# Chapter 11
# Cosmolinguistics: Necessary Components for the Emergence of a Language-Like Communication System in a Habitable Planet

**Kazuo Okanoya**

**Abstract** The emergence of human language is one of the biggest wonders in the universe. In this chapter, I define "a language-like communication system" and examine the components necessary for the emergence of such a system, not only on Earth but in any habitable planet. Human language is a unique system among animal communication. Language is a system of transmitting an infinite variety of meanings by combining a finite number of tokens based on a set of rules. Language is not only used in communication but also in thinking. Thus, language is a system that enables compositional semantics. I propose that at least three components are necessary for the emergence of a language-like system: segmentation of context and behavior, the association between them, and the honesty of the emitted signals. When a signal conveys sufficient information regarding the behavioral state of the sender, that signal is defined as "honest," meaning that its production incurs physiological, temporal, and social costs. I explain each of these conditions and discuss the possibility of "language as it could be" on other planets. I also extend my argument to the future of linguistic communication.

**Keywords** Segmentation · Association · Signal honesty · Ritualization · Communication · Extraterrestrials

## 11.1 Introduction

Communication with extraterrestrial existence is a favorite topic of science fiction novels. The novel *Solaris*, written by Stanislaw Lem in 1961, describes the entire planet Solaris as an intelligent being. The planet Solaris tries to study the very human researchers who are, in turn, trying to study it (Lem 1961). Solaris seems to

K. Okanoya (✉)
The University of Tokyo, Tokyo, Japan
e-mail: cokanoya@mail.ecc.u-tokyo.ac.jp

be sentient and reactive to human investigation; however, the attempt to "communicate" with the planet fails. This is partially due to Solaris's reflective nature: the aim of communication for Solaris is to mimic the inside of human mind, while for humans, communication is believed to be mutually beneficial. More recently, Ted Chiang wrote the novel *Story of Your Life*, in which heptapod extraterrestrials visit and try to communicate with humans in their specific "language" (Chiang 1998). Their language has a holistic nature that allows it to transcend the time dimension. By studying their language, the linguist acquires a unique perception of time; he transcends the time dimension and sees the past and future simultaneously. These novels present possible structures of extraterrestrial language.

Here, I define language as a system of transmitting an infinite variety of meanings by combining a finite number of tokens based on a set of rules. Each token, in turn, has multiple associations with specific meanings. Language is not only used in communication but also in thinking. In fact, in some schools of theoretical linguistics, language is considered to have originated as a tool for thought (Chomsky 2000). Thus, language is a system that enables not only communication but also compositional semantics. How could such a system have evolved on planet Earth, and what conditions would be necessary for such a system to evolve outside of Earth? By asking these questions, I aim to start a new branch of linguistics, namely, "Cosmolinguistics."

## 11.2    The Emergence of Language-Like Communicative Signals on Earth

Communication in the context of biology is defined as "the transmission of a signal from one animal to another such that the sender benefits, on average, from the response of the recipient" (Slater 1983). Since this definition does not include an intention on the part of the signaler or a benefit to the receiver, it is useful to avoid anthropomorphic interpretations of animal behavior. Anthropomorphic views include the false notion that communication is mutually beneficial and communication is an indication of self-awareness. Communication can, in fact, evolve without self-awareness and mutual benefit (Bradbury and Vehrencamp 2012). In this section, I briefly propose a set of hypotheses to account for the emergence of language on planet Earth. Here, I limit myself to the discussion of acoustic communication in vertebrates, because the principal medium for language remains speech communication. I am aware that this is a specific condition on Earth in which most animals require respiration for metabolism.

### 11.2.1    Ritualization of Respiratory Movements

Communicative acoustic signals have always started as a secondary trait in vertebrate animal behavior (Fitch and Hauser 2003). Acoustic signals often originate from respiratory actions because respiratory organs function as air passages.

Because respiration is an action that is absolutely vital for animal survival, the secondary use of respiratory energy for vocal production has a low physiological cost (Oberweger and Goller 2001). The respiratory tract is a pipe connecting the bilateral lungs and mouth opening. Because the respiratory tract extends into the body, physiological conditions affect its acoustical characteristics. Coughing is associated with infection and inflammation of the respiratory tract. Strong exhalation produces noise associated with the length of the respiratory tract (Morton 1977). Furthermore, because opening the mouth is preparatory behavior for biting or attacking in predatory animals, the exhalative noise associated with mouth opening could signal attack (Briefer 2012). In this way, respiratory noise is correlated with subsequent behavior by the signaler.

When such signals change the behavior of the receiver so that the change benefits the sender, the signals gain communicative value. For example, the pup isolation calls of rodents comprise short, repeating ultrasonic calls. This acoustic signal has a characteristic of easy localization because, due to the short wavelength of ultrasounds, there are many onset-offset cues with phase information available for the small rodent heads. Upon detecting the isolation call, the mother quickly approaches to retrieve the pup, who is the sender of the call (Ehret 2005). These calls must have originated from the respiratory noises arising from the short and shrunken tracheae of infant animals, whose body temperature has quickly fallen due to isolation from the mother. Calls must then have undergone natural selection for localizability. During this process, the noise that originated due to hypothermia must have become the isolation call (Fig. 11.1).

## 11.2.2   Emergence of Songs

Most land animals emit "calls" specific to behavioral contexts. Calls are monosyllabic, simple vocalizations. In addition to calls, some animals emit trains of various calls, and such vocalizations are often used in mating contexts. Because of the acoustic resemblance to human singing, these vocalizations are sometimes referred to as songs. It has remained an enigma how songs emerged in animals.

In rodents, when pups are out of the nest, they emit isolation calls that induce retrieval responses from the mother. When bird chicks are hungry, they emit food-begging calls to make their parents bring food to them. When human babies need physiological or social care, they emit baby cries. These care-inducing signals are always in the form of repeated calls. This is true in rodents, birds, and humans (Wright and Leonard 2007). While repeated signals may increase the chance of detection, they may also increase the risk of becoming habituated. Infant pygmy marmosets produce repeated vocalizations when seeking care from adult animals, but they do so by combining different calls (Elowson et al. 1998). These call-repeating behaviors in young animals might be a preadaptation of songs in adults. Because these behaviors mimic infantile behavior, a tendency to produce randomly repeated calls may induce a strong reaction in female listeners.
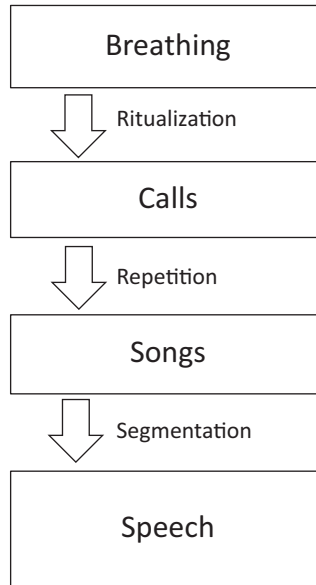
```
┌─────────────────────────────┐
│         Breathing           │
└─────────────────────────────┘
            ⇩  Ritualization
┌─────────────────────────────┐
│           Calls             │
└─────────────────────────────┘
            ⇩  Repetition
┌─────────────────────────────┐
│           Songs             │
└─────────────────────────────┘
            ⇩  Segmentation
┌─────────────────────────────┐
│          Speech             │
└─────────────────────────────┘
```

**Fig. 11.1** Schematic account of the set of hypotheses accounting for the emergence of language on planet Earth. Acoustic communication began as an expression of emotion associated with breathing. Such signals then became ritualized and the action patterns were fixed as calls. Repetitive calls were used by infant animals to intensify their signal value to mothers or parents. Similar signals were then mimicked by adult animals to relax female listeners in mating context. These signals comprise songs. Most animals sing innate songs, and receivers began to extract honest information about individual vigor from these songs. Songs then became sexually selected traits. In some species, complexity was preferred as a signal of vigor, and songs became a learned trait allowing further complexity. Such complex learned songs were shared in the societies of protohumans. The mutual segmentation of behavioral contexts and song phrases led to the emergence of speech

Supporting evidence for this infantile mimicry hypothesis comes from a neuro-anatomical study in the songbird brain (Liu et al. 2009). Chicks of chipping sparrows produce variable sequences of food-begging calls. When an expression of an immediate early gene (gene activated immediately after neural firing) was examined in the brain of these chicks, the area corresponding to the adult RA (robust nucleus of the arcopallium, homologues to the motor cortex in mammals) showed strong activation. Partial lesions of the same area resulted in a reduction in the variability of food-begging calls. The results indicate that food-begging and adult songs may utilize the same neural resources. This finding supports the hypothesis that food-begging calls may be a preadaptation to songs in birds.

Another line of evidence includes neurophysiological studies with mammalian isolation calls, including human cries. In rats and squirrel monkeys, lesioning the anterior cingulate cortex resulted in changes in the acoustic structures of isolation calls. In human babies, neural activity induced by crying was observed in the same brain area (Newman 2007). In adult mice, lesioning the anterior cingulate cortex resulted in changes in temporal and acoustical structures in courtship songs (Ariaga

unpublished observation). On the other hand, a mutant mouse that lacked neocortical and hippocampal areas sang normal songs, suggesting that only a part of the cortex may be necessary for courtship songs (Hammerschmidt et al. 2015).

Some species of bird, cetacean, and bat, in addition to one primate (only humans) demonstrate the additional faculty of vocal learning (Jarvis 2006). Vocal learning is the ability to acquire a new vocal repertoire through auditory-motor feedback learning. Vocal learning enables song complexity and eventually syllable variety in human speech. When and how vocal learning evolved is not known, but several hypotheses have been proposed, including mother-offspring interaction, sexual selection, domestication, and antipredator defense (Okanoya 2017).

Taken together, the idea that isolation calls and food-begging calls might be precursors to adult mating songs is consistent with the current data on neural mechanisms for vocal productions. Further studies are necessary to relate isolation and food-begging calls with adult mating songs in birds and mammals.

### 11.2.3   Emergence of Speech

Here, my challenge is to place the emergence of human speech in a continuous evolutionary line with the emergence of songs and the evolution of song complexity in nonhuman animals. To demonstrate the continuum of development with other primates, I will first examine song-like behavior in nonhuman primates and then propose a hypothesis related to the emergence of speech out of songs.

Gibbons are one of the five ape groups, of which the other four are humans, chimpanzees, gorillas, and orangutans. Because they are not great apes, gibbons are the most distant of the apes from humans. Gibbons do have song-like vocalizations (Geissmann 2002), but they are not learned, as indicated by cross-fostering studies. Cross-fostering studies involve exchanging babies of two species immediately after they were born and rearing the babies by the species different from their genetic species. Cross-fostering between two species of gibbons showed that gibbon songs are genetically determined and no effect of rearing environment was observed (Merker and Cox 1999). Nevertheless, gibbon songs are relatively diverse (Clarke et al. 2006) and are not only used in a mating context but also in many other social contexts (Inoue et al. 2012). In Muller gibbons, male calls consist of two simple types: a frequency-modulated "wa" call and a constant "o" call. Combinations of these calls and behavioral contexts have been correlated, meaning that gibbons might exchange contextual information via the combination of calls.

The gelada is a species of primate with a rich vocal repertoire. They also make a facial expression, with lip-smacking of 3–8 Hz used as an affiliative signal. On some occasions, their lip-smacking is presented with vocal sounds, making this behavior highly similar to human speech production (Bergman 2013). Other primates including macaques also show lip-smacking, and this behavior might be one of the precursors to human speech (Ghazanfar et al. 2012).

Both of these behaviors, vocal repertoire and lip-smacking, if combined with the bird-like ability of vocal learning, would provide a basis for the emergence of human speech.

## 11.3  Components for the Emergence of Language-Like Systems

I will now try to extend what might happen on Earth to habitable planets in general. Life evolved on planet Earth under highly specific conditions. Human language is the product of complex and arbitrary historical interactions that occurred only once on Earth. Nevertheless, it is possible to specify the boundary conditions that would lead to the emergence of language-like communication system on potentially habitable planets. I suggest that segmentation, association, and signal honesty are three key components necessary for such emergence.

Although vocal learning is considered a necessary condition for the emergence of human speech (Deacon 1998), I do not consider that condition at this point. This is because, in theory, language-like communication is possible if the agent has at least a binary (1 or 0) signal that is innately prepared. As is evident from digital computer architecture, a binary signal can emulate any degree of complexity. It is true that vocal learning and signal complexity can compress the time required to convey information, but these time constraints could vary depending on the agent's sensory and motor capacity.

## 11.4  Segmentation

Given a string of behavioral sequence, such as that for song on Earth, if there is a behavior that is sequentially or spatially emitted, it would provide the basis for segmentation and chunking. Segmentation is the process of cutting down longer or larger entities into shorter or smaller pieces. Chunking, on the other hand, is the process that does the opposite: amalgamating pieces to create a longer or larger entity. Segmentation occurs both in auditory and visual domains of the brain in vertebrate animals by means of lateral inhibition, in which the neurons that fired inhibit the activity of neighboring neurons (Meinhardt and Gierer 2000) or statistical learning, in which transition probabilities between two successive stimuli are learned (Saffran et al. 1999). The external environment would usually be more or less continuous, but living agents that move around must be able to segment or categorize the environment in order to reduce the load for sensory information processing. Segmentation is crucial in any agent that moves around a nonuniform environment.

When communication among similar organisms or conspecifics becomes beneficial, then the organisms output stimuli that are perceived by other organisms. These stimuli may be long and continuous on a physical domain but are packaged or chunked into pieces based on the motor constraints of their producing agent. The receiver may also segment the physically continuous stimuli into perceptual chunks based on the sensory constraints. The stimuli then become signals. In this way, the chunking of behavioral units and the segmenting of the perceived unit occur among the communicating agents.

Segmentation and chunking not only occur on stimuli but also in behavioral contexts. A behaving agent should know which behavioral situation is occurring in a given moment. This ability will also reduce the variety of its own behavioral state and make it easier to associate a stimulus with a behavior.

Consider what might have happened on Earth. How might song-like behavior in some primate species be connected with speech in humans? We proposed a conceptual model for this process (Merker and Okanoya 2007), in which each behavioral context is denoted by a particular song in a protohuman society. Consider the hypothesis that prior to language, protohumans developed singing behavior associated with several social contexts. If songs became a learned property, as they are in some species of bird and whale, a syllable phrase may be shared by more than one song. Then, likewise, parts of the behavioral contexts in which a song is sung may also be shared by more than one song (Fig. 11.2).
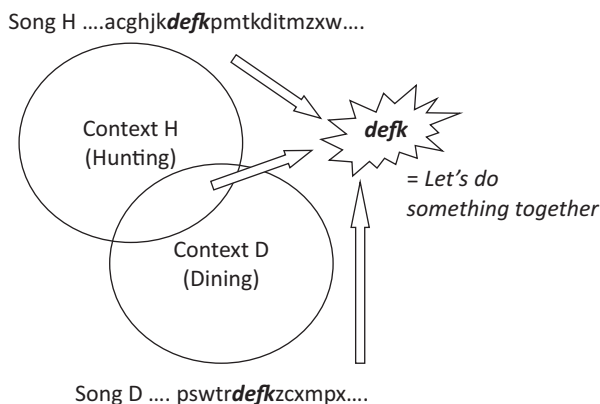
Song H ….acghjk**_defk_**pmtkditmzxw….

**_defk_**

Context H
(Hunting)

Context D
(Dining)

= _Let's do
something together_

Song D …. pswtr**_defk_**zcxmpx….

**Fig. 11.2** Mutual segmentation of song phrases and behavioral contexts. When two songs share a common phrase and context in which the songs are sung, the song (part) phrase and (part) context are mutually segmented and associated. The very short segmented song phrase then comes to denote the segmented specific context. In the specific example provided here, song H is sung when agents go hunting, and song D is sung when agents go dining. Because one of the contexts common to hunting and dining is "doing something together," the shared part of the two songs "_defk_" would likely be associated with the meaning of "doing something together" in the next generation of agents. In this way, holistic songs were gradually segmented into shorter pieces and associated with specific meanings through generations of agents

For example, a song sung when hunting (song H) and a song sung when dining (song D) might have shared the same phrase h&d. Furthermore, song H and song D shared the context of doing something together. After a while, by singing the shared phrase h&d, the singer could have specified the context of "let's do that together." By repeating this process, holistic songs might have been decomposed into specific phrases, which may in turn have become proto-words.

I call this the mutual segmentation hypothesis of song phrases and song contexts (Merker and Okanoya 2007; Okanoya and Merker 2007). Once the process of mutual segmentation commenced and segmented short utterances became associated with segmented restricted contexts, rudimentary forms of speech communication could have commenced. Subsequently, non-biological, cultural processes came to regulate the emergence of syntactical structures.

Segmentation and chunking in the signal and behavioral domain are thus essential for the evolution of language-like communication systems.

### 11.4.1   Association

In the previous section, I automatically assumed this faculty of associating stimuli and behavior. In all animals on Earth, associating given stimuli with given behaviors is an essential capacity for survival. This is shown to exist already in single-cell animals. The simplest form of such an association is habituation, in which repeated exposure to a given stimulus results in a reduction of behavioral response (Castellucci et al. 1970). A more complex form of association is known as Pavlovian conditioning, in which a neutral stimulus gains signal value through association with a key stimulus that can innately induce a certain response (Rescorla 1972). Operant conditioning is a further advanced form of associative learning in which the probability of occurrence of a defined behavior changes through external reward (Skinner 1990). For mutual segmentation of string and context to occur, I am assuming that at least associating a part of a string with a part of a context would be beneficial for the organism. Associative learning should be adaptive in any agents, as it affords the opportunity to predict what will occur next, as well as the selectivity to choose stimuli that result in positive reinforcement, and to avoid stimuli that result in punishment (Fig. 11.3).

### 11.4.2   Signal Honesty

For the receivers of the signal, it is crucial that the signal reflects the true behavioral state of the sender. If not, the signal loses its value and gradually ceases to function. Behavioral states include emotional, intentional, nutritional, and genetic (Brudzynski 2014; Searcy and Nowicki 2005). When a signal conveys sufficient information regarding the behavioral state of the sender, it is defined as "honest" (Searcy and
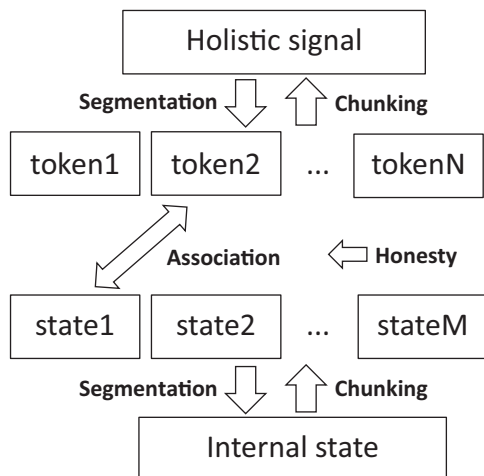
**Fig. 11.3** Necessary components to form a language-like communication system. Signals should have a hierarchy, which is achieved by chunking components (token1, token2, … tokenN) into a holistic signal or segmenting a holistic signal into tokens. Likewise, internal states should also have a hierarchy by chunking and segmentation. Internal states correspond to behavioral contexts interpreted by the agent. Behavioral tokens and internal states are associated via temporal proximity to form token-state pairs. Tokens must be honest signals if they are to guarantee the occurrence of a certain internal state. Thus, the production of each token is costly

Nowicki 2005). An honest signal bears "costs" of producing, such as physiological, temporal, and social costs. For example, birdsong incurs costs in terms of neural resources, metabolism, the risk of being located by a predator, and time costs (e.g., reduced time for foraging or other alternative behaviors). Thus, singing can be an honest signal to indicate the singer's resourcefulness and fitness. The above considerations on signal honesty should apply in any biological system that evolves not only on Earth but on any habitable planet.

## 11.5 Cosmolinguistics

I have discussed three components necessary to form a language-like communication system: segmentation, association, and signal honesty. Language is a system of transmitting an infinite variety of meanings by combining a finite number of tokens based on a set of rules. When language is defined as such, it enables the accumulation of knowledge. To form such a system, the segmentation of an external stimulus and internal state, their association, and maintaining signal honesty are considered necessary components.

I hypothesize that proto-speech emerged from the process of mutual segmentation of song string and behavioral contexts (Merker and Okanoya 2007; Okanoya and Merker 2007). Once speech had gained the combinatory property by which new

expressions became possible, the speech signal could now point to nonexisting or imaginary entities. This marked the beginning of imagination. By freely combining concepts that were not associated, humans came to develop their imagination and creative thinking. However, at the same time, this also marked the beginning of manipulative communication, because with language, anything could be expressed without grounding it in the traits of the speaker. This also made language a dishonest signal in the sense of signal honesty (Bradbury and Vehrencamp 2012). Nevertheless, humans continued to use language once it had been acquired evolutionarily. Why is this possible? This consideration would also give shape to extraterrestrial "language."

## 11.5.1 Language-Like Signals: Honest and Dishonest Components

One of the reasons why language, a dishonest signal, survived could be because language as expressed speech has multiple components. Speech comprises vocal behavior used in face-to-face contexts. This means that speech, in its original mode, is used in real time, in proximity, and together with visual information. Speech behavior includes emotional information such as prosody, facial expression, and bodily movement. This emotional information mostly consists of honest signals, because they cannot be manipulated intentionally (Zuckerman et al. 1979). At the same time, of course, speech content comprises editable information. In face-to-face communication, if the speech content intentionally contained false information, prosodic or facial emotion would convey that the content was untrue. Honesty of speech content was thus guaranteed by honesty of speech behavior (Fig. 11.4). In
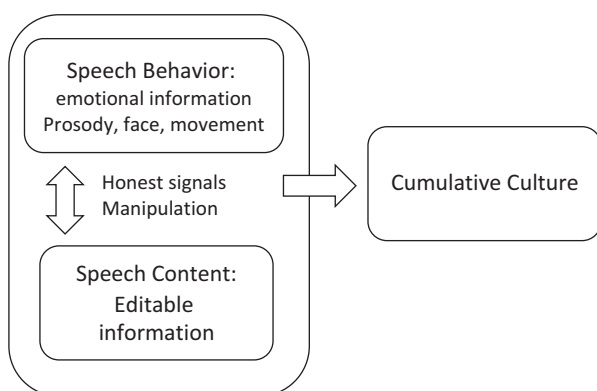


**Fig. 11.4** Since the emergence of language, communication content has been divided into linguistic content and emotional information. Since the invention of telecommunication, text content alone is often conveyed, but the accompanying emotional information is often discarded. This situation causes problems in maintaining linguistic communication

this way, human speech was utilized and evolved as a useful tool to accumulate knowledge.

If an extraterrestrial language exists, it should also contain multiple components, some of which should support the accumulation of information and the remainder of which should contribute to securing signal honesty.

### 11.5.2 The Future of Human Communication

The above scenario might account for the evolution of speech up to the invention of telecommunication in humans. Telecommunication first began with the invention of speech recording by means of non-acoustical, mostly visual notations. Sophisticated visual notations of speech ultimately led to the invention of letters. Because visual notations and letters continued to include emotional information and cost of production, the honesty of the contents was still not entirely violated (Lachmann et al. 2001) and the primary mode of communication continued to be face-to-face. As electrical devices for telecommunication advanced, however, the face-to-face mode of speech communication began to lose its position as the primary mode of communication. In modern society, a great deal of work is conducted through telecommunication devices in which most of the information is text based. We examined how emotional content could be transmitted in telecommunication devices and found that the sense of emotion transmission is very low in text-based communication (Arimoto and Okanoya 2015).

Although text-based communication is efficient in terms of the time, cost, and accuracy of both parties, it lacks the signal honesty necessary for fruitful communication. Additionally, since devices develop much more quickly than a single generation of humans, different generations are imprinted with different means of information transfer (Kelly 2016). Most current social problems are rooted in these simple facts. Now is the time to consider how we should design future means of communication.

### 11.5.3 Can Solaris Exist?

To revert to the introduction, I described two novels whose theme is language in nonhuman extraterrestrial intelligence. I think Solaris could not exist because it is a single organism that does not require communication and competition with other similar organisms. The planet Solaris does not require segmentation, association, or signal honesty. Thus, it does not require a system of communication, and no self-awareness would evolve in such a planet. Likewise, the heptapod in *Story of Your Life* could not obtain the time-transcending linguistic system. This is because language is based on the token-state association, and this association depends on the temporal co-occurrence of events (Rescorla 1972). Signal honesty is not supported

in heptapod communication because honesty is not judged in time-transcending situations. Of course, I am by no means criticizing these novels; in fact, I love them for the very reason that they trigger my imagination on important questions of what it is to be human.

## 11.6   Conclusion

In this chapter, I reviewed the literature on the evolution of acoustic communication in animals. I developed a set of hypotheses to account for the emergence of human speech and language in line with the evolution of animal communication. I found that a discontinuity occurred when humans began to use devices for telecommunication, since these remove the emotional information that supports the honesty of linguistic content. I considered that this might change the way humans use language. When this is extended to the language-like communication system of a hypothesized extraterrestrial one, I can suggest at least that the system should contain multiple components to support the contradictory needs of information accumulation and signal honesty. I can also suggest that such a system would need to function on the axis of time to enable the effective association of tokens and states.

## References

Arimoto Y, Okanoya K (2015) Mutual emotional understanding in a face-to-face communication environment: how speakers understand and react to listeners' emotion in a game task dialog. Acoust Sci Technol 36(4):370–373

Bergman TJ (2013) Speech-like vocalized lip-smacking in geladas. Curr Biol 23(7):R268–R269

Bradbury JW, Vehrencamp SL (2012) Principles of animal communication. Sinauer Associates, Sunderland

Briefer E (2012) Vocal expression of emotions in mammals: mechanisms of production and evidence. J Zool 288(1):1–20

Brudzynski SM (2014) Social origin of vocal communication in rodents. In: Witzany G (ed) Biocommunication of animals. Springer, Dortrecht, pp 63–79

Castellucci V, Pinsker H, Kupfermann I, Kandel ER (1970) Neuronal mechanisms of habituation and dishabituation of the gill-withdrawal reflex in Aplysia. Science 167(3926):1745–1748

Chiang T (1998) Story of your life. Stories of your life and others, 117–178

Chomsky N (2000) New horizons in the study of language and mind. Cambridge University Press, Cambridge MA

Clarke E, Reichard UH, Zuberbühler K (2006) The syntax and meaning of wild gibbon songs. PLoS One 1(1):e73

Deacon TW (1998) The symbolic species: the co-evolution of language and the brain. WW Norton & Company, New York

Ehret G (2005) Infant rodent ultrasounds–a gate to the understanding of sound communication. Behav Genet 35(1):19–29

Elowson AM, Snowdon CT, Lazaro-Perea C (1998) Infant 'babbling' in a nonhuman primate: complex vocal sequences with repeated call types. Behaviour 135(5):643–664

Fitch WT, Hauser MD (2003) Unpacking "honesty": vertebrate vocal production and the evolution of acoustic signals. In: Simmons AM, Popper AN, Fay RR (eds) Acoustic communication. Springer, New York, pp 65–137

Geissmann T (2002) Duet-splitting and the evolution of gibbon songs. Biol Rev 77(1):57–76. https://doi.org/10.1017/s1464793101005826

Ghazanfar AA, Takahashi DY, Mathur N, Fitch WT (2012) Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. Curr Biol 22(13):1176–1182

Hammerschmidt K, Whelan G, Eichele G, Fischer J (2015) Mice lacking the cerebral cortex develop normal song: insights into the foundations of vocal learning. Sci Rep 5:8808

Inoue Y, Sinun W, Yoshida S, Okanoya K (2012) Male gibbons change the note order according to the behavioral situations. In: The Evolution of Language: Proceedings of the 9th International Conference (EVOLANG9), World Scientific, Kyoto, Japan, p. 458, 13–16 March 2012

Jarvis E (2006) Selection for and against vocal learning in birds and mammals. Ornithol Sci 5(1):5–14

Kelly K (2016) The inevitable: understanding the 12 technological forces that will shape our future

Lachmann M, Szamado S, Bergstrom CT (2001) Cost and conflict in animal signals and human language. Proc Natl Acad Sci U S A 98(23):13189–13194

Lem S (1961) Solaris, trans. Joanna Kilmartin and Steve Cox. Solaris, the chain of chance, a perfect vacuum (London: Penguin, 1981):32

Liu W-C, Wada K, Nottebohm F (2009) Variable food begging calls are harbingers of vocal learning. PLoS One 4(6):e5929

Meinhardt H, Gierer A (2000) Pattern formation by local self-activation and lateral inhibition. BioEssays 22(8):753–760

Merker B, Cox C (1999) Development of the female great call in *Hylobates gabriellae*: a case study. Folia Primatol 70(2):97–106

Merker B, Okanoya K (2007) The natural history of human language: Bridging the gaps without magic, Emergence of Communication and Language. Springer-Verlag, London, pp 403–420

Morton ES (1977) On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. Am Nat 111(981):855–869

Newman JD (2007) Neural circuits underlying crying and cry responding in mammals. Behav Brain Res 182(2):155–165

Oberweger K, Goller F (2001) The metabolic cost of birdsong production. J Exp Biol 204(19):3379–3388

Okanoya K (2017) Sexual communication and domestication may give rise to the signal complexity necessary for the emergence of language: an indication from songbird studies. Psychon Bull Rev 24(1):106–110

Okanoya K, Merker B (2007) Neural substrates for string-context mutual segmentation: a path to human language, Emergence of Communication and Language. Springer-Verlag, London, pp 421–434

Rescorla RA (1972) A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement, Classical conditioning II: Current research and theory. Appleton-Century-Crofts, New York

Saffran JR, Johnson EK, Aslin RN, Newport EL (1999) Statistical learning of tone sequences by human infants and adults. Cognition 70(1):27–52

Searcy WA, Nowicki S (2005) The evolution of animal communication: reliability and deception in signaling systems. Princeton University Press, Princeton

Skinner BF (1990) The behavior of organisms: an experimental analysis. BF Skinner Foundation, New York

Slater PJB (1983) The study of communication. In: Halliday TRS, Slater PJB (eds) Communication, vol 2. Blackwell Scientific Publications, New York, pp 9–42

Wright J, Leonard ML (2007) The evolution of begging: competition, cooperation and communication. Springer Science & Business Media, Berlin

Zuckerman M, DeFrank RS, Hall JA, Larrance DT, Rosenthal R (1979) Facial and vocal cues of deception and honesty. J Exp Soc Psychol 15(4):378–339