Trends in
Cognitive Sciences

CellPress
OPEN ACCESS

Feature Review

The 'design features' of language revisited

Michael Pleyer[1], Marcus Perlman[2,*], Gary Lupyan[3], Koen de Reus[4,5], and Limor Raviv[6,7]

Language is often regarded as a defining trait of our species, but what are its core properties? In 1960, Hockett published 'The origin of speech' enumerating 13 design features presumed to be common to all languages, and which, taken together, separate language from other communication systems. Here. we review which features still hold true in light of new evidence from cognitive science, linguistics, animal cognition, and anthropology, and demonstrate how a revised understanding of language highlights three core aspects: that language is inherently multimodal and semiotically diverse; that it functions as a tool for semantic, pragmatic, and social inference, as well as facilitating categorization; and that the processes of interaction and transmission give rise to central design features of language.

Highlights

What makes human language unique? In 1960, Hockett proposed 13 'design features' of language to identify commonalities with, and differences from, other nonlinguistic signaling systems and animal communication systems. His framework was highly influential in shaping language evolution research and cross-species comparative work.

Yet, in the 65 years since its publication, major advances in our understanding of language and animal communication demand a rethinking of what characterizes language and what makes it unique.

We re-evaluate Hockett's framework, focusing on three key features of language: its multimodality and semiotic diversity, derived from our ability to transform any behavior into a communicative act; its use for ostensive communication, social signaling, and cognitive augmentation; and its nature as an adaptive system, stressing the role of interaction and transmission.

'Design features': a traditional comparative framework

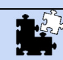
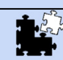
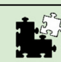
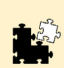
What, if anything, makes human language distinct from the communication of other animals? In 1960, the linguist Charles Hockett proposed a comparative, evolutionary approach to the study of language origins by enumerating a set of 13 'design features' that highlight the supposed commonalities and differences between human (spoken) language, other human signaling systems, such as music, and the communication systems of other animals (e.g., the honeybee waggle dance or gibbon calls) [1]. By breaking language into its constituent parts, Hockett aimed to determine which features are shared with other species and which combination of features makes human language unique. This framework [1,2] was extremely influential in shaping the field of language evolution and cross-species comparative work: over the past 5 years alone, his seminal paper [1] has been cited over 800 times. Hockett's 'design features' are also frequently included in introductory textbooks in linguistics and cognitive science [3], a testament to the appeal of this approach.

However, in the 65 years since its publication, many of Hockett's design features are in need of major revision (Figure 1). For example, instead of being synonymous with speech (see Hockett's Feature 1: 'vocal-auditory channel'), modern scholars now understand that sign languages are also fully fledged languages, requiring a multimodal framework for the comparative study of language [4,5]. Moreover, although Hockett characterized language as having 'duality of patterning' (i.e., the ability to combine meaningless sounds into meaningful words, and meaningful words into more complex meaningful sentences), we now know that sign languages can be fully productive without phonological combinatoriality [6]. In addition, contrary to Hockett's view that language fundamentally comprises 'arbitrary' signals (Feature 5), we now know that iconicity (i.e., motivated resemblance between form and meaning) is prevalent in languages, both spoken and signed [7,8].

Notably, although Hockett did not claim that his design features are unique to human language [1,2], recent research demonstrates that most are far more common across nonhuman animal communication systems than he originally envisioned (e.g., reviewed in [9]). For example, some nonhuman communication systems, including those of primates, such as Campbell's monkeys,

[1]Center for Language Evolution Studies, Nicolaus Copernicus University in Toruń, Toruń, Poland
[2]Department of Linguistics and Communication, University of Birmingham, Birmingham, UK
[3]Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA
[4]Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[5]Department Life Sciences, Erasmus University College, Erasmus University Rotterdam, Rotterdam, The Netherlands
[6]LEADS Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[7]Donders Centre for Cognition, Radboud University, Nijmegen, The Netherlands

*Correspondence:
m.perlman@bham.ac.uk (M. Perlman).

Relevant theme	Design feature	Original definition	Where it stands now	Status
Multimodality and semiotic diversity	(1) Vocal-auditory channel	Language is produced by articulating sounds with our vocal tract and perceived through our sense of hearing	Language is multimodal and not restricted to speech. Humans readily learn visual (signed) languages, and the use of language involves multiple modalities, including manual gestures and facial expressions	
	(2) Broadcast transmission and directional reception	When humans speak, sound emanates outward into the environment and can travel in different directions. Listeners are able to localize the direction of the source of the sound	The dynamics of signal transmission and reception depend on the modality of communication. For example, sign languages require line of sight and, thus, are broadcasted more narrowly than are spoken languages	
	(3) Rapid fading (transitoriness)	Speech exhibits a temporary quality because the sound waves of an utterance are only perceptible for a brief instant of time before they dissipate	Temporal dynamics depend on the modality of communication. For example, unlike speech and sign, which demand immediate processing before the signal fades away, written text is persistent in time, enabling longer and more complex sentences	
	(4) Total feedback	Speakers have immediate access to the speech they produce, allowing them to monitor and modify it	Feedback is present in all communication systems, not only in speech. Feedback comes from multiple senses (auditory, visual, and proprioceptive) and depends on the modality of communication	
	(5) Arbitrariness	Word forms bear no resemblance to their meaning	Language combines arbitrary components with motivated ones. Iconicity is widespread in both signed and spoken languages	
	(6) Discreteness	Language distinguishes meanings through categorical (discrete) rather than analog contrasts	Language includes a mix of discrete and graded features, such as prosody and iconic gesture	
Functions of language	(7) Inter-changeability	Language users can take on both sender and receiver roles and can equally produce and comprehend any message that is expressible in language	Proficiency in production and comprehension can be asymmetric. For example, in the case of language learning or passive bilingualism, individuals may be able to comprehend a language without being able to speak it proficiently	
	(8) Specialization	Language is designed and used specifically for the purpose of communication	Language serves multiple functions, including ostensive and non-ostensive communication, social signaling, and cognitive augmentation	
	(9) Displacement	Language enables the ability to communicate about things that are not present in the here and now	Displacement is fundamental to language and crucially depends on ostensive-inferential communication to identify relevant displaced referents	
Language as an adaptive system	(10) Semanticity	Words have specific meanings, which are characterized by fixed associations with recurrent features or situations of the world around us	Linguistic expressions are made meaningful through a process of inference, which draws richly on linguistic and extralinguistic context. Meanings are not fixed but flexible	
	(11) Productivity	Language enables the ability to produce and understand novel meanings	The expression of novel meanings is achieved by syntactic compositionality and conventionalized lexical innovation, which come about in the process of interaction and cultural transmission. Iconicity offers another route to productivity by enabling the creation of meaningful utterances without reliance on conventional form-meaning associations	
	(12) Traditional transmission	Languages and linguistic conventions are transmitted nongenetically by (social) learning and teaching	Language can be learned in the absence of explicit teaching. The processes of interaction and transmission within and across generations give rise to other design features, such as productivity, arbitrariness, and semanticity	
	(13) Duality of patterning	Meaningful messages exhibit two levels of structure: (i) phonological combinatoriality (combining meaningless elements into meaningful units, e.g., sounds to words); and (ii) syntactic compositionality (combining smaller meaningful elements into larger meaningful units. e.g., words to sentences)	Phonological combinatoriality and syntactic compositionality are two separate features of language that can develop independently of each other. For example, emerging sign languages can be fully productive without phonological combinatoriality. While all languages are compositional, they vary in how compositionality is divided between morphology, syntax and the lexicon	

Trends in Cognitive Sciences

Figure 1. Re-evaluation of Hockett's 'design features' of language. Illustration of Hockett's classic design features with their original definition, a short summary of where they stand today in light of modern research, and a status update for each one. Red 'thumbs down' indicates that an original feature is considered inaccurate. Puzzle pieces indicate that an original feature is considered incomplete. A 'mouth' indicates that this feature only applies to spoken language. Green 'thumbs up' indicates that an original feature still holds true. The background color of the cells indicates which of the three themes highlighted in the current paper is the primary reason for this update.

and of birds, such as pied babblers, also show evidence of combining basic units into more complex ones [10–12]. Similarly, communicating about referents not present in the here and now ('displacement', Feature 9), although rare in the animal kingdom, has been observed not only in honeybees, but also in ants and great apes [13–16].

Considering scientific findings such as these, Wacewicz and Żywiczyński concluded that Hockett's design features are a 'nonstarter', arguing that comparative investigations of human language and animal communication need to move away from thinking of language as a set of static features, and instead consider the capacities and processes that make language possible [17]. In this review, we offer a 'status update' on how well Hockett's original features capture the properties of language (Figure 1) and discuss why recent developments in our understanding of human language and nonhuman animal communication demand a radical rethinking of the nature of human language and the processes underlying it (Boxes 1 and 2). We also show that integrating new evidence and theoretical advances can nevertheless lead to a productive comparative approach.

Specifically, we focus on three broad themes that characterize our modern understanding of language and show how each of these lead to a fundamental re-evaluation and reconceptualization of Hockett's design features: (i) multimodality and semiotic diversity; (ii) the functions of language; and (iii) language as an adaptive system. For example, while Hockett conceived of language as characterized by a set of static features, modern research views language as a fundamentally dynamic and adaptive system, which is continuously changing [18]. In this sense, the design feature of 'traditional transmission' is best thought of not as a static feature, but as the social processes of interaction and transmission, which drive the creation of other design features, such as 'arbitrariness' (Feature 5), 'semanticity' (Feature 10), and 'duality of patterning' (Feature 13), as shown in experimental research on the cultural evolution of language [19–21]. Similarly, Hockett's feature of 'specialization' (Feature 8) suggested that the function of language is the transfer of semantic information using a discrete 'code'. However, a more contemporary understanding of the functions of language requires us to move beyond treating language as merely a system for communicating

Box 1. Lessons from nonhuman animals: the value of comparative work in language evolution

Years of animal communication research have shown no evidence of language in nonhuman animals. Yet, some of their communication systems share key features with human language (e.g., multimodality, social signaling, and combinatoriality), suggesting that these traits have not evolved exclusively in humans. Thus, a comparative approach to language can be highly informative in understanding which biological and social factors drive the evolution of these shared features. However, the communication systems of many animal species remain poorly understood, mostly due to a lack of species-appropriate paradigms and measures that take into account the unique biology and socioecology of each species [181].

Notably, comparative work has focused largely on vocal and gestural communication, probably because these modalities are most salient to humans. However, without properly considering how species actually perceive these signals, and without paying more attention to modalities or features to which we as humans are less sensitive but are highly prominent in other species (e.g., vibration), we may never be able to fully appreciate their communication systems. For example, zebra finches, which are seen as promising models for understanding the evolution of human vocal learning, produce songs characterized by sequences of elements arranged in fixed patterns [182,183]. However, recent psychoacoustic work suggests that zebra finches care less about song sequences and instead attend to the spectrotemporal features of smaller song units to which they are sensitive at a level beyond human perception [184].

Machine learning can help to overcome these biases, and can improve our analysis of animal signals by picking up on subtle regularities, integrating different modalities, identifying function by systematically mapping signals to their social and environmental contexts, and comparing signals from different species at an unprecedented level [185]. However, despite the potential far-reaching applications of machine learning, we should remain aware that our intrinsic human biases [186] can still influence their outcomes. Nonetheless, a comparative approach that relies on less subjective tools may be highly informative in uncovering the complexities of the communication systems of other animals, as well as the evolutionary origins of features that are shared with human language.

Box 2. What counts as language?

Notably, there is currently no consensus on the definition of ‘language’ and how much it overlaps with ‘communication’. Different theoretical frameworks range from more restricted to more expansive views of what counts as language [187] (Figure I). Different scholars prescribe to different views along the restrictive–expansive continuum, which guides their interpretation of neural and behavioral patterns as well as their expectations from theories on language learning, use, and evolution.

A longstanding approach, which is largely rooted in the generative linguistics program, holds that language comprises a mental dictionary of words and an abstract grammar for combining them [188,189]. In these more restricted views, processes, such as pragmatic inference, and communicative signals, such as iconic gestures and gaze, are outside the bounds of language. The influence of these views can be seen in recent work that aimed to map the ‘language network’ in the brain, which is thought to encompass language-specific representations of ‘sounds, words and syntactic structure, along with a large set of form-meaning mappings for words and constructions’ [190] that are used to encode and decode linguistic messages. These views have focused on the highly standardized, systematic properties of human communication, treating language as a modular system that exists independently of more spontaneous and multimodal elements of communication, such as iconic gestures, prosody, and facial expressions.

However, scholars favoring a more expansive approach consider spontaneous and multimodal aspects of communication to be just as distinctive of human languages as vocabularies and grammars. Within this expansive approach, language is seen as a multimodal system that includes both conventionalized and spontaneous elements, such as gestural-kinematic, and visual aspects, including co-speech gestures [4,24]. In this view, the uniqueness of human language is that it can use and recruit multiple semiotic resources through ostension, or ‘means for making meaning’ [191], including elements that would be considered ‘nonlinguistic’ under a restrictive view.

Regardless of whether one’s use of the term ‘language’ is reserved for only the most systematic and conventional elements of human communication, such as those found in vocabularies, grammars, and sound systems, the point is that humans’ extraordinary expressive power is derived from our (seemingly unique) capacity to turn nearly every behavior into a communicative act through ostension.

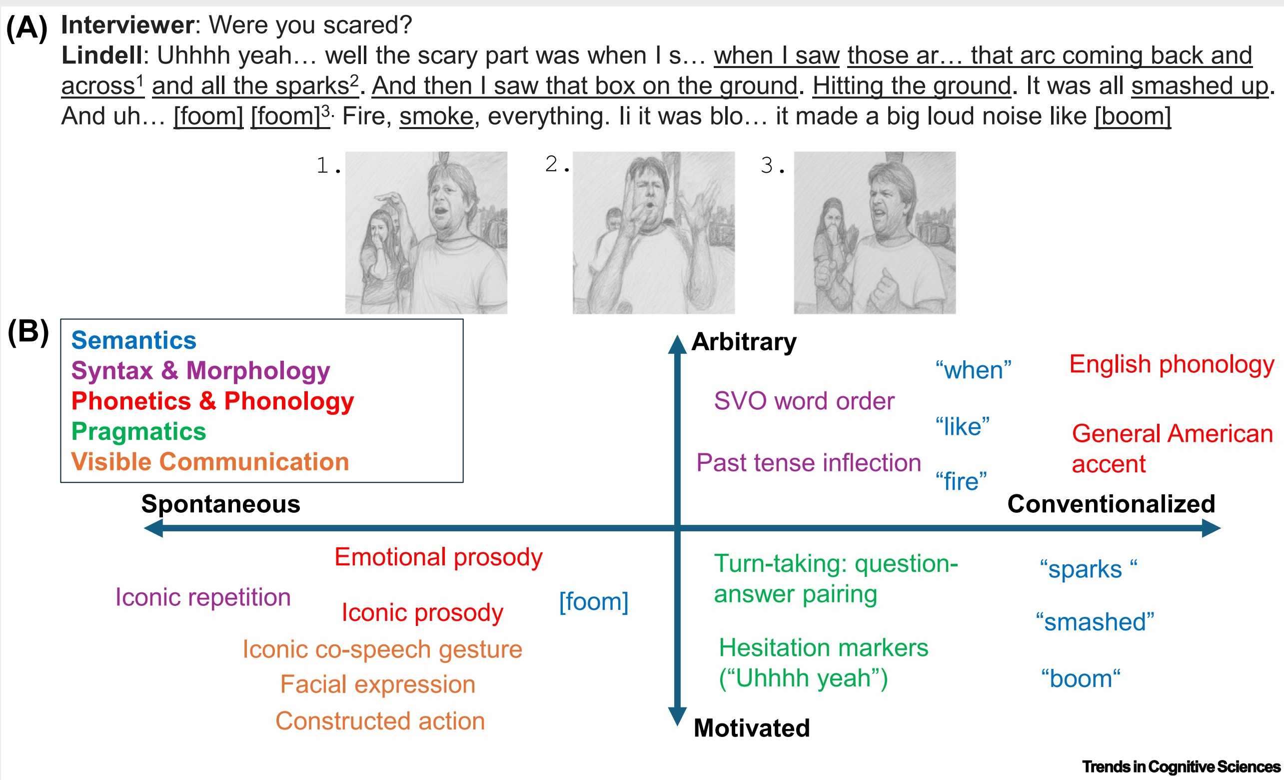


Figure I. Language in a restricted versus expansive perspective. (A) Annotated example of a real-life speech act with corresponding visual snippets, with underlined text indicating a co-occurring gesture and brackets indicating a nonstandardized vocalization. (B) How different aspects of communication (spanning lexical semantics, syntax and morphology, phonetics and phonology, pragmatics, and visual signals) can be characterized along two axes: from arbitrary to motivated, and from spontaneous to conventionalized. Different views would consider different elements as ‘language’. A more restricted view includes only conventionalized elements, which are treated as arbitrary components (e.g., words, word order, and phonetics; top-right quadrant). A more expansive view also includes spontaneous, pragmatic, visual, and clearly iconic components as ‘language’ (e.g., emotional and iconic prosody, facial expressions, co-speech gestures, and hesitation markers; bottom-right and bottom-left quadrants). The top-left quadrant of spontaneous and arbitrary components is empty, because a truly arbitrary and unconventionalized signal would not be understood due to the lack of any common ground. (A) Based on www.youtube.com/watch?v=6L94Qy_D998&ab_channel=FOX10Phoenix.

discrete, semantic information, and instead also consider the other functions it serves, such as social signaling and cognitive augmentation. Moreover, because meaning-making in language relies heavily on pragmatics, ostension, and inference [22], we highlight ostensive-inferential capacities as a central component of human language and cognition, which enable interlocutors to turn just about any behavior, whatever the modality, into a communicative act. In this view, it is the design features of human interaction [23] that are central to the emergence of many of the design features of human language.

Multimodality and semiotic diversity

The title of Hockett's article, 'The origin of speech', reflected the predominant view during the mid-20th century that language and speech were synonymous. This explains why the 'vocal–auditory channel', 'broadcast transmission and directional reception' (Feature 2), 'rapid fading (transitoriness)' (Feature 3), and 'total feedback' (Feature 4) were essential features of language for Hockett. Scholars of that era did not pay much attention to the role of visual information in speech and did not recognize that sign languages were fully fledged languages. Many researchers now consider multimodality and semiotic diversity as vital properties of language [4,7,24–27].

Language is multimodal in two ways. First, it exhibits modality flexibility and can be implemented in different modalities. Sign languages used in deaf communities, of which more than 200 are currently documented, with more documented each year, have all the linguistic complexity and expressive power of spoken languages [5]. Moreover, neuroscientific work shows that signed and spoken languages are processed similarly in the brain [28]. In some relatively isolated communities, an incidence of heritable deafness of only a few percent of the population is enough to spur the creation and maintenance of sign languages [29]. In the case of Nicaraguan Sign Language, interactions among only a few cohorts of students attending a newly established school for the deaf were sufficient for the emergence of a new sign language [30]. Moreover, sign languages may not be limited to only the visual modality. Around 2006, a tactile communication system, known as Protactile, began to emerge in a deaf–blind community in Seattle, USA, and is already developing core linguistic properties, including phonological and syntactic structures [31] (Box 3). The modality flexibility of language is also apparent in the ease with which people adapt to new technologies, such as video conferencing, texting, and emojis, which quickly develop systematic conventions similar to spoken and signed languages[i]. For instance, Emoji show emerging grammatical patterns, such as the use of reduplication as an intensifier, for example 💰💰💰 indicating 'a lot of money' [32].

Box 3. Protactile

Protactile is an emerging communication system used by a community of DeafBlind signers in Seattle, USA [31]. Distinct from visual American Sign Language, Protactile is created by DeafBlind people themselves, designed for tactile communication, and its users note that it enables dramatically more fluid and effective communication. Users communicate via reciprocal, tactile channels, with each person using their hands and arms, upper back, and, when in a seated position, knees and the top of the thigh as contact spaces for interaction. Signers coform utterances by using not only their own, but also their interaction partner's body, who they are in physical contact with (see [192] for examples from Norwegian and Swedish signers). Researchers and users of Protactile describe the emergence of a phonological system, implemented through proprioceptive constructions, comparable to classifier constructions in sign languages [31]. Proprioceptive constructions are articulated jointly between a conveyer and a receiver working together in standardized ways to articulate the language. Similar to visual sign languages, Protactile features a large amount of iconicity, which is rooted in spatial and tactile, rather than in visual, form-meaning resemblance.

Tactile sign languages are a testament to the incredible human ability to create language in different modalities, making use of whatever semiotic resources are available. They challenge long-held theories about the nature of language, providing compelling evidence for the need to re-evaluate the view of language expressed in Hockett's design features. Tactile sign languages demand a view where language is multimodal and semiotically diverse (see 'Multimodality and semiotic diversity' in the main text), where meaning is jointly constructed in interaction based on ostension, inference, and context (see 'Functions of language' in the main text), and where structure dynamically emerges over the course of (and is shaped by) interactive encounters (see 'Language as an adaptive system' in the main text).

Second, even within a particular modality, language is profoundly multimodal. For example, spoken language involves the tight integration of vocalizations with visual signals, including manual gestures, facial expressions, gaze, and movements of the head and torso [27,33–35]. Sign languages also make use of multiple channels, including the face and torso as well as the hands. Traditionally discounted as falling outside the realm of language, these communicative signals are not merely supplementary to language, but are critical to how meaning is constructed and derived during interaction [4]. For instance, co-speech gestures affect what mental model is constructed by receivers [36], and the visible information in manual gestures and facial expressions can influence the auditory perception of speech itself [37,38]. In this contemporary and more expansive understanding of language (Box 2), speech, gestures, and other visible signals are all complementary parts of an integrated multimodal system. While speech can be understood without these multimodal aspects, such as when talking on the telephone, face-to-face interaction has historically always been the primary context for communication in our species, emphasizing that language thus evolved as a multimodal system [39].

Once the multimodal aspects of language are integrated in our theoretical thinking, one is required to also consider the ‘semiotic diversity’ of languages, combining not only multiple modalities, but also multiple methods of signaling [8,25,26]. For example, Hockett focused on the ‘discreteness’ (Feature 6) of linguistic forms (i.e., words) and the contrastive function of phonemes to categorically distinguish one word from another, assuming that all word forms have an ‘arbitrary’ relationship to their meanings. While it is true that many spoken words (and signs) are discrete and appear arbitrary, it is now widely accepted that all languages regularly use signals with ‘motivated’ correspondences between form and meaning (i.e., iconicity; Figure 2), which serve important functions in language learning and development [40–42], as well as in language evolution [43,44] (Box 4). These iconic correspondences are often graded, with modifications in form conveying analogical differences in meaning. Across sign languages, many signs are depictive and many also index locations on the signer’s body or in space [45]. Major components of signed grammars are also iconic [46], including devices such as constructed action, in which signers use their bodies to physically enact the roles, actions, and expressions of people, animals, or other entities [47]. Furthermore, signs are often graded, and can be iconically modified to express a particular meaning in context, for example, modifying the shape and size of the sign for ‘shoe’ to depict a specific form of shoe [48], or modifying the form of ‘swim’ to show a specific manner of swimming [49]. While early sign language researchers downplayed these functions [50], the vital role of these semiotic modes is now increasingly recognized. In fact, some researchers have proposed that, if the study of language had begun with sign languages, multimodality and iconicity would have taken on a prominent role much earlier [51].

Spoken language also features much more iconicity than previously appreciated (Figure 2). First, many of the gestures that people produce when they speak are clearly motivated by the meaning being expressed. These include gestures such as pointing at relevant entities, enacting actions and events, depicting the shapes of objects [27], and even representing more abstract ideas through metaphor [52]. Crucially, speech itself also features a lot of iconicity. Spoken languages contain a substantial number of iconic words forms, including onomatopoeic words for sounds (e.g., ‘buzz’ or ‘crash’) and ideophones more generally, which are now understood to be a distinct lexical class of highly vivid, depictive words found in many, if not most, spoken languages [53]. Moreover, cross-linguistic statistical analyses show that iconicity exists all over the lexicon: across many different languages, words with the same meanings tend to share similar sounds (e.g., the high front vowel /i/ is frequently used in words for ‘small’, and rhotic consonants such as /r/ often appear in words for ‘rough’) [54,55]. These iconic patterns extend beyond isolated words to larger portions of the vocabulary within a particular semantic domain, such as size

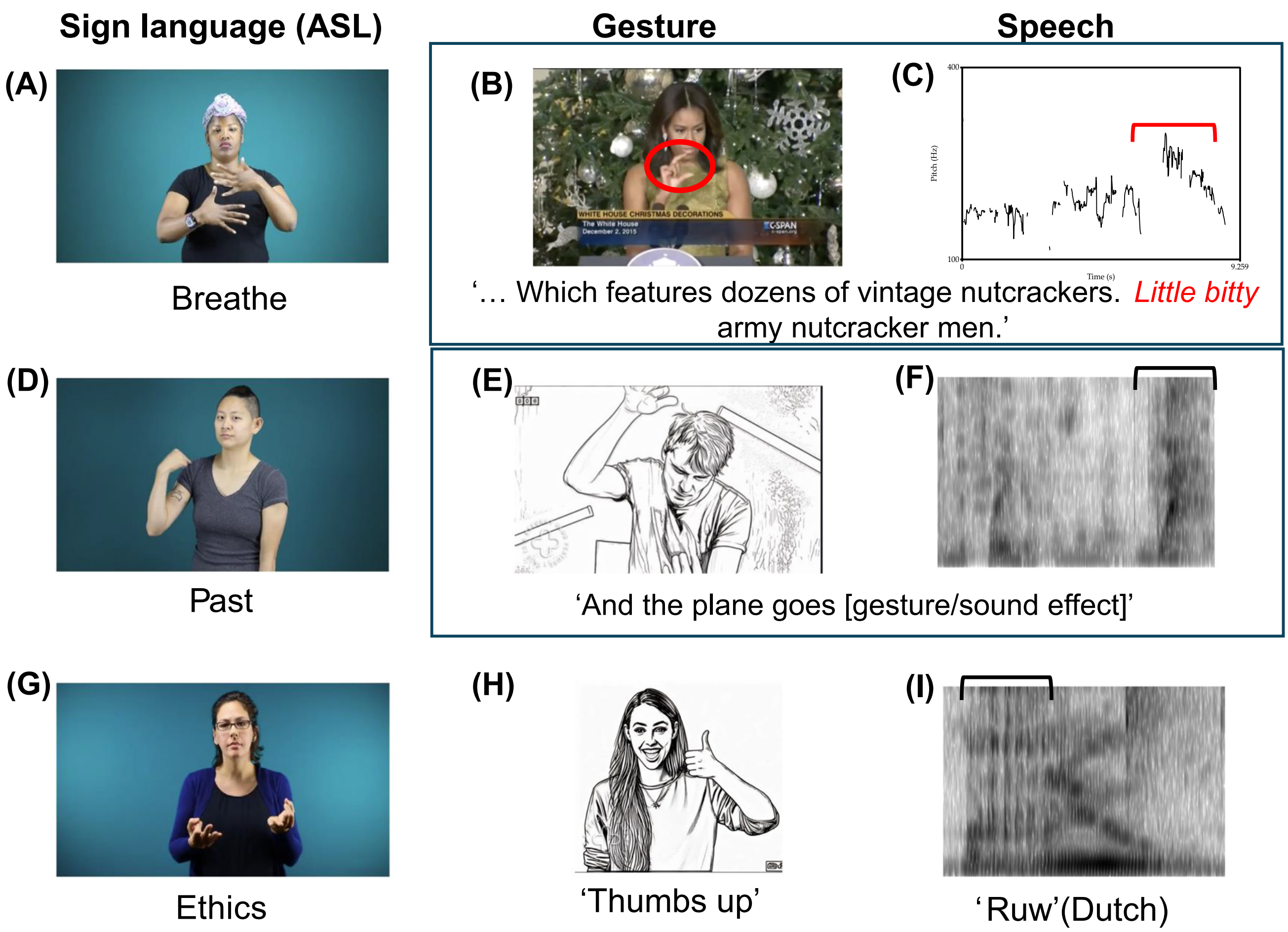


Figure 2. Iconicity across modalities. (A–C) Iconic signs from American Sign Language, illustrating how iconicity can depict both concrete and abstract meanings. (D, G) A speaker producing an iconic gesture as they utter the phrase 'little bitty' with heightened pitch. The iconic gesture (D) and the iconic high pitch corresponding to 'little bitty' (G) are marked in red. (E,H) A speaker describes the sudden descent of an airplane with a multimodal expression ('the plane goes…'), depicting the descent with an iconic gesture (E), and a corresponding sound effect marked in a bracket in the spectrogram (H). (F) A person producing an iconic 'thumbs up' emblem, where the upward orientation of the thumb aligns with a widespread conceptual metaphor in which 'up' is associated with positive valence. (I) Spectrogram of iconicity in speech, using the Dutch word 'ruw' (meaning 'rough') that includes trills in its initial phoneme (marked with a bracket); previous research showed that trilled /r/ is iconic of roughness across many spoken languages [57]. (A–C) Reprinted from https://aslsignbank.haskins.yale.edu/. Abbreviation: ASL, American Sign Language.

adjectives [56] and texture descriptors [57]. Moreover, spoken languages also make grammatical use of iconic strategies by, for example, using the order of words to represent the temporal sequence of events and placing forms closer together in correspondence to concepts that are more closely related [58].

Semiotic diversity is also reflected in how we say words (the prosody of speech), including intonation, rhythm, tempo, and timbre. Prosody is shaped by a range of factors, such as speakers' attitude on a topic, the prominence of particular words within a sentence or conversation, the syntactic structure of the utterance, and the type of speech act (e.g., question versus statement). While prosody is partly guided by conventional patterns, much of it is spontaneous and dynamically molded in graded ways that are tuned to context and the particular meaning being expressed. Similar to iconic gestures, speakers often modify their voice to depict meaning, Box 4. Iconicity and the evolution of language

Whenever people need to communicate without a shared language, they instinctively resort to iconic signals (pantomime, sound effects, or drawing) to get their meaning across [20,156,193]. Iconicity enables people to communicate flexibly and productively even without the use of shared conventions, a function familiar to anyone who has ever played 'charades'. This capacity to innovate meaningful signals is fundamental to the creation of symbol systems, including signed and, hypothetically, spoken languages [149].

While the role of iconicity in the origins of signed languages is obvious, its role in the creation of spoken languages is more opaque. 'Gesture-first' theories postulate that sign must have preceded, and served to bootstrap, the formation of spoken symbols because of the assumed superior potential for iconicity in gesture over speech [93]. However, there is now ample evidence for iconicity in spoken languages, and semiotics experiments show that people can create iconic vocalizations to express a range of meanings [194], and that these novel vocalizations are understandable by people from diverse linguistic backgrounds [195], showcasing the potential to use vocalizations to communicate without convention. This research suggests that iconicity is at the root of all languages, spoken and signed alike.

The importance of iconicity in language evolution raises questions about the phylogeny of humans' iconic faculty. While language-trained great apes can learn to use rudimentary systems of arbitrary symbols in different modalities [196], it is not clear whether they are able to make use of iconicity in their communication. Apes produce gestures that appear iconic [197], including directive visible movements [198] and pantomimes of different actions [199]. However, it is disputed whether they have any psychological awareness of the iconicity of their gestures as humans do [200], or whether the gestures are formed instead from the ritualization of functional actions in development [201] or evolution [202]. Recent work argues that great ape gestural forms are recruited from action schemas and natural bodily movements leading to an 'iconic' resemblance between the form of a gesture and its meaning [92], although it is debated whether apes are reflectively aware of this resemblance for communication [203]. Therefore, the extent to which we see precursors of iconic communication in our closest relatives remains to be determined.

such as slowing their tempo to represent a 'sloooow' event or raising their fundamental pitch to describe something 'teeny' [59].

As this overview shows, multimodality and semiotic diversity are ubiquitous in human language (Figure 2). Under this lens, language is an integrated system that combines conventionalized and less conventionalized forms that vary in the degree to which they are discrete versus graded and arbitrary versus iconic (Boxes 2 and 4). One counterintuitive consequence of such semiotic diversity and language being an adaptive system is that it allows for inventing new language modalities, such as written text, which strips away multimodal resources, including co-speech gesture and prosody (supplementing some of the latter with punctuation marks). It is tempting to conclude that the ability of humans to master written language means that language need not be multimodal. However, all used written languages are originally derived from spoken languages. Interestingly, even though written text alone appears to contain sufficient information/structure to enable a large language model (LLM) to learn a productive grammar (Box 5), the fact that such models require huge amounts of data compared with human learners may partially stem from the absence of multimodality [60].

Box 5. Language as training data for machines and humans

LLMs, such as ChatGPT, are trained largely through next-token prediction, wherein the model learns a set of weights for predicting what words are most likely given a previous linguistic context. The remarkable efficacy of this approach to learning basic aspects of word meanings was first demonstrated by Elman's simple recurrent networks [204]. The transformer architecture used by modern language models [205] descends from this tradition [206], but allows for training on orders of magnitude more data and to incorporate context in more flexible ways.

The idea that self-supervised systems can learn complex structure in the course of lowering prediction error is now a mainstay of cognitive science and neuroscience [207,208]. Previously, however, this learning was applied to specific domains. A system designed to lower prediction error in visual input can, for example, learn to integrate visual contours [209] and recognize objects [210]. Nevertheless, the productivity of language means that the information embedded in it is far broader, such that, when the same predictive approach is applied to language, a model ends up learning far more than language itself [211], but a surprisingly rich and varied amount of world knowledge. A system tasked with predicting 'Humpty ___' can succeed by learning simple bigrams; one tasked with predicting 'Paris is the capital of ___' can get by with learning isolated facts. However, as the variety of what needs to be predicted grows (e.g., 'After the accident, he was___'; 'The banana turned ____', 'A sequence is divergent if its limit ____'), success requires learning more general models [212]. Investigating these 'world models' is an active area of research [213]. The flexibility with which LLMs deploy such knowledge in the service of tasks such as answering questions from a range of standardized tests [214] and about novel documents [215], style transfer [216], and even creating a new language[ii], far exceeds any system that merely parrots its input [218] or 'auto-completes'[iii].

Showing that LLMs are good models of human cognition requires more than impressive performance on benchmarks. That said, the finding that exposure to language alone can lead neural networks to succeed on such a wide range of tasks raises the question of whether the human ability to excel at such a wide range of cognitive tasks may depend more on linguistic 'entrainment' than generally acknowledged [75,219].

Adopting a view of language as being characterized by semiotic diversity means that the entire integrated system of multimodal ways of meaning making should be the object of study. From this vantage point, we can see evolutionary precedents for some of the key features of language. More specifically, such a point of view can make productive contact with recent research in nonhuman animal communication, which has increasingly highlighted the role of multimodality [61], with evidence of multimodal communication found in primates, elephants, birds, lizards, frogs, and spiders, among others [62–64]. For example, nonhuman primates integrate gestural, facial, and vocal channels during communication [65–67]; anurans, such as squirrel treefrogs and poison-dart frogs, combine vocal signals with visual displays (e.g., besides helping vocal production, inflating their vocal sacs also serves as an important visual signal for mate attraction and territorial defense) [68]; and social insects, including ants, use multimodal signals with chemical and vibrational properties [69]. Thus, in the animal kingdom, multimodal communication appears to be the norm rather than the exception.

While there are many examples of multimodality in the communication systems of nonhuman animals, much less is known about the ability of other animals to flexibly adjust their use of different modalities during communication. Interestingly, recent work suggests that several species can flexibly adjust their signals, at least to some extent, based on their social and/or physical environment. For example, African savannah elephants and chimpanzees will appropriately switch modalities and use audible or tactile gestures if their intended receiver is not paying visual attention [62,70]. Moreover, male wolf spiders will use more seismic signals to attract females if substrates in their habitat have good vibration conduction properties, but will use more visual signals on substrates where seismic communication would be less effective [71]. However, more work is needed to determine the extent of modality flexibility in nonhuman animals.

With respect to iconicity, although the kind of 'open-ended' iconicity associated with human communication may be rare in the animal kingdom, it is not unprecedented. For example, in the waggle dance of the honeybee (touted by Hockett as an example of displacement), the duration of the waggle corresponds to the distance to a food source, and the direction of the dance maps the direction of the food relative to the azimuth of the sun [72]. Thus, the honeybee waggle dance constitutes a narrow but productive iconic communication system, encoding specific information about the location of food in a motivated and flexible way. This combination of fixed and variable iconic components may be seen as schematically analogous to some (vastly more flexible and semantically complex) classifier constructions in sign languages, whereby a conventionalized handshape representing an entity (compared with the genetically fixed 'waggle' of the honeybee dance) is combined with variable movement to depict aspects of the movement of the entity, such as its direction, speed, and manner [46]. However, the question of whether our closest relatives, the great apes, show evidence of iconicity in their communication remains controversial (Box 4).

Functions of language

In Hockett's view, the production of language 'serves no function except as signals'. This feature of 'specialization' was intended to capture the purely communicative nature of language, distinguishing it from behaviors such as the panting of a dog, which may inform others that the dog is hot, but primarily functions to cool the body rather than communicate (in animal behavior terminology, the panting is a cue rather than a signal [73]). The idea that language is 'for' communication has been a recurring theme in the language sciences. For example, attempts to explain the emergence of language in terms of increasing biological fitness (e.g., [18,74]) view communication as the central adaptive function of language. In more recent investigations of the neural substrates of language, researchers interpreted evidence of functional specialization for language processing (the 'language network') as supporting the idea that these regions specialize in mapping between forms and meanings of lexical items and constructions [75], a conclusion broadly supporting Hockett's claim that language is specialized for communication. Here, we outline three important revisions to the feature of 'specialization' concerning the functions of language (Figure 3A). First, while Hockett conceived of linguistic communication strictly in terms of a code model in which meaning is encoded by the producer and decoded by the recipient, linguistic communication is in fact best understood as a process of inference through ostension. Second, in addition to the sort of semantic information Hockett focused on, language also has an important role in social signaling. Third, language also has a cognitive function: augmenting nonverbal cognition.

Ostensive communication

Hockett viewed linguistic communication strictly in terms of a code model: a sender encodes a message in words, and a recipient decodes the message. The code model is behind commonly used expressions such as 'put your thoughts into words', 'I didn't catch that', and 'The meaning of that sentence is hard to unpack' [76]. This way of thinking about linguistic communication, and the broader information theoretic principles from which the code model derives [77], have been enormously useful in understanding language as a process of optimizing for competing pressures [78]. For example, different languages trade off ambiguity and complexity: having a different word for maternal versus paternal grandparents allows for more precise referencing, but requires more complex kinship systems [78,79]. Viewing linguistic communication in terms of codes that need

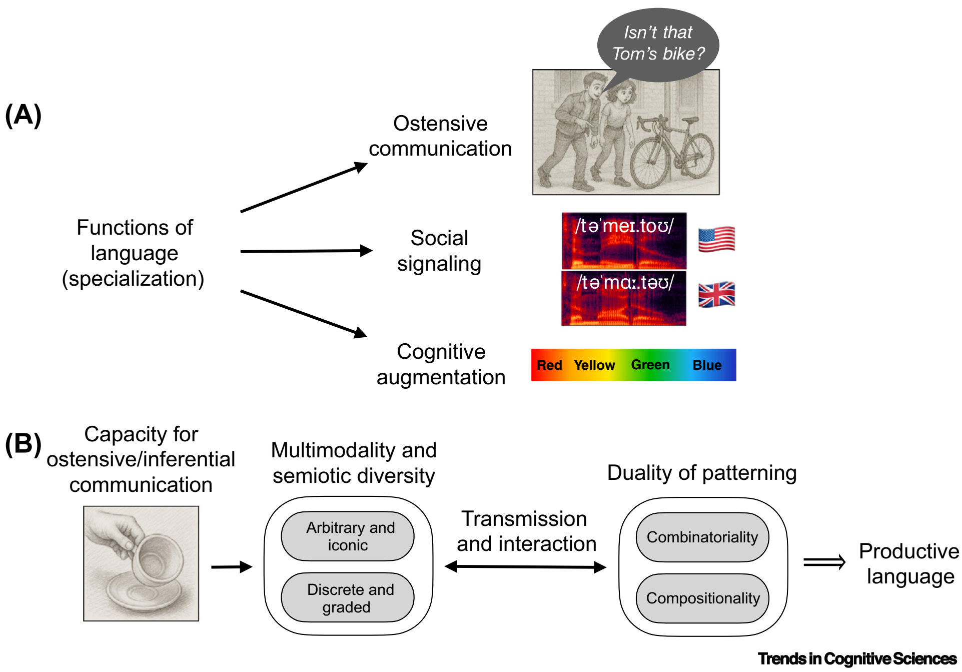


Figure 3. Functions of language and language as an adaptive system. (A) A major function of language is ostensive/inferential communication. The meaning of 'Isn't that Tom's bike?' is not in the words. Rather, the words are cues that enable a comprehender to construct a meaning using their background knowledge and context. Language also appears specialized for social signaling, as when one's accent conveys information about where the speaker is from. The third function is cognitive augmentation, as when continuous color representations are made more categorical through color labels, affecting performance in perceptual tasks, such as color discrimination. (B) Underlying the human capacity to use language is the capacity for ostensive/inferential communication enabling people to turn a variety of behaviors into communicative acts and giving rise to the semiotic diversity that characterizes language use. Combinatoriality and compositionality emerge over the course of language learning (vertical transmission) and interaction (horizontal transmission), enabling productivity which allows users to, in principle, communicate about anything that is within the human conceptual capacity. This includes displaced referents.

to be learned and effectively transmitted also accommodates a range of communication by non-human animals, such as mating, threat displays, and signaling systems, including the honeybee waggle dance. In all these cases, the goal of the signaler is to use signals to produce a change in the behavior of the recipient. However, developments in the philosophy of language, linguistics, and animal communication have shown that this code model fails to capture two key aspects of linguistic communication.

First, the 'conventional' meanings of words of the sort Hockett had in mind (utterance meaning) often depart from the message intended by the speaker (speaker meaning) [80–82]. For example, the intended meaning of a simple expression such as 'Pass the salt' could be a bona fide request to pass the salt, serve as the punchline of a joke, a quote of what someone said, or an attempt to change the subject ('So what'd you do this morning?' 'Uhm, pass the salt'). A view focusing on the importance of inference and pragmatics helps explain the pervasiveness of 'nonliteral' meaning in language, which was neglected in Hockett's framework. Metaphor, metonymy, idioms, and polysemy are now also recognized to be just as central to language as its more literal aspects [83,84]. The experimental and computational study of these inferential processes is one of the fastest-growing fields of linguistic inquiry, with its own formal frameworks, such as the Rational Speech Act framework, aiming to capture how context informs inferences [85,86]. Receivers interpreting signals based on combining information from multiple signals and social context are also increasingly found in animal communication [87,217,220]. For example, for the alarm calls of vervet and putty-nosed monkeys and for the gestural communication of great apes, interpretations are highly dependent on interpersonal, contextual and situational information [88–90].

Second, much of human communication is ostensive. In the field of pragmatics, communication is ostensive when one does not only communicate something, but also intends their interlocutor to recognize that they are trying to communicate something to them. In this view, human communicative acts have both informative intent (to inform the audience of something) and communicative intent (to inform the audience of one's informative intention; i.e., to communicate your intent to communicate [82]). Consider trying to communicate to a waiter that we would like another coffee. We can use language: 'I'd like another coffee'. However, we can also use a variety of nonverbal expressions, such as tilting the empty cup in the direction of the waiter, or raising a finger after our neighbor says that they would like a refill to indicate that we would also like one [23,91]. This adeptness at turning just about any behavior into an act of ostensive communication, be it tilting a cup, exaggerating a smile, or looking disapprovingly, suggests that underlying the ability of humans to use language is a more general capacity for ostensive communication, without which language would not be possible [22,92]. Underlying ostension is a suite of social cognitive capacities and motivations, including joint attention to the perceptual and conceptual common ground (e.g., background knowledge) that we share with our interlocutors [91]. For example, shared common ground can determine whether pointing to a bicycle parked in front of a bar will be interpreted as 'This is our friend Tom's bike, so let's go inside' or 'This is your ex-partners' bike, so let's avoid this bar!' [93].

Ostensive communication is characterized by a high level of voluntary control. For example, when we speak or tilt a cup to communicate wanting more coffee, we do so voluntarily. By contrast, non-ostensive communication tends to be more automatic, as in the case of natural smiles, laughter [94], and spontaneous screams [95] (Table 1). The ease with which people make use of ostensive communication may also help explain the feature of 'displacement'. Recipients try to interpret the communicative relevance of ostensive signals (such as 'Tom's bike'). If there is no relevant referent in their immediate surroundings, this might trigger in the recipient a search for the relevance of a signal beyond the here and now based on common ground between the

Table 1. A comparison of behaviors used non-ostensively versus ostensively[a]

Behavior	Non-ostensive use	Ostensive use
Laughter/smiling/screaming	'Duchenne smile' or natural laughter as a spontaneous response; a scream from an acute injury	Smiling, laughing, or screaming intentionally to convey something, for example ironically laughing in response to something that transparently was not funny
Using a particular linguistic variety	Speaking casually with one's natural sociolect (e.g., Southern US English, Irish English, Flemish, Austrian German, Shanghainese, Brazilian Portuguese, etc.)	Using an accent/sociolect in a way that makes clear that the choice of accent is intentional, for example switching to a nonstandard dialect after hearing one's interlocutor speak in that dialect
Yawning	Yawning due to tiredness	Directing an interlocutor's attention and then yawning in an exaggerated manner to communicate 'this is boring'
Bodily action	Tilting a cup to drink	Tilting a cup to signal to a waiter that we want more coffee

[a]Many behaviors that occur automatically and spontaneously, such as smiling, laughing, yawning, or using a particular accent, can also be used ostensively (i.e., in a way that makes it clear to others that they are being used for the purpose of communication). Our capacity for ostension allows us to turn ordinarily non-communicative behaviors, such as yawning and drinking, into communicative ones.

communicators (such as 'we saw Tom's bike in front of a bar yesterday'). Thus, ostensive-inferential communication and social cognition fundamentally underlie many of the processes that give rise to the design features of language [23,93,217].

Social signaling

Linguistic utterances convey not only semantic information, but also social information. For example, merely using a particular language or dialect already conveys one's status as a member of a particular community, and one's accent in a second language can often give away one's first language. Moreover, linguistic variation within language communities is often tied to gender, social class, and ethnicity [96]. This makes language a highly reliable marker of social identity [97,98]. Individual differences in language use are also distinctive enough to make us individually identifiable. Given the high reliability of such linguistic signals to mark social identity, it is not surprising that young children are very sensitive to them. Indeed, children as young as 5 months show a preference for speakers who use the same accent as them [99], and 5-year olds' social preferences are guided more by a speaker's accent than by visual cues of group membership, such as race [100].

That language is such a rich and persistent source of information of social identity, and that we are so sensitive to social identity information conveyed through language, hints that social signaling is one of the evolved functions of language. Although conveying social information every time we speak is not always desirable, the ubiquity of linguistic social signaling suggests that it is a feature, not a bug. That said, sociolinguistic variation can also be the basis for discrimination when users of more 'prestigious' varieties encounter users of nonstandard varieties or non-native language users [101].

Although social signaling of information in language is often unintentional (making linguistic social signaling an intriguing example of non-ostensive linguistic communication), our capacity for ostensive communication allows us to use linguistic emblems of social identity ostensively, deliberately shifting between different linguistic styles to intentionally express social identity, community, and group membership [102,103] and to strategically align with or distinguish ourselves from others [104]. In doing so, these emblems can serve as linguistic resources for audience design [105]. The fact that our language use is tied to social roles also has implications for the design feature of 'interchangeability' (Feature 7), according to which every language user can send and receive the same messages. However, language use is not directly interchangeable because not everyone can 'do' the same things with the same words [106]. From a social and pragmatic perspective, the function of certain utterances can be directly tied to the social roles of the sender and receiver. For example, speech acts such as 'I now pronounce you husband and wife', or 'I sentence you to 5 years in prison' only achieve their intended communicative function if spoken by someone whose social role is imbued with the power to actually marry a couple or sentence someone to prison.

Social signaling has also been investigated in nonhuman animals and shows many interesting parallels with human language. For example, vocal convergence (i.e., where the acoustic features of different individuals become more similar over time) is an affiliative behavior that helps to form and maintain social bonds [107], and has been shown in several nonhuman animals, including, among others, songbirds [108], bats [109], elephants [110], cetaceans [111], and primates [112]. Over time, such vocal convergence reduces the amount of acoustic variation within groups while increasing the acoustic variation between groups, leading to the establishment of group markers. Vocal behaviors that signal group identity can translate into group variation for specific call types (e.g., Guinea baboons [113]) or group-specific call orders (e.g., bonobos [114]). Some nonhuman animals even use family member-specific vocal labels to refer to other individuals in the group (e.g., marmosets [115] and elephants [116]). Moreover, acoustic signals often carry important information about individuals' biological characteristics, such as age (e.g., red deer [117]), sex (e.g., African clawed frog [118]), and body size (e.g., pig [119]), which help to mediate social interactions, such as mate competition, territory defense, and parent–offspring recognition [73]. Besides the spectral information in acoustic signals, temporal information can also be used for social signaling. For example, northern elephant seals can memorize the unique rhythm and timbre of their rivals' vocalizations, allowing them to recognize individual competitors, determine their status in the social hierarchy, and adjust their behaviors accordingly [120].

Cognitive augmentation

Being a competent speaker of a language requires, among other things, learning the set of meanings that comprise its basic vocabulary. Most words denote categories: 'red' denotes a region of color space; 'on' and 'aunt' denote spatial and kin relations, respectively; 'hundred' denotes a cardinality; and 'exasperated', a feeling. Learning the meanings of these words requires learning the categories they denote. Some of these categories would be learned with or without the benefit of learning their names. However, many other categories would not be learned (or would not be learned by everyone in a community) if they were not part of a linguistic system [121,122]. A growing literature shows that, beginning in infancy, words serve as 'invitations to form categories' [123], with language guiding children to learn categories they would otherwise struggle to learn [124–126], and having especially pronounced effects in tasks that require overriding visual appearance in favor of more abstract relations, such as 'on' and 'under' [127].

In addition to guiding people to learn categories that are culturally deemed worth knowing [128,129], experiments show that even informationally redundant labels can facilitate category learning [130,131], shape visual memory [132], and align people's mental representations [133]. Interfering with language produces categorization impairments, particularly when the task calls for selectively focusing on a specific dimension while ignoring other salient dimensions [134,135]. Such 'extra-communicative' effects of language are even evident in lower-level perceptual tasks. Contrary to classic claims that vision is immune to influences from language

[136], verbal descriptions can dramatically transform our ability to recognize an otherwise nonsensical image [137]. Moreover, color words affect the accuracy with which people can discriminate colors [138], and a label can make an otherwise invisible stimulus visible [139,140]. A common thread uniting these findings is that language transforms largely analog and continuous representations into more categorical ones. In turn, this enables more reliable reuse and recombination of representations [141]. The ability of language to augment cognition is powerfully demonstrated by the recent success of LLMs. These systems offer an existence proof of how much functional world knowledge can be extracted from language alone (Box 5).

Findings that language augments cognition are further supported by reports of augmented cognition in language-trained nonhuman animals. For example, language-trained animals, such as Alex (an African gray parrot) and Kanzi (a bonobo), showed cognitive behaviors far outside those of untrained conspecifics, for example with regard to higher-order abstractions, such as shapes, colors, or numerals, and concepts, such as presence/absence and same/different [142,143].

Language as an adaptive system: interaction and transmission

Hockett's framework conceives of language as a product, an endpoint of an evolutionary process. For Hockett, the cultural (i.e., 'traditional') transmission (Feature 12) of language over generations was a characteristic feature of language. However, recent advances in evolutionary linguistics [18] show that 'transmission' is not just one static feature among many others, but can be understood as two complex dynamic processes of interaction and transmission, which continuously shape language itself, even giving rise to other features such as 'productivity' [163] (Feature 11; Figure 3B). In this complex adaptive systems approach, cultural 'transmission' is not simply a design feature of language, but a key process responsible for the very design of language as a whole. In this approach, languages are best understood as dynamic products of ongoing, cumulative interactions between individual users over time, spanning multiple timescales: from the real-time production constraints and accommodation processes happening moment to moment in conversation, to the developmental trajectories of individuals acquiring language, to broad cultural changes in populations occurring within and across generations, to long-term biological changes over evolutionary time [104,144–146]. The empirical study of these dynamic interactions both within and across generations has fundamentally transformed our understanding of other design features, such as 'duality of patterning', 'arbitrariness', 'semanticity', and 'productivity' in that these features are now also seen as inherently dynamic instead of static, and these dynamic features themselves arise out of an interplay between interaction and transmission.

With regard to the design feature of 'semanticity', taking face-to-face interaction as the 'core ecology for language use' [39] has led to an increasing focus on the dynamic co-creation and negotiation of meaning based on pragmatics and inference during interaction [22,23,147]. Experimental work simulating the emergence of novel communication systems shows that symbols acquire their meanings over a repeated process of negotiation, repair, and use, with individuals slowly aligning on shared/agreed meanings based on their cognitive and communicative needs [21,148–152]. These results contrast sharply with Hockett's static view of semanticity, where words and phrases inherently 'have' or 'carry' meaning. It is now clear that the meaning of words and phrases are highly contextual and constantly shaped by interaction. For example, answering 'He is still walking around' to the question 'Is John healthy?' implies that John is in fact not very healthy [153]. The importance of inference also extends to comprehension of single words, such as knowing that 'or' in 'soup or salad' implies either but not both, while in 'cream or sugar' it implies neither, either, or both.

This dynamic view of the construction of meaning in interaction also helps explain the processes involved in ongoing lexical change, whereby words lose and acquire new meanings over time

[154,155]. It further helps explain the prevalence of iconicity alongside 'arbitrariness' across languages, by exposing how repeated use and transmission can transform highly iconic signs into more arbitrary and less transparent ones over time (Box 4) [149,156]. Repeated interactions, observational learning, and (attempts at) imitation are also important factors in the establishment of conventions in some nonhuman animals, such as Guinea baboons and chimpanzees [157–159]. Interestingly, in both human and nonhuman animals, the process of cultural evolution can give rise to diversification over time [96,160], such as when different social groups develop their own unique patterns of communication (e.g., accents or dialects; see 'Social signaling').

'Productivity' can emerge from the processes of interaction and transmission. In all human languages, there are two ways in which smaller units are recombined to create larger units: (i) phonological combinatoriality, where a limited set of meaningless building blocks (e.g., sounds or handshapes) are recombined to create a diverse set of meaningful lexical items (i.e., words, morphemes); and (ii) syntactic compositionality, where these meaningful lexical items are, in turn, systematically recombined to express an almost limitless array of more complex meanings (e.g., compounds and sentences). In Hockett's framework, these two features are collectively referred to as 'duality of patterning'. However, it appears that these features can evolve independently of each other. For example, emerging sign languages can be fully productive without phonological combinatoriality [6,161]. Moreover, multiple behavioral experiments and computational models have shown that both combinatoriality and compositionality, including the higher-level properties of hierarchical structure and recursion they enable, spontaneously emerge as a response to communicative and cognitive pressures operating during social interaction and cross-generational transmission [20,21,151,159,162–167].

The emergence of combinatorial and compositional structure is also affected by sociodemographic factors For example, combinatorial structure emerges faster in populations where peer-to-peer interaction is highly prevalent, and individuals learn most often from other learners as opposed to from more experienced individuals [168]. Similarly, compositional structure emerges faster in larger communities [169], with bigger and more diverse populations worldwide typically using languages with more systematic and regular grammars [170–173]. These results highlight the adaptive nature of language in response to different pressures, such as the need to interact with more people, with novices, or with people with whom there is less common ground. Together, these findings underscore the fact that at least some of the basic architecture of natural language, such as the duality of patterning and productivity of languages, can be explained as an ongoing and dynamic adaptation to communicative needs and processing limitations, emerging over time as part of the continuously changing processes of interaction and transmission.

Importantly, combinatorial and compositional structures have been found in the communication systems of several nonhuman animals, including mammals and birds [11,174,175]. For example, bird species such as chickadees and chestnut-crowned babblers combine meaningless sound elements to create functionally relevant vocalizations, with variations in call combinations linked to specific contexts and behavioral responses [176,177]. Other species, including southern-pied babblers and Campbell's monkeys, also show evidence of compositionality [178,179]. Specifically, male Campbell's monkeys produce distinct alarm calls for leopards ('krak') and eagles ('hok'), which can become alarm calls for more general disturbances or aerial threats by adding an '-oo' suffix [179]. The use of more sophisticated and objective methods to capture and analyze nonhuman animal signals in multiple modalities may reveal that compositionality is even more widespread than previously assumed (Box 1). Although it remains unclear how this combinatoriality and compositionality emerged in nonhuman animals, some communication systems, such as birdsong and whale song, are also culturally transmitted, and recent research suggests that human-like statistical properties similarly emerged in these systems thanks to the process of cultural transmission [180].

Concluding remarks

Developments in cognitive science, linguistics, and animal communication over the past 65 years require a revision of Hockett's 'design features of language'. Some of the original features, such as 'vocal-auditory channel' or 'arbitrariness', have become obsolete owing to a deeper understanding of human language, while features including 'duality of patterning' that were considered unique to humans are apparently more pervasive in animal communication systems than was known in Hockett's time (Figure 1). Other features, such as 'productivity', remain central to language, but are more dependent on cultural processes than originally envisioned. More generally, the picture of human versus nonhuman animal communication that has emerged over these 65 years is importantly different from Hockett's, raising new avenues of research (see Outstanding questions). Our focus was on three key aspects: the multimodality and semiotic diversity of language; the functions of language; and the nature of language as a complex adaptive system.

First, we emphasized the critical role of multimodality and semiotic diversity as a central design feature of language. Language can be implemented in different modalities, such as speech, visual sign languages, and even tactile sign languages (i.e., modality flexibility; Box 3), and typically integrates multiple modalities simultaneously, as, for example, in combining speech with gesture and facial expression (Figure 2). Language also exhibits semiotic diversity, featuring both discrete and arbitrary signals as well as graded, iconic ones. While the use of multimodal signals is found all over the animal kingdom, the extent to which other species are able to flexibly switch between modalities outside their regular repertoires or spontaneously generate iconic signals is less clear (Box 4).

Second, we highlighted the fact that there are multiple coexisting functions of language that go beyond exchanging information (Figure 3A). Although the centrality of this function is widely acknowledged, current evidence demands a more pluralistic view by which language is also used for signaling social information, such as group membership, as well as for cognitive augmentation. Due to its reliance on ostension and pragmatic inference, language makes use of multiple semiotic resources to flexibly convey meanings, allowing humans to turn just about anything into a communicative act (Table 1 and Box 2). While the communicative signals produced by nonhuman animals also often serve as important markers for group identity, less is known about the impact of newly learned signals on their cognition, or about their ability to turn non-ostensive behaviors into ostensive ones.

Lastly, while Hockett's view of language was characterized by a list of static features, a modern understanding of language views it as an inherently adaptive and dynamic system, requiring us to shift our focus to the processes of interaction and transmission that give rise to these features in the first place (Figure 3B). In this view, cultural transmission is not simply another feature of language, but rather encompasses two processes, transmission and interaction, which shape its structure and give rise to other features, such as 'productivity', 'semanticity', and 'duality of patterning'. This adaptive system view emphasizes that languages are best understood as dynamic products of ongoing, cumulative interactions between individuals over time and across broad timescales: from the real-time production constraints of daily conversation, to population-wide cultural and biological changes. Notably, this view highlights the fact that language is a social tool, used for social purposes and evolving under social pressures. While we know that combinatoriality and compositionality (which together create 'duality of patterning'), as well as 'semanticity' can also be found in the communication systems of other animals, the potential role of cultural transmission in shaping their evolution remains an intriguing possibility.

Outstanding questions

What was the evolutionary impetus for the emergence of ostensive communication in the human lineage?

How can we determine which signals have an important role in the communication systems of other species?

What methodological, biological, and ethical considerations should projects attempting to understand nonhuman communication using machine learning models (e.g., Earth Species Project) take into account?

To what degree do other animals make use of iconicity in communication?

Why do we not find language-like systems in other animals, despite their latent social and cognitive capacities? Could this be related to their ability (or lack thereof) to ostensively and spontaneously recruit multimodal semiotic resources?

LLMs succeed in performing a variety of tasks only from being trained on (written) language. Can the study of LLMs help us understand how exposure to different aspects of language works to entrain human cognition?

How and which social and environmental pressures shape languages? Do these pressures operate similarly across different timescales (e.g., day-to-day interaction versus cross-generational transmission of language)?

Throughout this review, we have shown how the study of language evolution and adaptation can gain important insights from research on animal communication (Box 1). Thus, multimodality and semiotic diversity, diverse functionality, and adaptiveness represent crucial directions for comparative work on human language and nonhuman animal communication, and provide a new roadmap for future research avenues.

Acknowledgments

M.Pl. was funded by the National Science Centre, Poland, grant number 2024/53/B/HS2/01366. G.L. was partially funded by NSF-PAC #2020969. We are very thankful to Bodo Winter for his meaningful contributions to the early stages of conceptualizing the paper, and for initiating this project. We thank Aslı Özyürek and Cedric Boeckx, as well as the anonymous reviewers, for their helpful comments and feedback.

Declaration of interests

None declared by authors.

Resources

[i]https://figshare.com/articles/online_resource/O5S5_Documenting_the_experiences_of_the_ASL_Communities_in_the_time_of_COVID-19/16983517

[ii]https://maximumeffort.substack.com/p/i-taught-chatgpt-to-invent-a-language

[iii]www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html

References

1. Hockett, C.F. (1960) The origin of speech. *Sci. Am.* 203, 88–97
2. Hockett, C.F. (1963) The problem of universals in language. *Univers. Lang.* 2, 1–29
3. Wacewicz, S. *et al.* (2022) The representation of animal communication and language evolution in introductory linguistics textbooks. *J. Lang. Evol.* 7, 147–165
4. Özyürek, A. (2021) Considering the nature of multimodal language from a crosslinguistic perspective. *J. Cogn.* 4, 42
5. Stokoe, W.C., Jr. (2005) Sign language structure: an outline of the visual communication systems of the American deaf. *J. Deaf. Stud. Deaf. Educ.* 10, 3–37
6. Sandler, W. *et al.* (2011) The gradual emergence of phonological form in a new language. *Nat. Lang. Linguist. Theory* 29, 503–543
7. Dingemanse, M. *et al.* (2015) Arbitrariness, iconicity, and systematicity in language. *Trends Cogn. Sci.* 19, 603–615
8. Perniss, P. *et al.* (2010) Iconicity as a general property of language: evidence from spoken and signed languages. *Front. Psychol.* 1, 227
9. Pleyer, M. and Hartmann, S. (2024) *Cognitive Linguistics and Language Evolution*, Cambridge University Press
10. Berthet, M. *et al.* (2025) Extensive compositionality in the vocal system of bonobos. *Science* 388, 104–108
11. Engesser, S. and Townsend, S.W. (2019) Combinatoriality in the vocal systems of nonhuman animals. *WIREs Cogn. Sci.* 10, e1493
12. Suzuki, T.N. *et al.* (2016) Experimental evidence for compositional syntax in bird calls. *Nat. Commun.* 7, 10986
13. Bohn, M. *et al.* (2015) Communication about absent entities in great apes and human infants. *Cognition* 145, 63–72
14. von Frisch, K. (1993) *The Dance Language and Orientation of Bees*, Harvard University Press
15. Hölldobler, B. and Wilson, E.O. (1978) The multiple recruitment systems of the African weaver ant *Oecophylla longinoda* (Latreille) (Hymenoptera: Formicidae). *Behav. Ecol. Sociobiol.* 3, 19–60
16. Lyn, H. *et al.* (2014) Apes communicate about absent and displaced objects: methodology matters. *Anim. Cogn.* 17, 85–94
17. Wacewicz, S. and Żywiczyński, P. (2015) Language evolution: why Hockett's design features are a non-starter. *Biosemiotics* 8, 29–46
18. Beckner, C. *et al.* (2009) Language is a complex adaptive system: position paper. *Lang. Learn.* 59, 1–26
19. Kirby, S. *et al.* (2015) Compression and communication in the cultural evolution of linguistic structure. *Cognition* 141, 87–102
20. Motamedi, Y. *et al.* (2019) Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition* 192, 103964
21. Raviv, L. *et al.* (2019) Compositional structure can emerge without generational transmission. *Cognition* 182, 151–164
22. Heintz, C. and Scott-Phillips, T. (2023) Expression unleashed: the evolutionary and cognitive foundations of human communication. *Behav. Brain Sci.* 46, e1
23. Levinson, S.C. (2025) *The Interaction Engine: Language in Social Life and Human Evolution*, Cambridge University Press
24. Cohn, N. and Schilperoord, J. (2024) *A Multimodal Language Faculty: A Cognitive Framework for Human Communication*, Bloomsbury Publishing
25. Ferrara, L. and Hodge, G. (2018) Language as description, indication, and depiction. *Front. Psychol.* 9, 716
26. Kendon, A. (2014) Semiotic diversity in utterance production and the concept of 'language'. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130293
27. McNeill, D. (1992) *Hand and Mind: What Gestures Reveal about Thought*, University of Chicago Press
28. Emmorey, K. (2021) New perspectives on the neurobiology of sign languages. *Front. Commun.* 6, 748430
29. Mudd, K. *et al.* (2020) An agent-based model of sign language persistence informed by real-world data. *Lang. Dyn. Change* 10, 158–187
30. Senghas, A. *et al.* (2004) Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* 305, 1779–1782
31. Edwards, T. and Brentari, D. (2020) Feeling phonology: the conventionalization of phonology in Protactile communities in the United States. *Language* 96, 819–840
32. Danesi, M. (2017) *The Semiotics of Emoji: The Rise of Visual Language in the Age of the Internet*, Bloomsbury Publishing
33. Holler, J. and Levinson, S.C. (2019) Multimodal language processing in human communication. *Trends Cogn. Sci.* 23, 639–652
34. Kendon, A. (2004) *Gesture: Visible Action as Utterance*, Cambridge University Press
35. Krahmer, E. and Swerts, M. (2009) Audiovisual prosody—introduction to the special issue. *Lang. Speech* 52, 129–133

36. Holler, J. *et al.* (2009) Do iconic hand gestures really contribute to the communication of semantic information in a face-to-face context? *J. Nonverbal Behav.* 33, 73–88
37. Bosker, H.R. and Peeters, D. (2021) Beat gestures influence which speech sounds you hear. *Proc. R. Soc. B Biol. Sci.* 288, 20202419
38. Mcgurk, H. and Macdonald, J. (1976) Hearing lips and seeing voices. *Nature* 264, 746–748
39. Levinson, S.C. and Holler, J. (2014) The origin of human multi-modal communication. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130302
40. Laing, C.E. (2014) A phonological analysis of onomatopoeia in early word production. *First Lang.* 34, 387–405
41. Motamedi, Y. *et al.* (2024) Language development beyond the here-and-now: Iconicity and displacement in child-directed communication. *Child Dev.* 95, 1539–1557
42. Perniss, P. *et al.* (2018) Mapping language to the world: the role of iconicity in the sign language input. *Dev. Sci.* 21, e12551
43. Armstrong, D.F. and Wilcox, S. (2007) *The Gestural Origin of Language*, Oxford University Press
44. Imai, M. and Kita, S. (2014) The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130298
45. Taub, S.F. (2001) *Language from the Body: Iconicity and Metaphor in American Sign Language*, Cambridge University Press
46. Liddell, S.K. (2008) *Grammar, Gesture, and Meaning in American Sign Language*, Cambridge University Press
47. Cormier, K. *et al.* (2015) Rethinking constructed action. *Sign Lang. Linguist.* 18, 167–204
48. Fuks, O. (2014) Gradient and categorically: handshape's two semiotic dimensions in Israeli Sign Language discourse. *J. Pragmat.* 60, 207–225
49. Ferrara, L. and Halvorsen, R.P. (2017) Depicting and describing meanings with iconic signs in Norwegian Sign Language. *Gesture* 16, 371–395
50. Wilcox, S. (2004) Cognitive iconicity: conceptual spaces, meaning, and gesture in signed languages. *Cogn. Linguist.* 15, 119–147
51. Vigliocco, G. *et al.* (2014) Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130292
52. Cienki, A. and Müller, C. (2008) *Metaphor and Gesture*, John Benjamins Publishing Company
53. Dingemanse, M. (2012) Advances in the cross-linguistic study of ideophones. *Lang. Linguist. Compass* 6, 654–672
54. Blasi, D.E. *et al.* (2016) Sound–meaning association biases evidenced across thousands of languages. *Proc. Natl. Acad. Sci. U. S. A.* 113, 10818–10823
55. Johansson, N.E. *et al.* (2020) The typology of sound symbolism: defining macro-concepts via their semantic and phonetic features. *Linguist. Typol.* 24, 253–310
56. Winter, B. and Perlman, M. (2021) Size sound symbolism in the English lexicon. *Glossa J. Gen. Linguist.* 6, 79
57. Winter, B. *et al.* (2022) Trilled /r/ is associated with roughness, linking sound and touch across spoken languages. *Sci. Rep.* 12, 1035
58. Haiman, J. (1980) The iconicity of grammar: isomorphism and motivation. *Language* 56, 515–540
59. Perlman, M. *et al.* (2015) Iconic prosody in story reading. *Cogn. Sci.* 39, 1348–1368
60. Frank, M.C. (2023) Bridging the data gap between children and large language models. *Trends Cogn. Sci.* 27, 990–992
61. Zhang, E.Q. and Pleyer, M. (2024) Toward interdisciplinary integration in the study of comparative cognition: insights from studying the evolution of multimodal communication. *Comp. Cogn. Behav. Rev.* 19, 85–90
62. Eleuteri, V. *et al.* (2024) Multimodal communication and audience directedness in the greeting behaviour of semi-captive African savannah elephants. *Commun. Biol.* 7, 1–12
63. Graham, K.E. (2024) Goal-directed bodily signals in birds and frogs. *Learn. Behav.* 53, 5–6
64. Higham, J.P. and Hebets, E.A. (2013) An introduction to multimodal communication. *Behav. Ecol. Sociobiol.* 67, 1381–1388
65. Liebal, K. *et al.* (2014) *Primate Communication: A Multimodal Approach*, Cambridge University Press
66. Tomasello, M. and Zuberbühler, K. (2002) Primate vocal and gestural communication. In *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition* (Bekoff, M. *et al.*, eds), pp. 293–299, MIT Press
67. Waller, B.M. *et al.* (2024) The face is central to primate multicomponent signals. *Int. J. Primatol.* 45, 526–542
68. Taylor, R.C. *et al.* (2011) Multimodal signal variation in space and time: how important is matching a signal with its signaler? *J. Exp. Biol.* 214, 815–820
69. Hölldobler, B. (1999) Multimodal signals in ant communication. *J. Comp. Physiol. A.* 184, 129–141
70. Leavens, D.A. *et al.* (2004) Tactical use of unimodal and bimodal communication by chimpanzees, *Pan troglodytes*. *Anim. Behav.* 67, 467–476
71. Gordon, S.D. and Uetz, G.W. (2011) Multimodal communication of wolf spiders on different substrates: evidence for behavioural plasticity. *Anim. Behav.* 81, 367–375
72. von Frisch, K. (1954) *The Dancing Bees: An Account of the Life and Senses of the Honey Bee*, Springer
73. Bradbury, J.W. and Vehrencamp, S.L. (2011) *Principles of Animal Communication* (2nd edn), Sinauer Associates
74. Pinker, S. and Bloom, P. (1990) Natural language and natural selection. *Behav. Brain Sci.* 13, 707–727
75. Fedorenko, E. *et al.* (2024) Language is primarily a tool for communication rather than thought. *Nature* 630, 575–586
76. Reddy, M.J. (1979) The conduit metaphor – a case of frame conflict in our language about language. In *Metaphor and Thought* (Ortony, A., ed.), pp. 284–297, Cambridge University Press
77. Shannon, C.E. (1948) A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423 623–656
78. Gibson, E. *et al.* (2019) How efficiency shapes human language. *Trends Cogn. Sci.* 23, 389–407
79. Kemp, C. *et al.* (2018) Semantic typology and efficient communication. *Annu. Rev. Linguist.* 4, 109–128
80. Grice, P. (1989) *Studies in the Way of Words*, Harvard University Press
81. Scott-Phillips, T.C. (2015) *Speaking our Minds: Why Human Communication is Different, and How Language Evolved to Make it Special*, Palgrave Macmillan
82. Sperber, D. and Wilson, D. (1996) *Relevance: Communication and Cognition* (2nd edn), Wiley-Blackwell
83. Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*, University of Chicago Press
84. Vicente, A. and Falkum, I.L. (2017) Polysemy. In *Oxford Research Encyclopedia of Linguistics* (Aronoff, M., ed.), Oxford University Press
85. Degen, J. (2023) The rational speech act framework. *Annu. Rev. Linguist.* 9, 519–540
86. Goodman, N.D. and Frank, M.C. (2016) Pragmatic language interpretation as probabilistic inference. *Trends Cogn. Sci.* 20, 818–829
87. Seyfarth, R.M. and Cheney, D.L. (2017) The origin of meaning in animal signals. *Anim. Behav.* 124, 339–346
88. Arnold, K. and Zuberbühler, K. (2013) Female putty-nosed monkeys use experimentally altered contextual information to disambiguate the cause of male alarm calls. *PLoS ONE* 8, e65660
89. Deshpande, A. *et al.* (2023) Context-dependent alarm responses in wild vervet monkeys. *Anim. Cogn.* 26, 1199–1208
90. Hobaiter, C. *et al.* (2022) Are ape gestures like words? Outstanding issues in detecting similarities and differences between human language and ape gesture. *Philos. Trans. R. Soc. B Biol. Sci.* 377, 20210301
91. Clark, H.H. (1996) *Using Language*, Cambridge University Press
92. Scott-Phillips, T. and Heintz, C. (2023) Animal communication in linguistic and cognitive perspective. *Annu. Rev. Linguist.* 9, 93–111
93. Tomasello, M. (2008) *Origins of Human Communication*, MIT press
94. Bonard, C. (2024) Underdeterminacy without ostension: a blind spot in the prevailing models of communication. *Mind Lang.* 39, 142–161
95. Gouzoules, H. *et al.* (2023) Loosening the leash: the unique emotional canvas of human screams. *Behav. Brain Sci.* 46, e10

96. Wardhaugh, R. and Fuller, J.M. (2021) *An Introduction to Sociolinguistics*, John Wiley & Sons
97. Eckert, P. (2019) The limits of meaning: social indexicality, variation, and the cline of interiority. *Language* 95, 751–776
98. Padilla-Iglesias, C. *et al.* (2020) Language as a marker of ethnic identity among the Yucatec Maya. *Evol. Hum. Sci.* 2, e38
99. Kinzler, K.D. *et al.* (2007) The native language of social cognition. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12577–12580
100. Kinzler, K.D. *et al.* (2009) Accent trumps race in guiding children's social preferences. *Soc. Cogn.* 27, 623–634
101. Roessel, J. *et al.* (2020) Modern notions of accent-ism: findings, conceptualizations, and implications for interventions and research on nonnative accents. *J. Lang. Soc. Psychol.* 39, 87–111
102. Eckert, P. and McConnell-Ginet, S. (1992) Think practically and look locally: language and gender as community-based practice. *Annu. Rev. Anthropol.* 21, 461–488
103. Hall-Lew, L. *et al.,* eds (2021) *Social Meaning and Linguistic Variation: Theorizing the Third Wave*, Cambridge University Press
104. Giles, H. and Ogay, T. (2007) Communication accommodation theory. In *Explaining Communication: Contemporary Theories and Exemplars* (Whaley, B.B. and Samter, W., eds), pp. 293–310, Routledge
105. Bell, A. (1984) Language style as audience design. *Lang. Soc.* 13, 145–204
106. Austin, J.L. (1962) *How To Do Things With Words*, Harvard University Press
107. Tyack, P.L. (2008) Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *J. Comp. Psychol.* 122, 319–331
108. Hile, A.G. and Striedter, G.F. (2000) Call convergence within groups of female budgerigars (*Melopsittacus undulatus*). *Ethology* 106, 1105–1114
109. Boughman, J.W. (1998) Vocal learning by greater spear-nosed bats. *Proc. R. Soc. Lond. B Biol. Sci.* 265, 227–233
110. Poole, J.H. *et al.* (2005) Elephants are capable of vocal learning. *Nature* 434, 455–456
111. Smolker, R. and Pepper, J.W. (1999) Whistle convergence among allied male bottlenose dolphins (Delphinidae, *Tursiops* sp.). *Ethology* 105, 595–617
112. Candiotti, A. *et al.* (2012) Convergence and divergence in Diana monkey vocalizations. *Biol. Lett.* 8, 382–385
113. Fischer, J. *et al.* (2020) Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. *Proc. R. Soc. B Biol. Sci.* 287, 20202531
114. Schamberg, I. *et al.* (2024) Cross-population variation in usage of a call combination: evidence of signal usage flexibility in wild bonobos. *Anim. Cogn.* 27, 58
115. Oren, G. *et al.* (2024) Vocal labeling of others by nonhuman primates. *Science* 385, 996–1003
116. Pardo, M.A. *et al.* (2024) African elephants address one another with individually specific name-like calls. *Nat. Ecol. Evol.* 8, 1353–1364
117. Reby, D. and McComb, K. (2003) Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Anim. Behav.* 65, 519–530
118. Vignal, C. and Kelley, D. (2007) Significance of temporal and spectral acoustic cues for sexual recognition in Xenopus laevis. *Proc. R. Soc. B Biol. Sci.* 274, 479–488
119. Garcia, M. *et al.* (2016) Honest signaling in domestic piglets (*Sus scrofa* domesticus): vocal allometry and the information content of grunt calls. *J. Exp. Biol.* 219, 1913–1921
120. Mathevon, N. *et al.* (2017) Northern elephant seals memorize the rhythm and timbre of their rivals' voices. *Curr. Biol.* 27, 2352–2356
121. Frank, M.C. *et al.* (2008) Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition* 108, 819–824
122. Lupyan, G. and Zettersten, M. (2021) Does vocabulary help structure the mind? In *Minnesota Symposia on Child Psychology* (Sera, M.D. and Koenig, M.A., eds), pp. 160–199, John Wiley & Sons
123. Waxman, S.R. and Markow, D.B. (1995) Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cogn. Psychol.* 29, 257–302
124. Althaus, N. and Mareschal, D. (2014) Labels direct infants' attention to commonalities during novel category learning. *PLoS ONE* 9, e99670
125. Casasola, M. (2005) Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Dev. Psychol.* 41, 183–192
126. Nazzi, T. and Gopnik, A. (2001) Linguistic and cognitive abilities in infancy: when does language become a tool for categorization? *Cognition* 80, B11–B20
127. Loewenstein, J. and Gentner, D. (2005) Relational language and the development of relational mapping. *Cogn. Psychol.* 50, 315–353
128. Bloom, P. (2002) *How Children Learn the Meanings of Words*, MIT Press
129. Wolff, P. and Holmes, K. (2011) Linguistic relativity. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 253–265
130. Lupyan, G. *et al.* (2007) Language is not just for talking: labels facilitate learning of novel categories. *Psychol. Sci.* 18, 1077–1082
131. Vanek, N. *et al.* (2021) Consistent verbal labels promote odor category learning. *Cognition* 206, 104485
132. Lupyan, G. (2008) From chair to 'chair': a representational shift account of object labeling effects on memory. *J. Exp. Psychol. Gen.* 137, 348–369
133. Suffill, E. *et al.* (2024) Mind melds: verbal labels induce greater representational alignment. *Open Mind* 8, 950–971
134. Lupyan, G. (2009) Extracommunicative functions of language: verbal interference causes selective categorization impairments. *Psychon. Bull. Rev.* 16, 711–718
135. Nedergaard, J.S.K. *et al.* (2022) Verbal interference paradigms: a systematic review investigating the role of language in cognition. *Psychon. Bull. Rev.* 30, 464–488
136. Pylyshyn, Z. (1999) Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav. Brain Sci.* 22, 341–365
137. Samaha, J. *et al.* (2018) Effects of meaningfulness on perception: alpha-band oscillations carry perceptual expectations and influence early visual responses. *Sci. Rep.* 8, 6606
138. Forder, L. and Lupyan, G. (2019) Hearing words changes color perception: facilitation of color discrimination by verbal and visual cues. *J. Exp. Psychol. Gen.* 148, 1105–1123
139. Lupyan, G. and Ward, E.J. (2013) Language can boost otherwise unseen objects into visual awareness. *Proc. Natl. Acad. Sci. U. S. A.* 110, 14196–14201
140. Lupyan, G. *et al.* (2020) Effects of language on visual perception. *Trends Cogn. Sci.* 24, 930–944
141. Edmiston, P. and Lupyan, G. (2015) What makes words special? Words as unmotivated cues. *Cognition* 143, 93–100
142. Pepperberg, I.M. (2013) Abstract concepts: data from a grey parrot. *Behav. Process.* 93, 82–90
143. Savage-Rumbaugh, S. *et al.* (2001) *Apes, Language, and the Human Mind*, Oxford University Press
144. Brighton, H. *et al.* (2005) Language as an evolutionary system. *Phys Life Rev* 2, 177–226
145. Kirby, S. (2017) Culture and biology in the origins of linguistic structure. *Psychon. Bull. Rev.* 24, 118–137
146. Steels, L. (2017) Human language is a culturally evolving system. *Psychon. Bull. Rev.* 24, 190–193
147. Christiansen, M.H. and Chater, N. (2022) *The Language Game: How Improvisation Created Language and Changed the World*, Penguin Press
148. Galantucci, B. and Garrod, S. (2012) *Experimental Semiotics: Studies on the Emergence and Evolution of Human Communication*, John Benjamins Publishing
149. Lister, C.J. and Fay, N. (2017) How to create a human communication system. *Interact. Stud.* 18, 314–329
150. Müller, T.F. and Raviv, L. (2024) Communication games: social interaction in the formation of novel communication systems. In *Oxford Handbook to Methods in Language Evolution* (Raviv, L. and Boeckx, C., eds), pp. 41–62, Oxford University Press
151. Nölle, J. *et al.* (2018) The emergence of systematicity: how environmental and communicative factors shape a novel communication system. *Cognition* 181, 93–104
152. Nölle, J. and Galantucci, B. (2023) Experimental semiotics: past, present, and future. In *The Routledge Handbook of Semiosis*

and the Brain (García, A.M. and Ibanez, A., eds), pp. 66–81, Routledge
153. Du Bois, J.W. (2014) Towards a dialogic syntax. *Cogn. Linguist.* 25, 359–410
154. Brochhagen, T. *et al.* (2023) From language development to language evolution: a unified view of human lexical creativity. *Science* 381, 431–436
155. Traugott, E.C. and Dasher, R.B. (2001) *Regularity in Semantic Change*, Cambridge University Press
156. Garrod, S. *et al.* (2007) Foundations of representation: where might graphical symbol systems come from? *Cogn. Sci.* 31, 961–987
157. Formaux, A. *et al.* (2022) The experimental emergence of convention in a non-human primate. *Philos. Trans. R. Soc. B Biol. Sci.* 377, 20200310
158. van Leeuwen, E.J.C. and Hoppitt, W. (2023) Biased cultural transmission of a social custom in chimpanzees. *Sci. Adv.* 9, eade5675
159. Saldana, C. *et al.* (2019) High-fidelity copying is not necessarily the key to cumulative cultural evolution: a study in monkeys and children. *Proc. R. Soc. B Biol. Sci.* 286, 20190729
160. Zürcher, Y. *et al.* (2019) Are dialects socially learned in marmoset monkeys? Evidence from translocation experiments. *PLoS ONE* 14, e0222486
161. Aronoff, M. *et al.* (2008) The roots of linguistic organization in a new language. *Interact. Stud.* 9, 133–153
162. Loreto, V. *et al.* (2016) On the emergence of syntactic structures: quantifying and modeling duality of patterning. *Top. Cogn. Sci.* 8, 469–480
163. Beecher, M.D. (2021) Why are no animal communication systems simple languages? *Front. Psychol.* 12, 701
164. Roberts, G. and Galantucci, B. (2012) The emergence of duality of patterning: insights from the laboratory. *Lang. Cogn.* 4, 297–318
165. Verhoef, T. *et al.* (2014) Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *J. Phon.* 43, 57–68
166. Winters, J. *et al.* (2015) Languages adapt to their contextual niche. *Lang. Cogn.* 7, 415–449
167. Zuidema, W. and de Boer, B. (2018) The evolution of combinatorial structure in language. *Curr. Opin. Behav. Sci.* 21, 138–144
168. Kirby, S. and Tamariz, M. (2021) Cumulative cultural evolution, population structure and the origin of combinatoriality in human language. *Philos. Trans. R. Soc. B Biol. Sci.* 377, 20200319
169. Raviv, L. *et al.* (2019) Larger communities create more systematic languages. *Proc. R. Soc. B Biol. Sci.* 286, 20191262
170. Lupyan, G. and Dale, R. (2010) Language structure is partly determined by social structure. *PLoS ONE* 5, e8559
171. Nettle, D. (2012) Social scale and structural complexity in human languages. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1829–1836
172. Reali, F. *et al.* (2018) Simpler grammar, larger vocabulary: How population size affects language. *Proc. R. Soc. B Biol. Sci.* 285, 20172586
173. Trudgill, P. (2002) *Sociolinguistic Variation and Change*, Georgetown University Press
174. Coye, C. *et al.* (2018) From animal communication to linguistics and back: insight from combinatorial abilities in monkeys and birds. In *Origins of Human Language: Continuities and Discontinuities with Nonhuman Primates* (Boë, L.-J. *et al.*, eds), pp. 187–232, Peter Lang
175. Townsend, S.W. *et al.* (2018) Compositionality in animals and humans. *PLoS Biol.* 16, e2006425
176. Freeberg, T.M. and Lucas, J.R. (2002) Receivers respond differently to chick-a-dee calls varying in note composition in Carolina chickadees, *Poecile carolinensis*. *Anim. Behav.* 63, 837–845
177. Suzuki, T.N. (2014) Communication about predator type by a bird using discrete, graded and combinatorial variation in alarm calls. *Anim. Behav.* 87, 59–65
178. Engesser, S. *et al.* (2016) Meaningful call combinations and compositional processing in the southern pied babbler. *Proc. Natl. Acad. Sci. U. S. A.* 113, 5976–5981
179. Ouattara, K. *et al.* (2009) Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proc. Natl. Acad. Sci. U. S. A* 106, 22026–22031
180. Arnon, I. *et al.* (2025) Whale song shows language-like statistical structure. *Science* 387, 649–653
181. Henry, M.J. *et al.* (2021) An ecological approach to measuring synchronization abilities across the animal kingdom. *Philos. Trans. R. Soc. B Biol. Sci.* 376, 20200336
182. Hyland Bruno, J. and Tchernichovski, O. (2019) Regularities in zebra finch song beyond the repeated motif. *Behav. Process.* 163, 53–59
183. Zann, R.A. (1996) *The Zebra Finch: A Synthesis of Field and Laboratory Studies*, Oxford University Press
184. Fishbein, A.R. *et al.* (2020) Sound sequences in birdsong: how much do birds really care? *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190044
185. Rutz, C. *et al.* (2023) Using machine learning to decode animal communication. *Science* 381, 152–155
186. Kokko, H. (2017) Give one species the task to come up with a theory that spans them all: what good can come out of that? *Proc. R. Soc. B Biol. Sci.* 284, 20171652
187. Wacewicz, S. *et al.* (2020) Language in language evolution research: in defense of a pluralistic view. *Biolinguistics* 14, 59–101
188. Jackendoff, R. (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford University Press
189. Pinker, S. (1998) Words and rules. *Lingua* 106, 219–242
190. Fedorenko, E. *et al.* (2024) The language network as a natural kind within the broader landscape of the human brain. *Nat. Rev. Neurosci.* 25, 289–312
191. Jewitt, C. *et al.* (2016) *Introducing Multimodality*, Routledge
192. Mesch, J. *et al.* (2015) Co-forming real space blends in tactile signed language dialogues. *Cogn. Linguist.* 26, 261–287
193. Perlman, M. *et al.* (2015) Iconicity can ground the creation of vocal symbols. *R. Soc. Open Sci.* 2, 150152
194. Perlman, M. and Lupyan, G. (2018) People can create iconic vocalizations to communicate various meanings to naïve listeners. *Sci. Rep.* 8, 2634
195. Ćwiek, A. *et al.* (2021) Novel vocalizations are understood across cultures. *Sci. Rep.* 11, 10108
196. Lyn, H. (2012) Apes and the evolution of language: taking stock of 40 years of research. In *The Oxford Handbook of Comparative Evolutionary Psychology* (Shackleford, T.K. and Vonk, J., eds), pp. 356–378, Oxford University Press
197. Köhler, W. (1925) *The Mentality of Apes*, Routledge & Kegan Paul
198. Tanner, J.E. and Byrne, R.W. (1996) Representation of action through iconic gesture in a captive lowland gorilla. *Curr. Anthropol.* 37, 162–173
199. Russon, A. and Andrews, K. (2011) Orangutan pantomime: elaborating the message. *Biol. Lett.* 7, 627–630
200. Perlman, M. *et al.* (2012) A mother gorilla's variable use of touch to guide her infant: insights into iconicity and the relationship between gesture and action. *Dev. Primate Gesture Res.* 6, 55–72
201. Tomasello, M. and Call, J. (2018) Thirty years of great ape gestures. *Anim. Cogn.* 22, 461–469
202. Byrne, R.W. *et al.* (2017) Great ape gestures: intentional communication with a rich set of innate signals. *Anim. Cogn.* 20, 755–769
203. Graham, K.E. *et al.* (2024) The origin of great ape gestural forms. *Biol. Rev.* 100, 190–204
204. Elman, J.L. (1990) Finding structure in time. *Cogn. Sci.* 14, 179–211
205. Devlin, J. *et al.* (2019) BERT: pre-training of deep bidirectional transformers for language understanding. *ArXiv* Published online October 11, 2018. http://x.doi.org/10.48550/arXiv.1810.04805
206. Boleda, G. (2020) Distributional semantics and linguistic theory. *Annu. Rev. Linguist.* 6, 213–234
207. Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204
208. Huang, Y. and Rao, R.P. (2011) Predictive coding. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 580–593
209. Boutin, V. *et al.* (2021) Sparse deep predictive coding captures contour integration capabilities of the early visual system. *PLoS Comput. Biol.* 17, e1008629

210. Spratling, M.W. (2017) A hierarchical predictive coding model of object recognition in natural images. *Cogn. Comput.* 9, 151–167
211. Mahowald, K. *et al.* (2024) Dissociating language and thought in large language models. *Trends Cogn. Sci.* 28, 517–540
212. Agüera y Arcas, B. (2022) Do large language models understand us? *Daedalus* 151, 183–197
213. Conmy, A. *et al.* (2023) Towards automated circuit discovery for mechanistic interpretability. *arXiv* Published online April 28, 2023. http://dx.doi.org/10.48550/arXiv.2304.14997
214. OpenAI *et al.* (2024) GPT-4 technical report. *arXiv* Published online March 15, 2023. http://dx.doi.org/10.48550/arXiv.2303.08774
215. Zou, A. *et al.* (2024) DOCBENCH: a benchmark for evaluating LLM-based document reading systems. *arXiv* Published online July 15, 2024. http://dx.doi.org/10.48550/arXiv.2407.10701
216. Mukherjee, S. *et al.* (2024) Are large language models actually good at text style transfer? *arXiv* Published online June 9, 2024. http://dx.doi.org/10.48550/arXiv.2406.05885
217. Cartmill, E.A. (2023) Overcoming bias in the comparison of human language and animal communication. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2218799120
218. Bender, E.M. *et al.* (2021) On the dangers of stochastic parrots: can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610–623, Association for Computing Machinery
219. Lupyan, G. (2016) The centrality of language in human cognition. *Lang. Learn.* 66, 516–553
220. Fröhlich, M. *et al.* (2025) Rethinking ambiguity across species. *Neurosci. Biobehav. Rev.* 179, 106401