The Origins of Spoken Language

Megan Terwilliger

SUNY New Paltz

Abstract

Research regarding the origins of spoken language is extremely broad and considers

many different topics in academia. These topics include: biology and anthropology, as

well as anatomy, non-linguistic neurological foundations, potential protolanguages, and

possible selective pressures that could have produced such a complex system of

communication. Considering these topics can assist in imagining what language may

have presented itself as throughout human ancestry, and how it could have developed into

the intricate system of modern language. Anatomically, the descent of the larynx,

increased thoracic innervation, disappearance of the laryngeal air sacs, and changes in

neuroanatomical structures could have served as exaptations or adaptations for spoken

language. Non-linguistic factors such as vocal learning and vocal imitation fostered new

ways of learning to communicate. Protolanguages could have taken a gestural, musical,

hybrid, or lexical form that increased the selective pressures to create today's language

system. Social changes, mainly increase in group size, had the potential to provide great

selective pressure for the development of more efficient communication. The

combination of some of these factors eventually fostered an environment for the human

vocal communication system to evolve into what it is today.

Keywords: communication disorders, evolution of communication, evolution of

language, evolution of speech, protolanguage

The origin of spoken language in humans is a much-debated topic in many disciplines including biology and anthropology. When considering the question of how spoken language developed in human ancestry, there is an abundance of information to take into consideration. First, there is the distinction between speech and language. Speech is the physical production of our spoken language system. It involves moving our articulators in the correct manner in order to produce the correct sounds in combination (American Speech-Language-Hearing Association, n.d.). In addition to articulation, speech also includes phonation (vibration at the vocal folds to produce voicing) and fine control of respiration. Language is more concerned with the use of our spoken system to convey the meaning we intend communicate (American Speech-Language-Hearing Association, n.d.). The processes of speech and language are extremely intertwined, making it difficult to consider them separately, especially from an evolutionary standpoint. As Locke (2017) explains:

> We cannot achieve a satisfying proposal on the evolution of language, which is generally spoken, without explaining how it became possible for our ancestors to use their voices flexibly and creatively, thus to make the kinds of sounds that, with other changes, would have allowed them to communicate symbolically. (p. 232)

In addition to the distinction between speech and language, there is other information to consider regarding the origins of spoken language. For example, evolutionary changes in anatomy allowed for spoken language to be both possible and more efficient than it may have been otherwise (Fitch, 2010). Also, cognitive adaptations supported language through novel thought processes that were vital to expression and

understanding (Fitch, 2010). There also are hypotheses of protolanguages to consider.

These protolanguages attempt to explain the different stages in which language could

have presented itself originally, and how it could have developed to reach the complexity

it has today.

Throughout these changes in the evolutionary history, it is believed that larger

communities with increased social demands were a major driving factor. New

environmental changes that drove humans to be more social are thought to have caused

selective pressures that resulted in many of the adaptations for spoken language, along

with other changes in the species (Kendon, 2017). This idea is a prominent topic of

discussion in the recent language evolution literature, potentially making sociality the

most important selective pressure in the development of our species' communication

system.

This paper will explore the major themes of research regarding the origin of

spoken language: anatomical changes, cognitive advances, hypotheses of protolanguages,

and selective pressures in hominin communities. This will function to provide an

overview of major themes in the field, and how these themes may work together to

provide plausible explanations for the evolution of spoken language. Throughout the

paper, these categories' relevance will be addressed and explained, beginning with the

anatomical changes that supported spoken language.

There were many different anatomical changes that supported the evolutionary

development of spoken language (Fitch, 2010). Three major anatomical features that are

assumed to have aided in speech production are the larynx, increased thoracic

innervation, and laryngeal air sacs. Major evolutionary changes in the larynx include: the

size exaggeration hypothesis, potential safety concerns of a descended larynx, possible

connection of descent of the larynx to the transition to bipedalism, and the evolution of

the thryoarytenoid (vocal) folds. Thoracic innervation will be discussed in relation to

breath control for speech. Less is known about the laryngeal air sacs, but the possible

reasons for their disappearance will be considered. Changes in neural structure will also

be discussed including information about cortical-laryngeal connections and Broca's

area.

The larynx is a major focus of speech production, specifically for the voicing of

sounds. The descended larynx that is present in humans allows for the "dynamically

changing resonances" and rapid articulatory movements required for speech (Ghazanfar

& Rendall, 2008, p. R458). However, there are other explanations suggested for the

adaptation.

There are two main arguments in the discussion of the larynx and its speech

adaptations. The first argument views laryngeal changes as exaptation. The exaptation

argument includes the size exaggeration hypothesis, the transition to bipedalism, and the

properties of the larynx itself. The second argument views changes in the larynx as

adaptations for speech and this argument is primarily centered on the safety repercussions

that the descent of the larynx caused in our species.

"The size exaggeration hypothesis for laryngeal descent holds that lowering

formants (for example by the retraction of the larynx) functions to increase the

impression of size conveyed by vocalizations" (Fitch, 2010, p. 321). Lowering the

formants, and therefore the dominant frequency range, would make animals appear to

have a larger body size, potentially decreasing their predation risk.

There is evidence in other mammalian species that formant values provide information about body size. It has also been determined that certain animals, "…perceive formants, without training, in their own species-specific vocalizations (Rendall et al., 1998; Fitch and Kelley, 2000; Hienz et al., 2004; Fitch and Fritz, 2006)" (Fitch, 2010, p. 321-322). The combination of these two facts would provide evidence that the impression of exaggerated body size could potentially be the cause of the descended larynx adaption, which later, according to this hypothesis, served as a preadaptation for speech in humans.

It is also thought the descent of the larynx could be a result of the transition to bipedalism. Since the orientation of the spine in relation to the skull had to change, the oral space had to be adjusted. With this, the tongue, "…is now thick, deep, and specially bent at its base due to the notable approximation of its opposite ancorages, the genial tubercles in the front and the hyoid bone in the back" (DuBrul, 1972, p. 633). Therefore, without the descent of the larynx, the base of the tongue could have potentially blocked off the entire airway (DuBrul, 1972).

Even in early mammals, there is evidence that the larynx and vocal tract are adapted for complex vocalizations. There is evidence that their anatomy included cricoid and arytenoid cartilages therefore suggesting a more differentiated thyroid cartilage (Wind, 1972). This suggests the development of thyroarytenoid (vocal) folds at this time. These folds made it possible for the creation of vibration (vocalization) when air moves through them and the opportunity for varying frequency (Wind, 1972). Therefore, "… the laryngeal and respiratory differentiation of mammals enabled them to produce more

frequent vocalizations that could be varied much more subtly in pitch and amplitude…"

(Wind, 1972, p. 623).

The development of thyroarytenoid folds does not necessarily serve as evidence

of a speech adaptation. Once evolved, these vocal folds must have been able to adduct

and abduct, because otherwise the animal would not have been able to breathe. These

movements could have been for additional protection against aspiration during

swallowing. However, the appearance of thyroarytenoid folds does suggest that the

variation in characteristics of vocalizations may have served some sort of adaptive

advantage. This may suggest that changes in the thryoarytenoid folds in mammals served

as a exaptation for speech.

There is an additional hypothesis that views the descent of the larynx as an

adaptation for speech, instead of an exaptation. Although there are many mammals with a

descended larynx that provide support for the size exaggeration hypothesis, it is difficult

to ignore the safety implications that the descended larynx causes for humans (de Boer,

2017). Humans not only have a lower larynx than nonhuman primates, "…but humans

also have a much bigger gap between the larynx and the velum than do other primates,

and even than do other mammals with permanently lowered larynges (increasing the risk

of choking on one's food: e.g. Heimlich, 1975)" (de Boer, 2017, p. 158). With this safety

concern, it is assumed that this gap was selected for in our evolutionary past, possibly for

a vocal function (de Boer, 2017). "Therefore, following Parker and Maynard Smith

(1990), who observed that optimization for a function often indicates selective pressure

related to that function, it seems likely that humans have undergone selective pressure

related to speech" (de Boer, 2017, p. 158). Therefore, configuration of the larynx within

the body is important. However, other aspects of anatomy also play an important role in this discussion.

The larynx is one of the more obvious parts of human anatomy that is connected to speech, but there are other anatomical features that are involved. For example, increased thoracic innervation is involved with breath control for speech. "The thoracic region of the spinal cord is involved in controlling respiration and the size of the thoracic vertebral canal gives an indication of the amount of innervation in this region of the spinal cord" (Ghanzanfar & Rendall, 2008, p. R458). When comparing modern humans to other extinct ancestors, it was found that humans have an expanded thoracic vertebral canal. Although there are other possible explanations, they are not possible because of "…evolutionary timing or because they are insufficiently demanding neurologically" (Ghanzanfar & Rendall, 2008, p. R458). It cannot be fully proven that this adaptation was for breath support, but it has been found that when vocalizing, nonhuman primates use a series of expirations and inspirations. This series of inhalations and exhalations contrasts with the way humans use breath support and vocalize entirely on an exhalation, possibly explaining this difference.

The last change in anatomical structure that could be connected to the evolution of speech is the laryngeal air sacs. There is little information known about the air sacs other than humans lost them somewhere in their evolutionary history. The primary hypotheses are they were used for amplification, for longer calling sounds, or to sound larger (much like a descended larynx) (Ghanzanfar & Rendall, 2008). If air sacs do in fact have a size-exaggeration function, this could be an explanation as to why they were lost since the descended larynx already was serving that function (Fitch, 2010). It has also

been discovered that air sacs could have potentially decreased intelligibility in

vocalizations, which could have been another reason for their disappearance in humans

(de Boer, 2012 as cited in de Boer, 2017). This decreased intelligibility hypothesis could

be linked to evidence in deaf speech, where it was found that increased resonance in the

chest cavity can decrease intelligibility (Boothroyd, 1987). The additional resonating

cavities these laryngeal air sacs could have provided could have increased resonance as

well, therefore decreasing intelligibility. Even with the information we have about

laryngeal air sacs, there is still much debate about their actual function and why humans

are lacking them.

      It is important to consider these changes in human anatomy and determine

possible selective pressures that influenced their development. However, it is possible

that some, many, or all of these features developed neither as adaptations or exaptations.

This is described by neutral theory. Neutral theory, "…suggests most of the genetic

variation in populations is the result of mutation and genetic drift and not selection"

(Understanding Evolution, n.d.). Therefore, although this discussion separates these

anatomical differences into finite categories, it is possible that there were other factors at

work.

      These three components of our anatomy: the larynx, increased thoracic

innervation, and the disappearance of laryngeal air sacs all contribute to the scientific

discussion surrounding the evolution of speech. Changes in the larynx such as its descent

and the appearance of thyroarytenoid folds could have served as exaptations for speech.

There is also the possibility that the increased risk of choking the descent of the larynx

caused proves that this evolutionary changed may have served as an adaptation for

speech since otherwise it would be extremely dangerous. Increased thoracic innervation

may have been an adaptation for better breath control during speech. The laryngeal air

sacs' potential redundancy or impact on intelligibility may have caused them to

disappear. It is also possible that the explanation for these changes are not relevant, and

there were no real direct selective pressures other than random mutations and genetic

drift. No matter the cause of these evolutionary changes, they all have an effect on

present-day spoken language. However, there are changes in human neuroanatomy that

are also important to consider.

The human neural structure is extremely similar to that of other mammals. More

specifically, the corticobulbar pathways that allow commands from the pre-central gyrus

to be relayed through the brainstem to the pons and medulla to control cranial nerves for

speech production is a similar pathway in all vertebrates. "Both the muscles of the larynx

and vocal tract, and their primary nervous supply, are shared virtually identically among

mammals, and many higher-order controlling structures in the brainstem and cortex are

also apparently shared" (Fitch, 2010, p. 347). The main differences in the neural control

of humans and other mammals exist at the highest level of control, otherwise known as

the "cortical control systems," and more specifically the "lateral cortical system" (Fitch,

2010).

"In particular, humans possess direct connections between the frontal motor areas

of lateral neocortex and important brainstem motor neurons, especially those in the

nucleus ambiguous involved in laryngeal control" (Kuypers, 1958; Deacon, 1992;

Jürgens, 1994 as cited in Fitch, 2010, p. 350). Damage to this area in humans causes total

loss of vocal control, while in other species it has no effect on vocalizations. There is

evidence in aphasic patients that these vocal connections are the connections associated with complex vocal imitation. Since some aphasic patients exhibit difficulty with controlled speech, yet still have appropriate emotional responses (i.e. crying, cursing, etc.), it is thought that these emotional responses could be preserved somewhere else in the brain. That would mean that these new connections are primarily linked to "new" vocal output or vocal output that is produced with this complex imitation ability (Fitch, 2010).

These connections are directly posterior to one of the brain's main speech areas known as Broca's area. Broca's area is directly involved with voluntary control of vocalizations in humans. However, it has not been found to have this same function in populations of monkeys. It has also been found as a major area of development during hominid evolution (Fitch, 2010). "For all of these reasons, neural changes in the general region of Broca's area have been seen by many commentators as critical in the evolution of human vocal motor control" (Fitch, 2010, p. 352).

Even with the abundance of information about Broca's area known today, it is still unclear what the exact function is. Broca's area is known to have some implications in motor control of the speech mechanism, but also has been shown to be a factor in comprehension, visual lexical processing, and phonological tasks (Fitch, 2010). With this uncertainty comes the same uncertainty in the role of Broca's area exact influence on the evolutionary development of speech, but it is clear that it definitely plays a role.

There have been differences noted between Broca's area in humans and in nonhuman primates. In the human brain, Broca's area is primarily located in the left hemisphere, where in nonhuman primates it is not primarily located in one hemisphere

(Friederici, 2017). Also, "…the connectivity between Broca's are and the superior temporal cortex is stronger in the human compared to the nonhuman primate brain" (Friederici, 2017, p. 44). These differences, due to Broca's area's and the superior temporal cortex's connection to language processing, could potentially be extremely important for the evolution of language (Friederici, 2017).

The cortical-laryngeal connections also provide humans with the unique ability to control their laryngeal musculature. It has been found that this is a major difference between humans and chimpanzees. Nonhuman primates can only control supralaryngeal articulators, not the actual laryngeal musculature (Hickok, 2017). "Thus, the development of voluntary laryngeal control has been argued to be the 'key innovation' in the evolution of speech" (Hickok, 2017, p. 58).

Overall, these connections are the best explanation of the development of speech that has been discovered. The connecting of these areas would have allowed further control over the larynx than what was present in earlier hominid species, creating greater coordination for phonation and speech (Fitch, 2010). The fact that these connections are not present in other nonhuman primates or mammals further exhibits the probable importance of these neural components in speech production (Fitch, 2010). It is logical to think that other cortical areas are possibly involved, but they have yet to be determined (Fitch, 2010). It is also possible that there are additional environmental or cultural factors to be considered.

Based on the data surrounding these anatomical changes, it appears that "…changes in the central nervous system, rather than peripheral vocal anatomy, were critical innovations in the evolution of speech" (Darwin, 1871, as cited in Fitch, 2010, p.

328). The changes that human anatomy went through were beneficial overall for clarity

and diverse production ability of speech, but languages are possible without it (de Boer,

2017). "For instance, the Rotokas language only uses 11 phonemic contrasts" (Firchow &

Firchow, 1969 as cited in de Boer, 2017, p. 161). This highlights the importance of the

non-linguistic neurological foundations that happened in order to have selective pressure

on the vocal anatomy (de Boer, 2017). As Fitch (2010) explains speech is, "…a

complex, learned signaling system involving rapid formant transitions, and the rapid

movements of the tongue and lips…" (p. 338). This complex system would clearly need

more than a strictly anatomical base to evolve. This is where the implications of the

complex neural system of our species becomes relevant.

Unlike the vocal systems found in nonhuman primates, the human vocal system

requires sophisticated control. "Speech requires much more than simply selecting among

a pre-existing set of innate calls" (Fitch, 2010, p. 339). It has been found that nonhuman

primates do have control over the use of their calls, but that control is limited (Fischer et

al. 2015 as cited in Fischer, 2017). Nonhuman primates do not have control over form or

structure of the call in action (Janik & Slater, 2000 as cited in Fischer, 2017). The use of

nonhuman primate calls helps to show that they are able to understand and use their calls

for their own form of communication, and the major difference between humans and

nonhuman primates in this realm has to do with the process of production (Friederici,

2009, Rilling et al., 2008 as cited in Fischer, 2017). For example, when humans get

louder, as a consequence their fundamental frequency increases. In this case, humans are

able to consciously control their fundamental frequency and decrease it to a typical level.

Nonhuman primates do not have this same type of control.

There are three main neural mechanisms that distinguish human vocalizations from primate vocalizations: vocal learning, vocal imitation, and complex vocal imitation. Although similar, these abilities provide humans with skills that are vital to speech and language. The first characteristic, vocal learning, supports the creation of new signals (Fitch, 2010). "Vocal learning is demonstrated when some aspect of call structure is modified due to specific experience with the environment (e.g. by imitating a perceived sound, or differentiating one's own vocalizations from it)" (Fitch, 2010, p. 339). This does not include any novel signals, but instead the altering of signals already pre-existing in the repertoire of that species (Fitch, 2010). For example, "In bats, the evidence for spontaneous vocal learning is limited to only a few species, and to the adjustment of a single call parameter of a species-specific call" (Jones and Ransome, 1993 as cited in Fitch, 2010, p. 339).

Speech requires the skill of vocal imitation, which is defined as, "…the capacity to incorporate into one's vocal repertoire novel sounds that are perceived in the environment" (Fitch, 2010, p. 339). Unlike vocal learning, this skill requires the vocalization to be completely different from vocalizations the organism already uses, and for this new vocalization to be added to the organism's repertoire (Fitch, 2010). Fitch (2010) also describes another form of this which he coined complex vocal learning. He claims this is the minimum necessary requirement to support a spoken language system, and it requires more skills than regular vocal imitation including, "…good memory for complex signals and an ability to differentiate their individual components (e.g. successive syllables)" (Fitch, 2010, p. 340).

Complex vocal imitation is thought to have been influenced by two main selective pressures. These selective pressures are sexual selection and territory maintenance. The influence of complex vocal imitation on sexual selection can be seen in songbirds. Songbirds, in order to attract a mate, must use complex song patterns that differ from other males but still identify them as a part of their own species. The fact that the songbird that has the most complex song is often the one that mates influences what is passed on to the next generation. This type of model is possibly similar to what the evolution of complex vocal imitation could have looked like in humans (Fitch, 2010).

There is also the role of territoriality in the development of complex vocal imitation. It was found both in songbirds and in humpback whales, using specific songs can help to keep other males out of the primary male's immediate area. It was found that more complex songs strengthen this effect. "Current data thus suggest that male song, in general, plays both an intersexual mate-attraction function and, simultaneously, an intrasexual male-repellent function" (Fitch, 2010, p. 342). The combination of these two effects could have potentially driven the evolution of complex vocal imitation, especially in males (Fitch, 2010). Complex vocalizations in females are mostly used for a territorial maintenance function, either singing alone or with a male with whom they share the area. Of species with similar abilities for both males and females, it has been observed that vocalizations can be used for social bonding or group communication (Fitch, 2010).

Through this data on vocal learning, there have been two potential models presented for the evolution of complex vocal learning abilities in humans. The first model involves separate development in both males and females. Using this viewpoint, the abilities would first be selected for in males by means of sexual selection, and females

would later acquire the abilities through either social bonding or territoriality pressures.

(Fitch, 2010). Conversely, in the second proposed model, males and females would

develop these abilities at the same time. This would suggest that the main selective

pressure in both sexes would be potentially social bonding or a related process (Fitch,

2010).

The neurological foundations of vocal learning, vocal imitation, and complex

vocal imitation are required for the development of modern-day human communication.

These mechanisms provide the base for the eventual transition to linguistic

communication. However, these skills exist in other organisms, suggesting that it is not

necessary for species with these neurological characteristics to use a linguistic

communication system. Therefore, these abilities may have evolved as exaptations, or co-

evolved with other adaptations to support the development of our spoken language

system.

Eventually, the exaptations and various changes in our human ancestors must

have led to some form of early language. These potential early languages are called

"protolanguages." Most researchers agree that some form of a protolanguage existed, but,

"…they disagree vehemently over the nature of protolanguage, and over how it

developed into modern language" (Smith, 2008 as cited in Zywiczynski, Gontier, &

Wacewicz, 2017, p. 3). These protolanguages can be separated by the different categories

of structure, function, and modality. These initial categories break down further into more

specific categories, outlining the different theories of the evolution of language through

protolanguages (Zywiczynski et al., 2017).

The four most prominent protolanguages in this debate are: the gestural protolanguage, the musical protolanguage, the multimodal protolanguage that utilized both gesture and vocalization, and the lexical protolanguage. These different ways of looking at the potential evolution of language help us to picture how such a development could have occurred, possible selective pressures involved, and the eventual transition to a more modern-day spoken language, in addition to the anatomical and non-linguistic neurological foundations discussed previously.

The gestural protolanguage hypothesis is partially centered around, "…the evolution of serial motor activity and its nervous control" (Steklis & Harnad, 1972, p. 447). The organs of serial motor activity in nonhuman primates are: the muscles of the eyes and head, the limb muscles, the vocal apparatus, and the facial muscles. Of these options, the vocal apparatus was used for the least amount of functions at this time (Steklis & Harnad, 1972). However, "…the neural control of the vocal activity of nonhuman primates is somehow not adapted to the kind of activity involved in language" (Steklis & Harnad, 1972, p. 447). Therefore, the vocal apparatus lacked the plasticity to learn new sounds. All vocalizations were assumed to be primarily fixed for each species of primate, similar to the way that primate vocal control was discussed previously.

Within the idea of a gestural protolanguage, it was assumed that since there was little control over the vocal apparatus, the use of the limbs for communication was a more functional starting place. At this time, the limbs were more easily controlled than the vocal apparatus. Even with the use of the forelimbs for communication, encephalization (increased neocortical control) still needed to take place to provide the opportunity for more complex motor activity (Steklis & Harnad, 1972). To evolve into speech and

language, this proposed gestural protolanguage would have involved not simply visual signals but a complex system. Unlike incidental primate vocalizations, these visual signals needed intention (Jaynes, 1972).

It has been proposed that limitations of a gestural communication system such as, not being able to see in darkness and not being able to simultaneously use one's hands for other tasks, eventually influenced the coupling of gesture and vocalization (Steklis & Harnad, 1972). Selective pressures such as bipedalism and tools have also been credited for the transition because of the need for the forelimbs for more involved activities (Jaynes, 1972). However, as explained below, these arguments have been discredited.

Many early explanations of the gestural protolanguage's transition to a vocal language have been invalidated. Although a spoken language would allow for communication in the dark, a signed language would allow for silent communication, which could be useful in hunting or avoiding being hunted, or for communication in environments with an excess of noise (Fitch, 2010). It is thought by the time language was developed that our ancestors had discovered fire, meaning that they could communicate using gesture in the light. Even if they had not discovered fire, there is no suggestion of any nocturnal behavior, suggesting that there would not be a large pressure to be able to communicate in the dark (Fitch, 2010). In regards to needing the forelimbs for more involved activities, it is important to note that, "…native signers are quite flexible in their use of sign in adverse conditions: they can sign with one hand, while driving, cooking, or operating equipment" (Fitch, 2010, p. 445). From a learning standpoint, "Most manual learning appears to occur via observation and imitation rather than vocal descriptions (Hewes, 1973)" (Fitch, 2010, p. 445).

Researchers of gestural protolanguage attempt to explain this gestural language's transition to vocal language. Both modern human's ability to communicate through only gesture and the great ape's capability to gesture far beyond their ability to use vocal language provide strong support for this argument. Advocates of this theory believe, "…a *primarily* gestural communication system once existed in our lineage, and that its existence was a *necessary precondition* for the evolution of spoken language" (Fitch, 2010, p. 465).

It has been found that ape gestures, although easily compared to human gesture, are not developed in the same way. Gestural development in apes, "… is a matter of exploring a latent repertoire conferred by biology" (Hobaiter & Byrne, 2011b as cited in Byrne & Cochet, 2017, p. 69). This suggests that ape gestures are not learned in the same way language is, but are genetically based. This biological component raises questions because it seems that nonhuman primates with common ancestors share some of their common ancestors' gesture repetoires. For example, chimpanzees and gorillas use some of the same gestures for the same functions (Byrne & Cochet, 2017).

This difference in ape and human gesture is highlighted in the way gestures are learned. As children learn to communicate, they also, "…gradually learn to convey intended meanings through gesture as a result of experience, either through social learning as a cultural product or through reinforcement learning from actions that initially do not involve communicative intention" (Byrne & Cochet, 2017, p. 70). This demonstrates that unlike apes, humans do not use the gestures that could potentially exist somewhere in our ancestry (Byrne & Cochet, 2017).

If somehow it was found that humans possess a similar biological gestural system, it could provide more evidence for the past existence of a gestural protolanguage. The connection to nonhuman primates through gesture has yet to be explored fully, but:

> …the fact that all living nonhuman great apes have an extensive, and extensively shared, repertoire of gestures that they use intentionally on a minute-by-minute basis, whereas intentional vocal communication in apes is restricted to the single situation of alarm in the chimpanzee, is strong evidence that the last common ancestor of *Homo* and *Pan* was also a largely gestural communicator. (Byrne & Cochet, 2017, p. 70)

The major criticism of a primarily gestural protolanguage is the problem regarding how it would have transitioned to a verbal language. "The specializations for speaking, in regard to both the production of speech and its reception, are complex and extensive, and would have required a long period of time to evolve" (Kendon, 2017, p. 165). As stated previously, there have not been any identified selection pressures that could have produced the change from an entirely gestural protolanguage to the vocal language we use today (Fitch, 2010). However, it is possible that the connection simply has yet to be found, and gesture may be a more serious option considering the implications of our relatives' gestural systems (Byrne & Cochet, 2017). As Fitch (2010) explains, "Evolution lacks foresight, any hypothesized intermediate form must have its own contemporary adaptive value for its bearers" (p. 484). This does not necessarily mean that human ancestors did not use gesture as a form of communication, but it does suggest that the use of a completely gestural language, which then transitioned to a spoken language, is unlikely (Fitch, 2010).

Although explaining the transition appears to be potentially troubling, supporters of this idea have offered possible explanations. For example, Arbib, Liebal, and Pika (2008) suggest that cognitive changes allowed humans to develop a system of "*complex imitation* of manual praxic actions", as opposed to simple imitation, through gesture (p. 1061). They use Broca's area as evidence for this possible connection. It is explained that, "…there is a human *mirror system for grasping*-i.e., a brain region activated for both grasping and observation of grasping-in or near Broca's area" (Arbib et al., 2008, p. 1053). Therefore, the possible hypothesis is based on the idea that language developed in a manner that began with gesture transitioning from simple to complex imitation and later developed into vocalizations. The proposed transition is:

> …(S5a) pantomime of grasping and manual praxic actions; (S5b) pantomime of actions outside the pantomimic's own behavioral repertoire (e.g., flapping the arms to mime a flying bird); (S5c) *protosign*, conventional gesutres used to formalize, disambiguate, and extend pantomime (e.g., to distinguish 'bird' from 'flying') that establish an ability for the free creation of arbitrary gestures to support an open-ended semantics; and, finally, (S6) *protolanguage*, a multimodal system created once protosign, after having achieved some critical mass, and *protospeech* (the capacity to use vocal communicative gestures)… (Arbib et al., 2008, p. 1061)

In summary, modern-day spoken language beginning with a gestural protolanguage seems plausible when first considered. The connection of human communication and nonhuman primate communication is obvious, and it is reasonable to believe if the vocal apparatus was not optimal for speech and language that the body

would be the next option. However, when studied more closely, there is a major

difference in the way that humans and apes gesture. It is also difficult to adopt when

supporters of this theory find it difficult to explain the transition to verbal

communication, which is a vital component of any protolanguage. If explanations such as

the one that Arbib and colleagues propose were to gain more concrete evidence, a

gestural protolanguage may prove to be a viable starting place for the human

communication system.

The next protolanguage hypothesis, the musical protolanguage, also does not use

linguistic form. Instead, it uses prosodic changes, similar to bird song to communicate.

Prosody is the rhythmic change in pitch within a vocalization. This protolanguage is

suggested to have developed in various stages, eventually ending in a more linguistically

based stage. Supporters find evidence for a musical protolanguage in various topics of

research including other species, the way we speak to infants, and how music is

processed in the brain.

The idea surrounding musical protolanguage is a paradox, "…that the generative

aspect of phonology might have emerged before it was put to any meaningful use" (Fitch,

2010, p. 471). This would suggest that phonology is considered an exaptation for this

type of communication. It was proposed that meaningful lexical language evolved in

three steps. First, a greater cognitive power was developed by means of selective

pressures influencing social and technological advancement such as increase in group

size and tool use. Second, vocal imitation through a means similar to bird song must have

influenced survival among our ancestors and more specifically sexual selection. Finally,

true language would develop out of this musical protolanguage (Fitch, 2010). The

transition to verbal language is thought to have been supported by vocal communication

and meaning coming together. The general idea is that, "…once proto-humans had the

capacity to imitate vocally, and to combine such signals with meanings, any of the much-

debated sources of meaningful words would be available…" (Fitch, 2010, p. 472).

When first developing this theory, there was not access to the abundance of

information that we have today. However, much of the data that has been gathered since

has helped to support some of the original ideas. For example, Darwin discussed birds

and bird song as a similar mechanism, but since then it has been discovered that whales,

seals, and hummingbirds all "…evolved vocal learning in the context of 'song'…" (Fitch,

2010, p. 477). As discussed previously, this vocal learning ability would have supported

the creation of novel signals.

Modern researchers describe musical protolanguage as being developed by our

ancestors in four steps. The first step invoked phonology coming before anything else.

"This system of 'bare phonology' provided a learned, complex, generative vocal

communication system, with multiple units being combined into a hierarchial, but

propositionally meaningless signaling system" (Fitch, 2010, p. 503). The second step

involved the introduction of meaning. It is during this step that whole phrases or songs

were linked with whole ideas or "semantic complexes" (Fitch, 2010). This is where the

protolanguage could have been used to influence others. In the third step the whole

phrases were broken down into smaller units of meaning that could be combined into

proto-words. The fourth and final stage led language to a more contemporary state with

the use of words and morphemes (Fitch, 2010). It also is when, "…pressure for rapid

analytic learning by children became strong" (Fitch, 2010, p. 504). This suggests that this stage was mostly influenced by information sharing between relatives (Fitch, 2010).

Further research has indicated, "…that in different animal species the ability to vary prosodic features in the voice, in conjunction with the ability to coordinate sound production with others-expressing emotions, and possibily triggering emotional reactions – has an adaptive value" (Filippi, 2016, p. 5). In nonhuman primates, prosody of calls is not voluntarily controlled. However, the prosodic cues are still received and responded to appropriately depending on the different features of the call (Filippi, 2016). Prosodic cues are used similarly in non-primate mammals where, "…the effective expression and perception of emotional arousal may allow individuals to respond appropriately, based on the degree of urgency or distress encoded in the call" (Filippi, 2016, p. 6). This supports the second step described above. The utilization of vocalizations in this way helps to explain the possible functionality of a musical protolanguage. Selective pressures such as, "…group cohesion, territory defense, pair bonding, and prenatal care…" could have made this type of communication advantageous for our hominid ancestors (Filippi, 2016, p. 6).

In addition to observing primate and non-primate mammals, researchers have studied human infants to test the hypothesis of a musical protolanguage. Common practices among caregivers of different cultures have raised questions regarding language development. For example, caregivers often use infant-directed speech (Filippi, 2016). Characteristics of infant-directed speech (IDS) are, "…shorter utterances, longer pauses, higher pitch, exaggerated intonational contours (Fernald and Simon, 1984; Fernald et al., 1989) and expanded vowel space (Kuhl, 1997; de Boer 2005)" (Filippi, 2016, p. 9). This

speech pattern is beneficial to the infant in terms of perception. It was found that,

"…infants tested in an operant auditory preference procedure showed a strong listening

preference for the frequency contours of IDS, but not for other associated patterns such as

amplitude or duration" (Fernald and Kuhl, 1987; Cooper et. al, 1997 as cited in Filippi,

2016, p. 9).

Infant-directed speech could be connected to the first use of vocalizations in

humans (Filippi, 2016). It has been hypothesized that the enlarging of the human brain

and the complications it caused for childbirth may have selected for infants with less

developed brains, allowing greater potential for development later on. This would have

lead to greater dependence of infants on their caregivers, making attention-seeking

behaviors, such as infant-directed speech, beneficial (Falk, 2004 as cited in Filippi,

2016). "Interestingly, this would explain why humans are the only species where tutors

exaggerate the prosodic features of the signal when speaking to immature offspring"

(Filippi, 2016, p. 9). Filippi (2016) connects this, "…use of prosody for emotional

communication and interactional coordination…" with the "…evolutionary emergence of

the first vocalizations in humans (p. 9)."

Evidence in support of the musical protolanguage hypothesis has been gathered

through musicological and brain imaging studies. Listeners, "…show strong biological

reactions to, and developmental readiness for, musical stimuli" (Fitch, 2010, p. 506). At

the same time, "…mechanisms specific to language (in particular all those associated

with propositional semantics) should be disjunct from those involved in music…" unless

an individual is assigning meaning to specific musical characteristics" (Fitch, 2010, p.

506).

This point is supported by a study that connects music and speech to the same area in the brain (Thompson, Marin, & Stewart, 2012). This study, "…examined sensitivity to emotion in speech prosody in a sample of individuals with congenital amusia, a neurodevelopmental disorder characterized by deficits in processing acoustic and structural attributes of music" (Thompson et al., 2012, p. 19027). The participants displayed difficulty judging emotional states including: happy, tender, afraid, irritated, sad, and no emotion. The participants also reported difficulty with comprehending vocal prosody in their everyday lives (Thompson et al., 2012). Thompson et al. (2012) explains:

> The results of this investigation confirm that amusic individuals are less accurate than matched controls at classifying emotions conveyed by speech prosody, supporting the hypothesis that music and speech are associated with shared resources for decoding acoustic signals of emotions. (p. 19029)

This is particularly interesting because amusic individuals do not consistently have any other difficulties in speech perception. This suggests a possible connection between music and speech, as well as emotional communication within the brain (Thompson et al., 2012).

This connection has been confirmed through brain imaging studies. It has been found that, "…the ability to process prosodic variations in language plays a vital role in the comprehension of both verbal and musical expressions" (Filippi, 2016, p. 10). This relationship has also been found in the study of whistled languages where both hemispheres are utilized, instead of primarily the left (Carreiras et al., 2005; Güntürkün et. al., 2015 as cited in Filippi, 2016). Research such as this suggests, "…that the abilities

underpinning linguistic prosody and music share cognitive and neural resources" (Filippi, 2016, p. 10).

In addition to the two major pre-lexical hypotheses of protolanguage, there is an additional popular idea surrounding the combination the musical and gestural protolanguages into a non-linguistic multimodal protolanguage. This idea is based on the fact that human and other ape communication is generally multimodal (Kita, Özyürek, Allen, & Ishizuka, 2010)" (de Boer, 2017, p. 160). The supporters of a hybrid protolanguage use the difficult transition from a fully gestural protolanguage to provide evidence for this idea. They think the idea of a multimodal protolanguage is more plausible since it does not require an explanation of the transition to a vocal system (Kendon, 2017). Kendon (2017) explains, "Evolution does not work in terms of possible but as-yet-unrealized advantages" (p. 166). Therefore, it seems more logical to adopt a protolanguage hypothesis that does not involve this transition.

The previous three protolanguages, gestural, musical, and multimodal are expected to have developed extremely early in the history of language. Another major theory, lexical protolanguage, is expected to have happened after the models that were previously discussed (Fitch, 2010). "A lexical protolanguage assumes, as prerequisites, an ability for vocal imitation (necessary to develop a shared spoken vocabulary) and a capacity and drive for referential communication" (Fitch, 2010, p. 401). This proposed protolanguage lacks syntactic structure, and the general idea has been incorporated into the work of many theorists in the field. Although differences exist in among theorists' ideas, "…they share a focus on reciprocal information exchange, among unrelated group-

living adults, as the crucial driving force underlying the evolution of propositional

communication" (Fitch, 2010, p. 424).

Much of the debate surrounding a lexical protolanguage is centered on how

syntax was developed. Many researchers agree on how a lexical protolanguage would

have presented itself, but disagree on how syntax could have been developed. One

popular idea is that language evolved from a representational system and that,

"…complex modern syntax is the final, and most crucial, step in the evolution of

language" (Fitch, 2010, p. 407). This idea about syntax is extreme, claiming that the

development of syntax was "catastrophic" (happened very suddenly) (Fitch, 2010). The

other popular view on how syntax could have developed is through a multi-step process.

The multi-step model includes stages both before and after the traditional stage of

lexical protolanguage that was originally suggested (Fitch, 2010). The first stage is the

ability to use symbols. Within this stage, the, "…voluntary use of symbolic vocalizations

(or other signals such as gestures)…" is a vital component (Jackendoff, 1999, p.p. 272).

Therefore, this stage could possibly be linked with previously discussed protolanguages.

Words with "…sudden high affect…" such as: "*ouch!, dammit!, wow!, and oboy!"* are

described as some of the earliest symbols of this stage (Jackendoff, 1999, p. 273). This is

not to suggest that human ancestors were using these specific markers, but that similar

phrases that like these have, "…no syntax and therefore cannot be integrated into larger

syntactic constructions…" were likely to have been used first (Jackendoff, 1999, p. 273).

The second stage involves the development of a large network of symbols. These

symbols would be developed into what we consider as a vocabulary. This increased

number of symbols would eventually demand humans to develop a way to store these

elements in long-term memory, which is supported by, "…the uniquely human ability to imitate…" (Donald, 1999 as cited in Jackendoff, 1999, p. 274).

The third stage is "proto-phonology" (p. 274). It was during this stage that the combination of syllables could be utilized to create novel utterances. Even with an extremely rudimentary set of syllables to employ, there could be a large diversification of utterances, creating the ability to communicate more effectively through more specific vocalizations (Jackendoff, 1999). The next step builds on this one, explaining the use of these syllabic combinations into more utterance-like form, still without the use of formal syntax (Jackendoff, 1999).

Another step toward syntax is described in the purposeful linear position of words to describe relationships. Without consistent position of words, it is possible for confusion and communication breakdown. For example, not knowing which word is performing the action, and which is being acted on, it could greatly change the meaning of the sentence (Jackendoff, 1999). This explains why, "…using the principles of word order would be communicatively adaptive" (Pinker & Bloom, 1990, as cited in Jackendoff, 1999, p. 275).

The next step builds on this idea, but instead of word order uses phrase structure. "The grouping of words into higher-order headed units is a crucial design feature of modern language. It allows principles of word order to be expanded to principles of phrase order" (Jackendoff, 1999, p. 277). This allows phrases to be made up of smaller phrases in addition to words (Jackendoff, 1999).

Following phrase structure, the next developments included relational vocabulary, inflection, and greater development in syntax. Relational vocabulary allows for more

expression of complex thought such as, "…to wonder if *p* and suppose that *p*, and to give reasons and purposes for actions, with a tremendous effect on the power of the individual and communal reason and planning" (Jackendoff, 1999, p. 277). Inflection and greater syntax development would further ensure the correct expression and perception of one's communication, making language an even more useful tool (Jackendoff, 1999).

Through these steps, the complete definition of a lexical protolanguage is reached, "… a lexical protolanguage with full vocal complexity and a large phonologically structured lexicon, used in a communicative context among adult members of some primitive hominid group" (Fitch, 2010, p. 412). The advances that these steps describe would make communication more clear and understandable, potentially increasing its adaptive value with each added level of complexity. Although this level of protolanguage appears more similar to current language than the gestural and musical protolanguages, there would still be a large number of cognitive and other modifications necessary to arrive at modern-day language.

The protolanguages described outline several possible methods as to how humans and their ancestors could have developed language. However, what they do not describe is possible selection pressures that could have caused these protolanguages to emerge. Currently there is a movement toward a social explanation of the development of speech and language capabilities in humans. The four main hypotheses explained below are the social brain hypothesis, the human self-domestication hypothesis, grooming as a form of social bonding, and the roles of imagination and deception in the evolutionary development of language.

The first hypothesis, the social brain hypothesis, helps to explain how increasing size of social groups could have served as a possible selective pressure for speech and language development. The social brain hypothesis states that:

> …across primates (including modern human) social group size correlates with brain size (Dunbar, 1992, 1998, 2001): since we know where we started (as a great ape) and where we ended up (as modern humans), it follows that hominin community size must track the changes in brain size in between (Gowlett, Gamble, & Dunbar, 2012; Dunbar, 2014a). (Dunbar, 2017, p. 209)

It is from this general background of increasing group size that different ideas of possible social selection pressures for language are based.

Hare (2017) proposes another socialization theory that possibly influenced the selection pressures surrounding the origin of communication systems, the human self-domestication hypothesis. "The human self-domestication hypothesis (HSD) draws on comparative, developmental, fossil, and neurobiological evidence to show that late human evolution was dominated by selection for intragroup prosociality over aggression" (Hare, 2017, p. 157). He uses other domesticated animals to support this hypothesis, explaining that humans demonstrate similar physiological and morphological changes (Hare, 2017).

It is suggested that the domestication of humans has changed how we interact with one another. For example, one of our closest ancestors, the chimpanzees, have been shown in studies to be able to either work together or communicate, but they have difficulty using these skills in conjunction with each other (Bullinger et al., 2014, Herrmann & Tomasello, 2006, Melis et al., 2009, as cited in Hare, 2017). "Only a small

minority of chimpanzee dyads within a group are tolerant enough to work for sharable

food and no dyad can cooperate once rewards require active sharing or turn taking (Hare

et al., 2007, Melis et al., 2006b)" (Hare, 2017, p. 160). This could suggest that working

together in humans is a result of our self-domestication.

It has been found that with the domestication of certain species, there has been,

"…increases in tolerance and social cognitive flexibility…" (Hare, 2017, p. 164). It is

suspected that this increase in tolerance has also happened in humans, possibly

influencing new cognitive developments (Hare, 2017). It is proposed that this increase in

tolerance was combined with an increase in inhibition in humans, creating, "…the

human-specific adaptation for more flexible tolerance and unique forms of human social

cognition" (Hare, 2017, p. 165).

These statements are supported by morphological data. Some examples of,

"…increased tolerance and cooperative-communicative abilities in *Homo sapiens*" are:

"Reduction in brow ridge and facial length", "2nd digit to 4th digit length ratio (2D4D)",

"White sclera", "Globular cranial development", "Extended synaptic pruning", and

"Absolute brain size increase" (Hare, 2017, p. 166). These morphological changes are not

necessarily related to communication, but indicated hormonal changes that are associated

with domestication and have been seen in other domesticated species (Hare, 2017).

These morphological changes allow for easier social development. For example,

the white sclera is what allows infants to use joint attention, which is, "…a central

psychological mechanism allowing for the development of unique human forms of

cooperative communication (Tomasello et al., 2005)" (Hare, 2017, p. 169). The white

scleras also suggest that face-to-face communication was becoming more of a common characteristic of socialization in humans.

The changes in the human brain indicate interesting possible social developments as well. "All of the brain regions leading to globular expansion are also involved in human social cognition, including the attribution of mental states to others" (Hare, 2017, p. 174). This could link the expansion of the brain with the development of a more complex social system (Hare, 2017).

Overall, the human self-domestication hypothesis does not directly suggest a connection to language, but it does demonstrate the great selective pressures that a changing social system could have had on our species. Morphological and physiological changes associated with prosociality had the possibility to change the way humans interact with each other, potentially driving major changes in the way we communicated. This hypothesis highlights how social theories are potentially valid explanations.

Another idea surrounding the social development of language is its connection to grooming. Dunbar (2017) explains that with the increase in group size, there would have also been a necessary increase in grooming among our ancestors to maintain social relationships. Eventually, this one-on-one relationship preservation would not have been able to be maintained with other responsibilities, which would have been detrimental to social systems. Therefore, there needed to be a more efficient system in place to address relationship maintenance (Dunbar, 2017).

The first possible transition away from grooming and towards language is laughter. Laughter releases the same endorphins as grooming would and is a social activity that includes more individuals (Dunbar, 2017). "The average size of naturally

occurring laughter groups is about three (Dezecache & Dunbar, 2012) and, since all three

people laugh, all get the endorphin 'hit'" (Dunbar, 2017, p. 210). Laughter is not thought

to have replaced social grooming practices altogether, but supplemented the practice to

dramatically decrease time spent socializing (Dunbar, 2017).

     The next transition in social grooming actually is connected to the idea of musical

protolanguage. As social groups continued to grow larger with increasing brain size,

laughter was also becoming an inefficient mechanism of socialization (Dunbar, 2017).

Dunbar (2017) connects this shift to the same time where many of the anatomical

changes supporting speech were happening, making singing or "musical chorusing" a

probable next step toward language (p. 210). "Importantly, singing triggers the same

endorphin mechanism as grooming and laughter, *and* at the same time increases the sense

of belonging or social bonding (Pearce et al., 2015, 2016; Weinstein et al., 2016)"

(Dunbar, 2017, p. 210). Chorusing also does not seem to have a limit on the number of

individuals it can affect, making it exponentially more efficient than laughing and social

grooming combined (Weinstein et al., 2016 as cited in Dunbar, 2017).

     It is thought during this time period of chorusing that the use and control of fire

began. Dunbar (2017) considers this to be the final step in the social development of

language. It would have been difficult to use laughter or grooming at such a large scale

during the nighttime hours. Dunbar (2017) explains:

     However, if wordless chorusing began to be used to allow communal chorusing

     on a conversational or even a camp-wide scale, it would have provided a natural

     template for the evolution of voiced speech, and hence language, by the very short

additional step of mapping meaning onto sound (as originally proposed by

Darwin 1871 and Jespersen 1922). (p. 211)

This new ability to use the nighttime hours allowed for additional time to socialize that

previously did not exist (Dunbar, 2017). "The night time *is* social time, and humans are

the only anthropoid primate capable of being active both day and night" (Dunbar, 2017,

p. 211). Through research involving hunter-gatherer groups today, it has been found that

mostly factual conversations happen during the day, while story-telling and jokes make

up majority of nighttime conversations, increasing the evidence for the probability of

fire's selective pressure on the evolution of language (Dunbar, 2017).

There is another hypothesis addressing the social development of language that is

referred to as the social complexity hypothesis. The idea is that the more complicated the

social system, the more complicated the communication needs to be in order to be

effective (Freeberg, Dunbar, & Ord, 2012 as cited in Kendon, 2017). Kendon (2017)

explains:

That is, there will be a wider and more diverse repertoire of communicative

signals, both vocal and visible, in societies that are larger, have differentiated

social roles and more complex interaction networks, and include maintained pair

relationships (as between mates, but also long-term friend relationships), than in

societies that have fewer of these features. (p. 167)

This hypothesis is linked to another idea about the connection of social

advancements to the evolution of language involving social cognition and the knowledge

of social hierarchy in groups of nonhuman primates. Nonhuman primates, like humans,

have complex social systems (Seyfarth & Cheney, 2017). "Long-term field studies

demonstrate that an important predictor of a baboon's or a chimpanzee's reproductive

success is an individual's ability to form close, long-term bonds" (Seyfarth & Cheney,

2017, p. 82).

Seemingly simple baboon vocalizations have several characteristics such as:

analyzing the vocalizing individual's intention to communicate, facilitating

communicative social exchanges, and the assessment of the meaning of the vocalization.

"A baboon's assessment of call meaning thus constitutes a *discrete, combinatorial, rule-*

governed, and open-ended system of communication (Cheney & Seyfarth, 1998;

Worden, 1998) in which the assessment of meaning depends upon the call type and

pragmatic interference based on social knowledge" (Seyfarth & Cheney, 2017, p. 82).

This does not mean that human language and nonhuman primate communication systems

are necessarily that closely related, but it does suggest, "…that several of the cognitive

mechanisms that have long been thought to mark a clear separation between language

and nonhuman primate communication can, in fact, be found-in admittedly simpler form-

in the communication and social cognition of nonhuman primates" (Seyfarth & Cheney,

2017, p. 82).

Although somewhat different, this hypothesis is similar to the grooming

hypothesis in that it relies on strong social bonds and relationships. These social bonds,

whether formed through social cognition or social grooming, are thought to increase the

likelihood of reproductive fitness. Therefore, greater communicative ability would be an

advantageous adaptation for our hominid ancestors.

Dor (2017) connects imagination and deception in social groups to the evolution

of language. He explains language as a tool for coordinated action, "…where the survival

of both speaker and interlocutor depended on the speaker's ability to provide the

interlocutor with the information that allowed for the appropriate action" (Dor, 2017, p.

46). With this sharing of information, the experiences of one person can be explained to

another, building "…imagination-based bridges over the experiential gaps between

communicators" (Dor, 2017, p. 47). In this case, the major selective pressure driving the

development of language would be an increase of dependency within social groups (Dor,

2017).

Within these new communities, new trends appeared in societal tasks. These

trends included, "…the control of fire and invention of cooking (Wrangham, 2009);

collaborative foraging and tool manufacturing industry it required (Sterelny, 2012;

Tomasello et al., 2012); the emergence of alloparenting (Hrdy, 2009)…" any many more

(Dor, 2017, p. 48). The introduction of these tasks and the interactions they created

developed social bonds and social dependency that earlier in our ancestry did not exist

(Dor, 2017).

The selection pressure of dependency on one another in a community could have

potentially created the environment where language could have developed. Dor (2017)

explains:

> …the growing dependency on experiential mutual identification locked humanity
>
> in a vicious circle: the ever-growing dependency of the community members on
>
> mutual identification required a constant rise in the amount and quality of
>
> communication within the group; the rise in the level of communication, however,
>
> only contributed to the deepening of the dependency. (p. 48)

This cycle could potentially influence the development of protolanguages, the transition of protolanguage to more modern-like spoken language, and the progression to what spoken language has become today. However, it is suggested that these steps would not have been completed without the introduction of lying and deception (Dor, 2017).

It is thought that language originally had to be stabilized prior to the introduction of lying, but that lying could have increased trust and therefore language complexity as time went on. As social groups became larger, it became more difficult for individuals to be a part of everything that was going on in a group (Gowlet et al., 2012 as cited in Dor, 2017). "The dispersed nature of the community forced individuals to 'think about absent individuals and their potential influence' (p. 697), which required a new system for the communication of information about people in their absence…" (Dunbar, 2004 as cited in Dor, 2017, p. 54). This new system allowed for gossip to be introduced, possibly furthering the complexity of language.

Lying was likely to have increased the strength of social relationships (Dor, 2017). "Language would be much simpler had it evolved just for honest communication, and we would be much less imaginative, suspicious and inquisitive and emotionally-controlled" (Dor, 2017, p. 57). Completely honest communication, although seemingly harmless, would have created a society with less social grace and most likely additional hostility. Even in modern-day communication, it is often the people who give their honest opinion who end up in the most altercations on a day-to-day basis. Therefore, although it may seem counterintuitive that lying and deceit could have advanced humans as a species, as well as our system of spoken language, the support for this idea is quite convincing.

Overall, the multitude of factors that could have influenced spoken language to become what it is today is overwhelming. There are anatomical, physiological, cognitive, and social selective pressures that seem to have had an effect, but it is also clear that the general consensus is that no one exactly understands how this complex system came to be. The mere fact that we are unsure of whether anatomical changes happened because language was selecting for them or whether it was something else in the environment selecting for these changes shows how much more research is to be done in this field.

It is unclear whether the debate regarding the different types of protolanguages will ever truly be solved. There is evidence for all of them, including the combinatorial examples, creating seemingly endless possibilities of what early language possibly might have presented itself as. The complexity surrounding the arguments for each protolanguage show how many different communicative strategies could have been beneficial, effective, and probable for our ancestors.

The social selective pressures surrounding the evolution of spoken language are extremely interesting arguments. Language itself is a social tool, so the connection between increasing group size and resulting consequences driving more complex language skills is plausible. The different approaches to this argument, although similar, do suggest that further research may provide additional vital information. This could shed light on how changes in social environments can affect a species' communication system.

References

American Speech-Language-Hearing Association. (n.d.). What is speech? What is

language?. Retrieved from

https://www.asha.org/public/speech/development/language_speech/

Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture and the

evolution of human language. *Current Anthropology, 49(6),* 1053-1076.

Boothroyd, A. (1986). Speech of deaf people. In *The Gallaudet Encyclopedia of*

Deafness and Deaf People (p. 191-195). New York: McGraw Hill.

Byrne, R., & Cochet, H. (2017). Where have all the (ape) gestures gone?. *Psychonomic*

Bulletin & Review, 24(1). 68-71. https://doi.org/10.3758/s13423-016-

1071-0

de Boer, B. (2017). Evolution of speech and evolution of language. *Psychonomic*

Bulletin & Review, 24(1), 158-162. doi: 10.3758/s13423-016-1130-6.

Dor, D. (2017). The role of the lie in the evolution of human language. *Language*

Sciences, 63. 44-59. http://dx.doi.org/10.1016/j.langsci.2017.01.001

Du Brul, L. (1976). Biomechanics of speech sounds. *Origins and Evolution of*

Language and Speech, *280*, 631-642.

Dunbar, R. I. M. (2017). Group size, vocal grooming and the origins of language.

Psychonomic Bulletin & Review, 24(1), 209-212.

Filippi, P. (2016). Emotional and interactional prosody across animal

communication systems: A comparative approach to the emergence of

language. *Frontiers in Psychology, 7.* DOI:10.3389/fpsyg.2016.01393

Fischer, J. (2017). Primate vocal production and the riddle of language evolution.

Psychonomic Bulletin & Review, 24(1), 72-78*.*

https://doi.org/10.3758/s13423-016-1076-8

Fitch, W. T. (2010). *The evolution of language.* Retrieved from https://ebookcentral-

proquest-com.libdatabase.newpaltz.edu

Fitch, W. T. (2017). Preface to the special issue on the biology and evolution of

language. *Psychonomic Bulletin & Review, 24(1),* 1-2.

Friederici, A. (2017). Evolution of the neural language network. *Psychonomic*

Bulletin & Review, 24(1), 41-47*.* doi: 10.3758/s13423-016-1090-x.

Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Current*

Biology, *18(11)* R457-R460).

Hare, B. (2017). Survival of the friendliest: *Homo sapiens* evolved via selection for

prosociality. *Annual Review of Psychology, 68*. Retrieved from

http://www.annualreviews.org.libdatabase.newpaltz.edu/doi/full/10.1146/

annurev-psych-010416-044201#_i5

Hickok, G. (2017). A cortical circuit for voluntary laryngeal control: Implications for

the evolution language. *Psychonomic Bulletin & Review, 24(1),* 56-63 doi:

10.3758/s13423-016-1100-z.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity.

Trends in Cognitive Sciences, 3(7), 272-279.

Jaynes, J. (1976). The evolution of language in the late Pleistocene. *Origins and*

Evolution of Language and Speech, *280*, 312-325.

Kendon, A. (2017). Reflections on the "gesture-first" hypothesis of language origins.

Psychonomic Bulletin & Review, 24(1), 163-170.

Locke, J. L. (2017). Emancipation of the voice: Vocal complexity as a fitness

indicator. *Psychonomic Bulletin & Review*, *24(1)*, 232-237.

Seyfarth, R. M., & Cheney, D. L. (2017). Precursors to language: Social cognition and

pragmatic interference in primates. *Psychonomic Bulletin & Review, 24(1),* 79-

84.

Steklis, H., & Harnad, S. (1976). From hand to mouth: Some critical stages in the

evolution of language. *Origins and Evolution of Language and Speech*, *280*,

445-455.

Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to

emotional prosody in congenital amusia rekindles the musical protolanguage

hypothesis. *Psychological and Cognitive Sciences, 109(46),* 19027-19032.

https://doi.org/10.1073/pnas.1210344109

Understanding Evolution. (n.d.). The neutral theory. Retrieved from

https://evolution.berkeley.edu/evolibrary/article/misconcep_08

Wind, J. (1976). Phylogeny of the human vocal tract. *Origins and Evolution of

Language and Speech*, *280*, 612-630.

Zywiczynski, P., Gontier, N., & Wacewicz, S. (2017). The evolution of (proto-

)language: Focus on mechanisms. *Language Sciences, 63,* 1-11.

https://doi.org/10.1016/j.langsci.2017.06.004