolved, at least at the muscle movement level, was that some of the cranial nerve, secondary motor neurons in the brainstem - which were receiving afferents from the monosynaptic pyramidal tract to innervate the lips, tongue, jaw, and lower facial muscles - needed to make contact with the muscles of the human vocal folds. Other than the cricothyroid (laryngeal) muscle, which is innervated by a branch of the Xth cranial vagus nerve called the superior laryngeal nerve, all of the other laryngeal muscles used in vocalizing are innervated by another branch of the Xth cranial vagus nerve, the recurrent laryngeal nerve, whose asymmetrical and circuitous route to make contact with the laryngeal muscles is often cited by evolutionary apologists as evidence against intelligent design (Dawkins, 2010).

Ackermann et al. (2014) have proposed a two stage evolutionary model by which this occurred. The first stage is what they call the monosynaptic refinement of the projections of the motor cortex to the brainstem nuclei that control laryngeal muscles, which they contextualize within the phylogenetic trend associated with increasing brain size during hominin evolution. The second stage was the vocal-laryngeal elaboration of the cortico-basal ganglia circuitries (Lieberman, 2002), now believed to be driven by a human-specific FOXP2 mutation.

Ploog (2002), writing before some of the newer work on FOXP2 mutations, believed that the last step in the evolution of human speech was an augmentation of the fiber portion of the pyramidal tract synapsing directly with the lower motor neuron nuclei for the vocal cords and tongue, making volitional control of vocal behavior possible.

Fitch (2010, p. 351, in the caption to Figure 9.2) explains the neuroanatomy as the "Kypers/Jürgens hypothesis". It is worth quoting (almost) verbatim.

> Direct connections between the cortex and vocal motor neurons are hypothesized to underlie speech motor control in humans. The indirect connections from lateral motor cortex to brainstem interneurons are typical of most mammals, where cortical neurons make no direct connections to the actual motor neurons that control the muscles of the tongue and larynx. In addition to such indirect connections, primates add direct connections from the cortex to the motor neurons controlling the tongue, jaw, and lips. Only humans, among primates, possess direct connections to the laryngeal motor neurons that control the muscles of the larynx.

To give a specific example, in contrast to humans, Old World rhesus monkeys have no direct monosynaptic connection between the upper motor neurons in the laryngeal motor cortex and the brain stem nucleus ambiguus, whose lower motor neurons control the laryngeal muscles (Simonyan & Jürgens, 2003).

Compared to old world monkeys and chimpanzees, prosimians and New World monkeys have only (non-pyramidal) polysynaptic corticobulbar projections to all their brain stem lower motor neurons, including those that innervate the orofacial muscles (Sherwood (2005). They can make expressive behavior facial affects but can't voluntarily control the muscles of their face. In the squirrel monkey, a New World primate, electrical stimulation of limbic, thalamic, hypothalamic, and brainstem structures elicit various extrapyramidally-mediated species typical Type I vocal calls; but no such species-typical Type I vocal calls could be elicited by stimulation of the neocortex (Jürgens & Ploog, 1970), which is neuroanatomically why their sounds are not considered volitional.

Although the rudiments of a direct cortico-bulbar connection from the motor cortex to the nucleus ambiguous (which innervates vocalizing muscles) was anatomically identified by Kypers (1958) in a chimpanzee; and both New World squirrel and Old World rhesus monkeys posses some non-direct neocortical representation of the laryngeal muscles bordering the area associated with the lips and tongue (Hast et al., 1974), among extant primates, it has been presumed that only in Homo sapiens is there a direct, monosynaptic, functional cortico-bulbar, pyramidal-tract-mediated, connection between the upper motor neurons in the laryngeal representation on the neocortical motor strip and the lower motor neurons of the nucleus ambiguous and hypoglossus, which send efferent cranial nerve axons to the laryngeal muscles and tongue. Such a pathway can initiate voluntary vocal behavior. Recently, this presumption has been challenged.

In addition to some group-novel sounds recorded from orangutans in the wild (Hardus et al., 2009; van Schaik et al., 2003), more recently Lamerira et al. (2016) have shown that a single, captive raised, adolescent male orangutan was able to exercise some degree of learned, volitional pitch control over his vocal folds to produce sounds that were beyond the Type I species-specific vocal repertoire. This is important because individual-unique sounds can be made in orangutans with extra-vocal cord movements in the tongue and lips.

Thus, based on very limited data from Kypers' (1958) anatomical studies on a chimpanzee and Hast et al.'s (1974) anatomical studies on squirrel and rhesus monkeys, and both field and laboratory studies on the orangutan, it is at least possible that a common ancestor to the orangutan, chimpanzee, and human, who would have existed 11-16 million years ago, could have had some degree of volitional control of the vocal cords, which would have been the neural innervations to create Type II intentional vocalizing behaviors and sounds. This ability is found today in a few living mammalian taxa (e.g., bats, cetaceans, pinnipeds, elephants) and some taxa of birds, just not among the extant primates other than Homo sapiens and possibly the orangutan.

When many people think of human speech, they often think of Broca's speech area of the human brain, which is in the dominant, posterior inferior frontal lobe. Neurons in Broca's area, after reverberations with basal ganglia (Lieberman, 2002), send efferent output to the laryngeal muscle representations on the motor strip to initiate syntax-containing vocal speech (Ingram, 2007). Broca's area neurons also influence and coordinate the body language that accompanies symbolic speech (Skipper et al., 2017).

Response to learning

Not only does there need to be a direct, monosynaptic pyramidal-tract-mediated connection between the upper motor neurons on the motor strip of the neocortex and the articulation-generating lower motor neurons in the brain stem and the controlled-breathing-generating lower motor neurons in the spinal cord to produce Type II vocal speech, all these muscles of speech have to be very capable of responding collectively to learning by which pitch and resonant frequency (formant) and other features of vocalization can be changed. Skinner (1986) conceptualizes this as the vocal musculature coming under operant control. The great apes cannot learn to make different pitch sounds vocally but, as previously mentioned, human beings, bats, cetaceans, pinnipeds, elephants and some birds can (Jarvis, 2007).

From hand to mouth

Because Type II non-vocalizing behaviors are so evident in the fine movements in the fingers and hands of the great apes, and because the mirror neuron system in primates is essentially a hand-to-mouth feeding system, an argument has been made that human speech evolved by transferring the same type of pyramidal-tract-mediated, monosynaptic, fine muscle control in the hands and fingers (as well as the lower face, lips and tongue) to the breathing and vocalizing musculature (Gentilucci & Corbalis, 2006). As evidence for this, homologous facial/vocal displays are used similarly in chimpanzees and bonobos, but manual gestures, which can be both flexible and volitional, varied enormously both between and within these two species. Also, in humans, speech-accompanying manual gestures can be both species-universal (Church et al., 2017) as well as culture-specific in form and meaning (Morris, et al., 1979). Even though chimpanzees cannot learn to speak or even learn to make different pitch sounds vocally, they can learn American Sign Language with their hands (Pollick & de Waal, 2007). A temporal connection between the development of human vocal speech and manual gestures is found in ontogeny as well as in phylogeny (Bates & Dick, 2002).

Type II human intentional vocal sounds

The Type II sounds themselves

As compared to the spontaneous, nonverbal affective vocalizations of Type I sounds, Ackermann et al. (2014) call Type II intentional sounds simply "articulate speech". They could also be called (more) "volitional speech." Let's again take as an example the class of Type II functionally defined vocal sounds that symbolize people who give birth to babies in different cultural groups. All instances of these sounds have the same meaning. On a species-wide basis, for a particular class function, all instances of the sounds can only be described by their form in the different human languages.

Instances of Type II intentional vocal sounds can easily be modified in form through ontogenic learning, which is how we learn a human symbolic language. It is common knowledge that pre-pubertal children are much better at this than adults. Pre-pubertal children can learn a new language much faster than an adult and can learn to speak a new language without an accent of their native language. Pubertal hormones bathing the child's brain close this critical period. Tinbergen type questions are not even considered by most linguists for this phenomenon. The Type II intentional vocal sounds are all the

sounds of different human symbolic languages that are used to communicate non-emotional-laden information between conspecifics. Once emotions are concomitantly expressed by the sounds, the Type II intentional vocal sounds also become Type II expressive vocal sounds through prosody. When this occurs the sounds also become more spontaneous.

Development of the Type II sounds

In the two organizing principles for the development of vocalizing behaviors in both non-human primates and human beings that were developed by Owren et al. (2010) and were previously discussed, Type II intentional vocalizing sounds would be developmentally acquired by "reception-first mechanisms." This means that these sounds require sensory auditory-motor experience to operantly develop during the lifetime of the individual. Among the primates, such sounds are found in human beings and at least to date, only in one orangutan in the laboratory.

Evolutionary history of Type II sound prosody

As explained, in order for Type II sounds to have evolved, there had to have been direct, monosynaptic, pyramidal-tract-mediated innervations from the neocortex to the breathing and articulation musculature. Whereas most scholars believe this occurred in the service of Type II symbolic speech in Homo sapiens, there is an intriguing, alternative hypothesis recently proposed by Pisanski et al. (2016). This alternative hypothesis, which will be explained using the new Type I and II terminology, is that the ability to modulate pitch and resonant frequency evolved first as a spontaneous, sexually selected ability to affectively influence extra-pyramidal-tract-mediated Type I sounds. As an example, think of the facultative expression of both male and female, human Type I vocal copulation sounds today where pitch and resonant frequency can be volitionally altered. Low pitched Type I sounds are interpreted as both more masculine and more dominant. High pitched sounds are interpreted as more feminine and more submissive. Male and female sexual behaviors are highly connected phylogenetically to vertebrate dominance and submission (Medicus and Hopf, 1990).

Being able to volitionally, rather than spontaneously, alter the pitch of one's Type I sounds through the monosynaptic pyramidal tract could have evolved as a way for individuals to exaggerate socially relevant signals. By volitionally making lower pitched grunting sounds, males would give the impression of being larger and more dominant. By volitionally making higher pitched pleasure-acknowledging sounds, females would give the impression of being smaller and more submissive. Once the monosynaptic pyramidal tract made contact with the vocal fold musculature for deceptively altering the pitch and resonant frequency of Type I sounds, this created the condition by which Type II symbolic speech could occur following the human-specific FOXP2 mutation and the reverberating cortico-basal ganglia articulate speech circuitry and the re-functioning of what is now called Broca's speech area that generates syntactic symbolic speech.

VIII. HOW TYPE II INTENTIONAL BEHAVIORS AND SOUNDS CAN ACQUIRE EXPRESSION

Giving expression to Type II vocalizing behaviors

There are extra-pyramidal, polysynaptic neural connections between the more primitive, emotion-generating anterior-cingulate limbic CPGs and the pyramidal tract pathways used to produce Type II intentional vocalizing behaviors and speech. These pathways, which also involve the non-dominant parietal lobe, produce the emotional (i.e., prosodic) aspects of neocortically generated, Type II human symbolic speech (Cramon & Jürgens, 1983; Mitchell et al., 2016; Ploog, 2002). More will be said about how this works under "Giving expression to Type II human vocal sounds", which follows.

Giving expression to Type II vocal sounds

Let's take as an example instances of the Type II intentional vocal sounds used in an agonistic verbal argument to produce "Go to Hell!" or "Fu*k you!" in English. The words themselves are produced by the same Type II intentional vocalizing behaviors that would be used to produce "It is three o'clock" with the exception that the prosody would be different. Prosody influences the pitch, length of sounds, loudness, emphasis, and timber of speech sounds. Prosodies of Type II vocal sounds are species-universal in meaning "our" sounds. Even when just listening on the radio one can tell that someone vocalizing in a non-understood foreign language is in an angry mood by the rate, volume and pitch of words spoken. Type II expressive vocal sounds are not limited to angry insults. There are also soft and higher pitched Type II expressive vocal sounds used when talking to babies in "motherese."

Spontaneous and volitional Type II expressive vocal sounds

Whereas prosody is spontaneously (i.e., non-volitionally) expressed by spontaneous Type I vocal sounds (e.g., all genuine screams are loud and high pitched), in Type II vocal sounds the prosody can be either spontaneously expressed, as when one is genuinely emotional about what one is saying; or, the prosody can be more intentionally expressed as in some religious and political rhetoric, where it is used by the speaker to manipulate the listeners' emotions.

Expressive body language accompanies Type II expressive vocal sounds

Prosodically modified human expressive Type II vocal sounds are almost always accompanied by spontaneous, mood-revealing body language expressive behaviors. They include Type I facial expressive behaviors (e.g., angry affect) and various spontaneous species-universal hand gestures (Church et al., 2017), which are influenced by efferent output from Broca's area (Skipper et al., 2017). Interestingly, when Type II expressive language is executed in writing, as in digital media communication, capital letters, explanation marks, and emoticons can be used to symbolically represent the prosody and body language of the sender.

IX. SUMMARY

Type I and Type II human vocalizing behaviors and sounds are summarized in Table 1.

Table 1. Main Characteristics of Type I and II Vocalizing Behaviors and Sounds

	Type I Vocalising Behavior	Type I ("Our") Vocal Sounds	Type II Vocalising Behavior	Type II ("Your") Vocal Sounds
Definable by Form in All Individuals for the Same Function Cross Cultural	Yes	Yes	No	No
Describable by Form in All Individuals for the Same Function Cross Cultural	Yes	Yes	Yes	Yes
Definable by Function Cross Cultural	Yes	Yes	Yes	Yes
Describable by Function Cross Cultural	Yes	Yes	Yes	Yes
Species Universal in Form for the Same Function Cross Cultural	Yes	Yes	No	No
Heritable Coordinated Pattern of Movement or Sound Cross Cultural	Yes	Yes	No	No
Very Easily Modified by Ontogenic Learning That Is More Than Cultural Variation of Kind on a Common (i.e., species-universal) General Theme	No	No	Yes	Yes

X. CONCLUSION

A biological classification has been proposed for an ethology of human vocalizing behaviors and the sounds they produce. The use of the word "toward" in the subtitle of this article acknowledges that this is just a first attempt. Such an endeavor, when improved and refined, could serve as a foundation for higher order endeavors, such as being able to contribute more, from a human ethology perspective, to our understanding of the evolution of human communication, speech, linguistics, and language development, such as the behavioral changes associated with when and how species-universal infant babbling turns into culture-specific human speech..

Human vocalizing behaviors and the sounds they make are just as amenable to ethological study as other behaviors if the observable and measurable vocal sounds are used as proxies for the behaviors that produce them. Some of the methods would need to be different. Tinbergen's classic questions, only some of which are applicable to Type II behaviors and vocal sounds, would have to be just locally applied in a particular culture (language group). Take for example the use of the prosodic expression of deity and

expletive curse words by Type II expressive vocal behaviors and sounds when something very terrible just happened or is about to happen in the western industrialized democracies. Three of Tinbergen's four questions are applicable, albeit limited to a particular culture (language group). One could then investigate if the findings are applicable in the same contexts in other non-western cultures (language groups).

For example, the Type II vocalizing behaviors that produce the prosodically affected "Oh, my God!" or "Fu*k you!", although culture (language group) specific, might have the same answers to the first three of the four Tinbergen (1963) classic questions across cultures: (1) proximate causation, (2) survival value (adaptedness), (3) ontogeny (development). The same with what appears to be a human tendency, at least in the western industrialized democracies, to spontaneously utter a Type II vocalizing-behavior-generated, expressive explicative having to do with excrement or copulation when hitting one's thumb when hammering or dropping and breaking something very valuable. Why and how is the specific semantic content generated?

There are many Type II vocalizing behaviors and sounds that are similar by analogy across human culture groups. They include those that are associated with comments on sneezing, greeting, departure, and toasting ceremonies, as well as the vocal sounds associated with formal and informal social interactions and with social interactions with peers versus those above or below one in a social hierarchy, as well as terms of endearment among lovers, apologies, acknowledgements, common courtesies (i.e., "Thank you" in English) etc.

Human proceptive courtship behaviors in heterosexual, reproductive age women, especially when in the presence of eligible males with relatively high mate value, are almost all spontaneous Type I non-vocal expressive behaviors (i.e., body language); but in the male, proceptive courtship is also executed through Type II intentional vocalizing behaviors and sounds. It has been said, somewhat tongue in cheek, that in heterosexual human courtship women are as vulnerable with their ears as men are with their eyes.

How (and more interestingly, "Why?") did spontaneous, species-universal, heterosexual human female proceptive body-language courtship displays evolve together with intentional, culture-specific Type II vocal components in the male? We appear to be the only primate that uses Type II intentional vocalizing behaviors in courtship. Is one possible reason because Type II mediated vocal gossip is the human analogy (i.e., same function but different evolutionary history) as Type II non-vocal hand and finger grooming behaviors that are used in other extant primates to increase the probability of copulating (Dunbar, 1998)? Type II mediated gossip-transmitted information, also used in courtship, could be thought of as a nuptial gift. Males also make status references about themselves with Type II behaviors and sounds. What are the rules or laws that govern the execution of interacting Type I non-vocal and II vocalizing behaviors (identified through their proxy sounds) in courtship and other essential human acts? That is a question that human ethologists are in a unique position to answer.

There is also the potential to learn something about normal human vocalizing behaviors and the sounds they make by studying the abnormal. Tourette Syndrome is where spontaneous, tic-like Type I and Ia behaviors often are combined with spontaneously executed, expressively vocalized, culture-specific Type II expletive curse words (Martino & Leckman, 2013). How is the specific semantic content spontaneously generated with the same cursing theme in different human languages? Perhaps one clue

is that non-Tourette patients with lesions (e.g., strokes) in their motor cortex lose the ability to generate intentional Type II speech but retain the ability to vocally curse (Van Lancker and Cummings, 1999). Positive emission tomography (PET) brain scans during the tics and involuntary cursing speech sounds of persons with Tourette's Syndrome (Stern et al., 2000) show interesting findings that are beyond the scope of this article.

In closing, a conceptual and methodological framework, advanced from the foundation of classical ethology and within which human vocalizing behaviors and the sounds they make can be addressed, has been proposed. There is much unchartered territory having to do with the vocalizing behaviors and sounds that humans make. And most important, there is no longer a reason to treat humans as though they don't make such sounds in future behavioral studies in human ethology. We now have a first order approximation of what these human vocalizing behaviors and the sounds they make are, broken down into two much more homogeneous objects of study. The next steps are to learn more about what rules or laws govern how they develop, how they are executed proximately, how they interact, and what they do functionally and adaptively from a unique, human ethology perspective.

REFERENCES

Ackermann, H., Hage, S.R., & Ziegler, W. (2014). Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary Perspective. *Behavioral and Brain Sciences, 37*, 529-604. DOI.

Ahmed, A. & Simmons, Z. (2013). Pseudobulbar affect: prevalence and management. J*ournal of Therapeutics and Clinical Risk Management, 9*, 483-489. DOI

Allis, C.D., Caparros, M-L., Jenuwein, T., & Reinberg, D. (Eds.), (2015). *Epigenetics, Second Edition*. Long Island, NY: Cold Spring Harbor Press.

Amal, L.H, Flinker, A., Kleinschmidt, A., Giraud, A-L., & Poepper, D. (2015). Human screams occupy a privileged niche in the communication soundscape. *Current Biology, 25*(15), 2051-2056. DOI

Arshavsky, Y.I., Deliagina, T.G. & Orlovsky, G.N. (1997). Pattern generation. Current *Opinions in Neurobiology, 7*(6), 781-789 DOI .

Bates, E. & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology, 40*, 293-310. DOI

Bryant, G.A. & Aktipis, C.A. (2014). The animal nature of spontaneous human laughter. *Evolution and Human Behavior, 35*, 327-335. DOI .

Bucher, D. et al. (2015). Central Pattern Generators. In *Encyclopedia of the Life Sciences*, eLS. Chichester: John Wiley & Sons, Ltd. DOI

Church, R.B., Alibali, M.W., & Kelly, S.D. (Eds.) (2017). *Why Gesture? How the hands function in speaking, thinking and communicating*. Amsterdam: John Benjamins Publishing Company.

Clay, Z., Archbold, J., & Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behavior. *PeerJ , 3*:e1124. DOI

Cramons, D. von & Jürgens, U. (1983). The anterior cingulate cortex and the phonatory control of monkeys and man. *Neuroscience and Biobehavioral Reviews, 7* (3), 423-425, DOI

Crowson, R.A. (1970). *Classification and Biology*. New Brunswick, New Jersey: Aldine Transaction.

Dawkins, R. (2010). *The Greatest Show on Earth: Evidence for Evolution*. New York: Free Press.

Dennett, D.C. (1998). *The Intentional Stance*. Cambridge, Massachusetts: The MIT Press.

Dimitrijevic, M.R., Gerasimenko, Y., & Pinter, M.M. (1998). Evidence for a spinal central pattern generator in humans. *Annals of the New York Academy of Sciences, 860*(1), 360-376. DOI

Dunbar, R. (1998). *Grooming, Gossip, and the Evolution of Language*. Cambridge, Mass.: Harvard University Press.

Eibl-Eibesfeldt, I. (1975). *Ethology: The Biology of Behavior, 2nd Edition*. New York: Holt, Rinehart and Winston, Inc.

Eibl-Eibesfeldt, I. (1989). *Human Ethology*. New York: Aldine de Gruyter.

Ekman, P. & Friesen, W.V. (1975). *Unmasking the Face*. Englewood Cliffs, NJ: Prentice-Hall.

Ewert, J-P. (1980). *Neuroethology: An Introduction to the Neurophysiological Fundamentals of Behavior*. Heidelberg: Springer-Verlag.

Feierman, J.R. (2009a). How some major components of religion could have evolved by natural selection. In E. Voland & W. Schiefenhövel (Eds.), *The Evolution of Religious Mind and Behavior* (pp. 51-66). Berlin: Springer-Verlag.

Feierman, J.R. (2009b). The evolutionary history of religious behavior. In. J.R. Feierman (Ed.), *The Biology of Religious Behavior: The Evolutionary Origins of Faith and Religion* (pp. 71-86). Santa Barbara, CA: Praeger/ABC-CLIO.

Fitch, W.T. (2000). The evolution of language: A comparative review. T*rends in Cognitive Science, 4*(7), 258-267. DOI

Frank, M.G. & Ekman, P. (1996). Physiological effects of the smile. *Directions in Psychiatry, 16*(25), 1-7. Retrieved from https://www.paulekman.com/wp-content/uploads/2013/07/Physiological-Effects-Of-The-Smile.pdf.

Gentilucci, M. & Corballis, M.C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews, 30*, 949-960. DOI

Grillner, S. (1985). Neurobiological bases of rhythmic motor acts in vertebrates. *Science, 228*, 143-149. DOI

Grillner, S. (2006). Biological pattern generation: The cellular and computational logic of networks in motion. *Neuron, 52*(7), 751-766. DOI

Hast, M.H., Fischer, J.M., Wetzel, A.B., & Thompson, V.E. (1974). Cortical motor representation of the laryngeal muscles in Macca mulatta. *Brain Research 73*(2), 229-240. DOI

Hinde, R.A. (1970). *Animal Behavior: A Synthesis of Ethology and Comparative Psychology*. New York: McGraw-Hill.

Hardus, M.D., Lameira, A.R., van Schaik, C.P., Wich, S.A. (2009). Tool use in wild orangutans modifies sound production: A functionally deceptive innovation. *Proceedings of the Royal Society b 276*, 3689-3694. DOI

Hauser, M.D. (1997). *The Evolution of Communication*. Cambridge, Massachusetts: The MIT Press.

Immelmann, K. & Beer, C. (1989). *A Dictionary of Ethology*. Cambridge, Massachusetts: Harvard University Press.

Ingram, J.C.L. (2007). *Neurolinguistics: An Introduction to Spoken Language Processing and Its Disorders*. Cambridge: Cambridge University Press.

Jarvis, E.D. (2007). Neural systems for vocal learning in birds and humans: A synopsis. *Journal of Ornithology, 148*(1), 35-44. DOI

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience & Biobehavioral Reviews 26*(2): 235-258. DOI

Jürgens, U. & Ploog, D. (1970). Cerebral representation of vocalization in the squirrel monkey. *Experimental Brain Research, 10*(5), 532-554. DOI

Kuypers, H.G.J.M. (1958). Corticobulbar connections to the pons and lower brain-stem in man. *Brain, 81,* 364-388. DOI

Lameira, A.R., Hardus, M.E., Mielke, A., Wich, S.A., & Shumaker. R.W. (2016). Vocal fold control beyond the species-specific repertoire in an orangutan. *Scientific Reports, 6*, 30315. DOI

LeDoux, J.E. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster.

Lehrman, D.S. (1953). A critique of Konrad Lorenz's theory of instinctual behavior. *The Quarterly Review of Biology, 28*(4), 337-363. Retrieved from http://www.jstor.org/stable/2813453?seq=1#page_scan_tab_contents.

Lieberman, P. (2002). On the nature and evolution of the neural bases of human language. *Yearbook of Physical Anthropology, 45*, 36-62. DOI

Lieberman, P. (2006). *Toward an Evolutionary Biology of Language*. Cambridge, Massachusetts: The Belknap Press of Harvard University.

Llinas, R.R. (2001). *I of the vortex: From neurons to self.* Cambridge: MIT Press.

Longobardi, G., Buch, A., Ceolin, A., Ecay, A., Guardiano, C., Irimia, M., Michelioudakis, D., Radkevich, N., & Jaeger, G. (2016). Correlated evolution or not? Phylogenetic linguistics with syntactic, cogency, and phonetic data. In S.G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.). *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG 11)*. Retrieved from http://www.sfs.uni-tuebingen.de/~gjaeger/publications/evolang11ComparisonFinal.pdf

Lorenz, K.Z. (1974). Analogy as a source of knowledge. *Science, 185*(4147), 229-234. DOI

Lorenz, K.Z. (1981). *The Foundations of Ethology*. New York: Springer-Verlag.

Ludlow, C.L. (2006). Central nervous system control of the laryngeal muscles in humans. *Respiratory Physiology & Neurobiology, 147*(2-3), 205-222. DOI

MacLarnon, A.M. & Hewitt, G.P. (1999). The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology, 109*, 341-363. DOI

Martino, D. & Leckman, J.F. (2013). *Tourette Syndrome*. Oxford: Oxford University Press.

Medicus, G. and Hopf, S. (1990). The phylogeny of male/female differences in sexual behavior. In J.R. Feierman (Ed.). *Pedophilia: Biosocial Dimensions* (pp. 122-149). New York: Springer.

Mitchelll, R.L.C. (2016). Recruitment of language-, emotion-, and speech-timing associated brain regions for expressing emotional prosody: Investigation of functional neuroanatomy with fMRI. *Frontiers in Human Neuroscience*, 10.518. DOI

Morris, D. et al. (1979). *Gestures: Their origins and distributions.* New York: Stein and Day.

Owren, M.J., Amoss, R.T., & Rendall, D. (2010). Two organizing principles of vocal production: Implications for non-human and human primates. *American Journal of Primatology, 71,* 1-15. DOI

Panchen, A.C. (1992). *Classification, Evolution, and the Nature of Biology*. Cambridge: Cambridge University Press.

Perlman, M., Patterson, F.G., & Cohn, R.H. (2012). The human-fostered gorilla Koko shows breath control in play with wind instruments. *Biolinguistics, 6*(3-4), 433-444. Retrieved from http://www.biolinguistics.eu/index.php/biolinguistics/article/view/243.

Pisanski, K., Cartei, V., McGettigan, C., Raine, J. & Reby D. (2016). Voice modulation: A window into the origins of human vocal control. *Trends in Cognitive Science 20*(4), 304-318. DOI

Ploog, D. (2002). Is the neural basis of vocalization different in non-human primates and Homo sapiens? *Proceedings of the British Academy, 106,* 124-135. DOI

Pollick, A.S. & de Waal, F.B.M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences, 104*(19), 8184-8189. DOI

Rinn, W.E. (1984). The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin, 95*(1), 52-77. DOI

Rubin, M. & Safdieh, J.E. (2016). *Netter's Concise Neuroanatomy Updated Edition*. Amsterdam: Elsevier.

Sherwood, C.C. (2005). Comparative anatomy of the facial motor nucleus in mammals, with an analysis of neuron number in primates. *The Anatomical Record, Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology 287*(1), 1067-1079. DOI

Simonyan, K. & Jurgens, U. (2003). Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. *Brain Research, 974,* 43-59. DOI

Skipper, J.I., Goldin-Meadow, S., Nusbaum, H.C., & Small. S.L. (2017). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language 101*(3), 260-277. DOI

Skinner, B.F. (1986). The evolution of verbal behavior. J*ournal of the Experimental Analysis of Behavior, 45,* 115-122. DOI

Stern, E., Silbersweig, D.A., Chee, K.Y., Holmes, A., Robertson, M.M., Trimble, M., Frith, C.D., Frackowiak, R.S.J., Dolan, R.J. (2000). A functional anatomy of tics in Tourette Syndrome. *Archives of General Psychiatry, 57*(8), 741-748. DOI

Tinbergen, N. (1951). *The Study of Instinct*. Oxford: Oxford University Press.

Tinbergen, N. (1963) On the aims and methods of ethology. *Zeitschrift für Tierpsychologie, 20,* 410-433. DOI

Turner, J. (2012). Logic and ontological pluralism. J*ournal of Philosophical Logic, 41*(2), 419-448. DOI

Van Lancker, D.& Cummings, J.L. (1999). Expletives: neurolinguistic and neurobehavioral perspectives on swearing. *Brain Research Reviews, 31*, 83-104. DOI

van Schaik, C.P., Ancrenaz, M, Borgen, G., Galdikas, B., Knott, C.D., & Singleton, I. (2003). Orangutan cultures and the evolution of material culture. *Science, 299*, 102-105. DOI

von Holst, E. (1973). *The Behavioral Physiology of Animals and Man: Selected Papers, Vol. 1*. Coral Gables, Florida: University of Miami Press.

Zeifman, D. (2014). An ethological analysis of human infant crying: Answering Tinbergen's four questions. *Developmental Psychobiology, 39*(4), 265-285. DOI