R

Ritual and Speech Coevolution

Camilla Power
Department of Anthropology
University of East London, London, UK

Synonyms

Conventionalization; Metaphor; Mimesis; Symbolic culture

Definition

Thesis originally developed by anthropologist Chris Knight on the important consideration of novel forms of group cooperation, and corresponding conflict, in the emergence of human language, analogous to signaling in other species.

Introduction

In this entry, I discuss the ritual/speech coevolution hypothesis for the evolutionary emergence of spoken language. This brings the main theory on the evolution of animal signals to bear on the question of language origins. Our closest primate relatives are fundamentally constrained in their ability to mimic or produce novel vocalizations,

under pressure to maintain reliability in their signals. The ritual/speech coevolution theory explains how those selection pressures were relaxed in the case of human evolution through the generation of novel levels of in-group trust.

The Problem of Cooperation in Communication

Ritual/speech coevolution is a hypothesis explaining the evolutionary emergence of spoken language. It rests on the assumption that language can only evolve given novel conditions of social cooperation widespread across human communities – effectively cooperation between strangers. In the contexts of human evolution, the only medium for securing such cooperation within and between groups was costly ritual (cf Durkheim 1912; Knight et al. 1995; Maynard Smith and Szathmary 1995; Deacon 1997; Rappaport 1999; Irons 2001; Alcorta and Sosis 2005; Knight and Lewis 2017).

Ritual/speech coevolution brings to the problem of language origins the main body of theory applied to the evolution of animal signals. Signal evolution theory deals with the emergence, function, and design of animal signals (Maynard Smith and Harper 2003). Focused on social interaction and behavior in the world, signal evolution is necessarily political in its approach. Central debates concern the honesty and reliability of signals (Zahavi and Zahavi 1997), manipulation

and mindreading (Krebs and Dawkins 1984), and the effects of shared versus conflicting interests in outcomes (Maynard Smith and Harper 2003).

Among primates generally, tactical use of deception is predicted by neocortex volume (Byrne and Corp 2004), a finding consistent with the Machiavellian intelligence hypothesis – the idea that intelligence evolves under pressure to deal with social challenges. Humans are clearly Machiavellian in terms of political alliance formation and capacities for deception. Issues of trust and reliability are critical to any understanding of language origins.

While apes and monkeys have rich repertoires of calls, they seem to be severely restricted in their ability to mimic or to produce novel vocalizations (Fitch and Zuberbühler 2013). In particular, when phonating, they lack neural control over articulators such as the lips, tongue, soft palate, etc. This means that our closest relatives are unable to manipulate vocalizations at will. No monkey or ape has an inflexible tongue, but when communicating it leaves the tongue out. Knight and Lewis (2017) draw on signal evolution theory to explain why. While sound – unlike visible gesture – carries over distances, goes around corners, and works in the dark, in these contexts listeners lack corroborating evidence of reliability. It makes sense then that vocalizations tied to bodily and emotional states will be perceived as more reliable than those which can be altered at will. Mistrusting one another's scheming, Machiavellian minds, primates ignore the all too-flexible tongue, preferring to rely on the evidence of their own eyes and ears. In this view, apes are "too clever for words" (Knight 1998: 72).

So the conundrum of speech and language evolution is to explain how and why natural selection, in the human case, switched from quarantining the primate tongue – excluding it from all but a marginal communicative role – to developing and fine-tuning that same tongue's role as the most important speech articulator of all.

Signal evolution theory contrasts two divergent trajectories depending on the degree to which signaler and receiver share interest in the same outcomes. Where we find high-cost, repetitive, multimedia displays, we may infer a function

in terms of social manipulation, conflict, and exploitation. Resistance by receivers sets up selection pressures acting on signal design. Signalers who encounter "sales resistance" are driven to respond by prolonging and repeating signals, increasing amplitude, and resorting to costly multimedia displays. Such extravagant advertisers include peacocks displaying to would-be mates or caribou bulls bellowing in the rut. Zahavi (1987) shows how the discernible costs of such displays enhance their credibility by tapping and hence testing the very reservoirs of quality that signalers are attempting to advertise. The evolution of high-cost signaling is driven by skeptical receivers pushing signalers to ever greater competitive effort to prove their quality.

By contrast, where we find low-cost, quiet, and efficient signals, a cooperative audience can be inferred. If signalers can afford to cut their emission costs, it is only because listeners are investing corresponding effort in receiving, decoding, and acting on signals. This implies that signalers and receivers significantly share interests. For such "conspiratorial whispering" (Krebs and Dawkins 1984) to evolve, signalers must be imparting useful information to receivers and those receivers are not expecting to be deceived.

The ultimate cost-cutting strategy would be to resort to purely tokenistic, wholly conventional signals that can be processed categorically at speed – relieving listeners of the need to evaluate gradations in physical performance and signalers of such costly performance. According to Zahavi (1993), however, animal "conspiracies" are never sufficiently cooperative. Internal conflict and skepticism precludes selection for reliance on tokenistic "paper money." Nowhere in the living world do we find purely conventional signaling – with the puzzling exception of human speech.

Turning to human symbolic communication, Knight (1998, 1999) argued that our ancestors developed divergent, formally contrastive types of communication along these two "high-cost" and "low-cost" trajectories: ritual and speech. Table 1 shows the diametric opposition of characteristics of each. On the basis of signal evolution theory, we can infer that speech emerged in a cooperative context while ritual did

Ritual and Speech Coevolution, Table 1 Speech versus ritual (After Knight 1999: 231)

Speech	Ritual
Cheap signals	Costly signals
Interpersonal	Group-on-group
Two-way communication	One-way signals
Low amplitude	High amplitude
Dispassionate	Emotive
Vocal-auditory	Multi-media
Digital	Analog
Discrete-combinatorial	Holistic
Productivity/creativity	Repetition/redundancy
Stress on novelty	Stress on conservatism
Conventionally coded	Iconic and indexical
Focus on underlying intentions	Focus on body-boundaries and surfaces

not. For speech to have evolved, "conspiratorial whispering" in the human case must have been anomalously trusting. By contrast, ritual – with its costly, inefficient features of redundancy and display – can only have emerged from a dynamic of conflict, manipulation, and exploitation.

The paradox here, in the light of nonhuman primate vocalization, is that the extremely low-cost, conventional codes of speech can easily leave listeners vulnerable to deception. With no intrinsic link between sounds and their purely arbitrary meanings, words are routinely decoupled from emotional veracity or real-world stimuli. Despite speech having no intrinsic reliability, human conversation works on Grice's "cooperative principle" (Grice 1969), with participants, even where they may be in some degree of social conflict, cooperating at the level of mutual conversational ends. If humans are on "speaking terms," they expect intentional honesty. Somehow this new default of honesty – honesty in the deployment of volitional, conventional signals – became established. In the course of human evolution what were once frequency-dependent tactical deceptions became increasingly harnessed to a reversed social function: group-wide sharing of socially useful information.

This paradox can be unraveled if we consider that whereas speech is fundamentally

interpersonal, ritual operates group-on-group and may function to demarcate boundaries of those groups (Cohen 1985; Harrison 1993). Experience of the high-cost, multimedia signals involved in ritual will differ according to whether the receiver is outside or inside the performing group. For outsiders, the costly signals are manipulative and exploitative, needing to overcome "sales-resistance" by impressing observers with the quality of performance. Among insiders, the "collective effervescence" (Durkheim 1912), aroused by singing, dancing, and musicking together, will intensify in-group trust and generate a sense of group identity as "We" (Rappaport 1999). Focusing on the emotional responses of individuals to ritual experience, Alcorta and Sosis (2005) show that the very costliness of undergoing ritual enables those individuals to demonstrate in hard-to-fake terms their commitment to the group.

There are two major emergent effects from such proto-ritual experience. First, the novel levels of in-group trust can now permit conventionalization and shorthand reference – volitional, conventional signals – in communicating among group members, now on the way to becoming a speech community. Secondly, in representing their coherent and enduring solidarity, ritual performers generate sacred, supernatural, and counter-intuitive concepts – the first gods (Durkheim 1912; Knight et al. 1995; Rappaport 1999; Alcorta and Sosis 2005).

Roy Rappaport analyzes the form and features of ritual that make it "the social act basic to humanity" (Rappaport 1999: 31). Ritual overcomes the divisiveness of individually held beliefs. Defining ritual as "the performance of more or less invariant sequences of formal acts and utterances not entirely encoded by the performers" (Rappaport 1999: 24), Rappaport focuses on the unique conjunction of features in ritual. Ritual form permits the intersection of the symbolic with what is not purely symbolic. Ritual combines formal canonical messages – encoded by other than the performers – with the indexical signals, the current state, mood, and emotions of the flesh and blood performers. This "substantiation of form and ...informing of substance" mean that the establishment of convention is intrinsic to

ritual. Without this, no social contract, moral order, or language could exist.

The function of ritual is not to differentiate between lexical meanings but to establish, for everyone, an overarching meaning – a meta-performative or Word – from whose subsequent fragmentation a limitless multiplicity of subsidiary meanings can be derived. Rappaport explains how apparently irrational nonsense – perhaps the endless repetition of just a few meaningless sounds – may "provide the ground, deeper than logic and beyond logic's reach" upon which to establish sufficient collective authority and mutual trust to build up "the usages and rules of social life," in turn enabling words to make sense.

While Rappaport did not have any very satisfactory model for the evolution of ritual (hence speech), Knight and Lewis (2017), within the ritual/speech coevolution paradigm, provide a specific evolutionary account. Words and grammar, they claim (Knight and Lewis 2017: 435), are "means of navigating within a shared virtual world. Singing, dancing, and other forms of communal ritual are necessary to join people together in such ideal or imagined worlds." Language then will not even begin to evolve unless ritual action has already begun to establish intensified levels of community-wide trust in association with a shared virtual domain. This explains why nonhuman primates, confined as they are to the brute world of physical facts, have no need of it.

Knight and Lewis explain how the generative principle for both words and grammar is the ostensive-inferential process underlying metaphor. In this view, symbolic communication "rests on the ability of listeners to infer relevant communicative intentions from expressions that, interpreted literally, are inadequate or untrue" (Knight and Lewis 2017: 436). Not only does language comprise zero or very low-cost volitional vocal signals, but at the level of both words and rules, lexicon and grammar, these amount to conventionally agreed "lies" or deceptions. Whereas nonhuman primates would simply reject these sound sequences, we humans instead search out the hidden implications and communicative intentions.

Knight and Lewis draw on hunter-gatherer ethnography to work out a model for the emergence of such socially agreed "deceptions." Human abilities to manipulate pitch were first exercised in fooling outsiders: animals, whether these were prey being hunted (by men) or predators being evaded (by women and children). Song first – a communal polyphony of meaningless sounds – used to confuse and deter predators regarding numbers of hominins in groups, would have developed capacities for intentionally varying pitch. In such a context, the emotional solidarity among chorusing groups would override any dangers of deceptive use. Shorthand snatches of song, animal-sound mimicry, or pantomime dance steps could now begin to take on specific shared meaning. From there, a fully grammaticalized language could emerge with extraordinary speed, as long as the freedom to innovate – freedom to "say" one thing while "meaning" another – was maintained. What had blocked earlier hominins and hominids from any such grammaticalization process was "the burden imposed on all signals to incorporate some costly component to demonstrate reliability" (Knight and Lewis 2017: 445). Grammaticalization is based on a tendency to efficiency, but requirement of reliability in signals demands the opposite, increased cost. The release from this constraint frees the tongue and other articulators so they can be recruited to increasingly efficient, conventionalized communication: speech.

Knight and Lewis (2017) aim to make their model testable by specifying, in detail, the "world's first metaphor," identifying it as a gendered ritual performance in which the core principles of primate politics are overturned. They argue that as female hominin ancestors faced increasing reproductive costs with very large-brained offspring, critical points of conflict occurred whenever an alpha male attempted to monopolize an imminently fertile (menstruating) female. This would have conflicted sharply with reproductive interests of both other females and other males. Ritual display developed as female coalitions mounted resistance to any would-be alpha, with the whole female group signaling their non-availability to males through song and pantomime

dance: we are the wrong species (animals); we are the wrong sex (males); and this is the wrong time (blood). By identifying their own "menstrual" blood as the blood of the game animals they wanted the men to hunt, women's first morally authoritative ritual construct established categoric rules encompassing sex, kinship, and economics.

The specific model proposed here has the merit of testability. But the general ritual/speech coevolution paradigm has advantages (Power 2014). Firstly it offers a coherent basis for continuity with animal communication systems through theory of ritualized behavior. Secondly, regarding the archaeological record, it makes the clear prediction that the origins of language will be marked by evidence for ritual behavior (Watts 2014).

Cross-References

▶ Communication, Cues, and Signals
▶ Cooperative Coalitions
▶ Evolution of Cooperation
▶ Female Resistance
▶ Grammaticalization
▶ Handicap
▶ Manipulation and Dishonest Signals
▶ Reliability and Deception in Language
▶ Sexual Conflict Theory
▶ Signal Reliability.

References

Alcorta, C. S., & Sosis, R. (2005). Ritual, emotion and sacred symbols. *Human Nature, 16*, 323–359.

Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society of London B, 271*, 1693–1699.

Cohen, A. P. (1985). *The symbolic construction of community*. London: Tavistock.

Deacon, T. (1997). *The symbolic species: The co-evolution of language and the human brain*. London: Penguin.

Durkheim, E. (1912). *Les formes élémentaires de la vie religieuse*. Paris: Alcan.

Fitch, W. T., & Zuberbühler, K. (2013). Primate precursors to human language: Beyond discontinuity. In

E. Altenmüller, S. Schmidt, & E. Zimmermann (Eds.), *Evolution of emotional communication* (pp. 26–48). Oxford: Oxford University Press.

Grice, H. P. (1969). Utterer's meanings and intentions. *Philosophical Review, 78*, 147–177.

Harrison, S. (1993). The mask of war. In *Violence, ritual and the self in melanesia*. Manchester & New York: Manchester University Press.

Irons, W. G. (2001). Religion as hard-to-fake sign of commitment. In R. Nesse (Ed.), *Evolution and the capacity for commitment* (pp. 292–309). New York: Russell Sage Foundation.

Knight, C. (1998). Ritual/speech co-evolution: A solution to the problem of deception. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language. Social and cognitive bases* (pp. 68–91). Cambridge: Cambridge University Press.

Knight, C. (1999). Sex and language as pretend-play. In R. Dunbar, C. Knight, & C. Power (Eds.), *The evolution of culture* (pp. 228–247). Edinburgh: Edinburgh University Press.

Knight, C., & Lewis, J. (2017). Wild voices: Mimicry, reversal, metaphor, and the emergence of language. *Current Anthropology, 58*, 435–453.

Knight, C., Power, C., & Watts, I. (1995). The human symbolic revolution: A darwinian account. *Cambridge Archaeological Journal, 5*, 75–114.

Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (2nd ed., pp. 380–402). Oxford: Blackwell Scientific Publications.

Maynard Smith, J., & Harper, D. (2003). *Animal signals*. Oxford: Oxford University Press.

Maynard Smith, J., & Szathmary, E. (1995). *The major transitions in evolution*. Oxford: W. H. Freeman.

Power, C. (2014). The evolution of ritual as a process of sexual selection. In D. Dor, C. Knight, & J. Lewis (Eds.), *The social origins of language* (pp. 196–207). Oxford: Oxford University Press.

Rappaport, R. A. (1999). *Ritual and religion in the making of humanity*. Cambridge: Cambridge University Press.

Watts, I. (2014). The red thread: Pigment use and the evolution of collective ritual. In D. Dor, C. Knight, & J. Lewis (Eds.), *The social origins of language* (pp. 208–227). Oxford: Oxford University Press.

Zahavi, A. (1987). The theory of signal selection and some of its implications. In V. P. Delfino (Ed.), *International symposium of biological evolution* (pp. 305–327). Bari: Adriatic Editrice.

Zahavi, A. (1993). The fallacy of conventional signalling. *Philosophical Transactions of the Royal Society of London, 340*, 227–230.

Zahavi, A., & Zahavi, A. (1997). *The handicap principle*. New York and Oxford: Oxford University Press.