

MIT Press • Open Encyclopedia of Cognitive Science

Language

Adele E. Goldberg

MIT Press

Published on: Apr 21, 2025

DOI: <https://doi.org/10.21428/e2759450.40bf93ad>

License: [Creative Commons Attribution-NonCommercial 4.0 International License \(CC-BY-NC 4.0\)](#).

Human language has long provided a crucible for new theories and methods in nearly all areas of cognitive science. Fiery debates have arisen, and critical advances have been made through investigations into how best to characterize relationships between form and function, the extent and implications of variation across languages, and the cognitive and social prerequisites for language. New methodological advances continue to inspire debates about how natural language is learned, produced, and understood and how these skills relate to what it means to be human.

History

For centuries before language was studied through the lens of cognitive science, scholars around the globe were describing and comparing languages. There was a general appreciation of the fact that each language displayed systematic regularities; historical linguists revealed systematic changes in sound patterns and word order over time, and philologists studied changes in word meanings across time. These insights facilitated the early recognition of relationships among languages, which provided evidence of historical connections between communities.

In the 1950s, experimental psychology was primarily focused on observable behaviors; the human mind was considered impossible to study scientifically because doing so would rely on untrustworthy intuitions. Yet, as computers entered the scene, an inevitable march toward trying to understand cognition computationally and scientifically began. [Chomsky's \(1959\)](#) critique of behaviorism along with [Miller's \(1956\)](#) observations regarding limits on working memory catalyzed the nascent field of cognitive science. Chomsky observed that human language cannot be based simply on behavioral responses to external stimuli. That is, he noted that that language is more than one word (a stimulus) predicting the next word (a response). Instead, human language requires at least two mental constructs: (1) working memory, because of discontinuous meaningful units (e.g., memory allows us to relate the bold bits in “**He** went for a swim **naked**”), and (2) hierarchical structure to capture the fact that meaningful units (*constituents*) of language can be subparts of larger units, which can themselves be subparts of still larger units.

Perhaps less widely remembered in this context is the fact that [Tolman \(1948\)](#), a leading behaviorist, had already documented the need for mental representations, even in rats, a decade earlier. In any case, the widespread recognition that human language requires mental representations played a key role in the emergence of cognitive science as a discipline.

A new generation of language researchers generally assumed that the required mental representations took the form of algebraic rules, presumably on analogy to the functions in computer programming or to formal logic. For instance, to capture word order generalizations, *phrase structure rules* were used (e.g., VP → V NP: a verb phrase can consist of a verb followed by a noun phrase); to capture regularities in pronunciation and sound change, phonological rules were proposed (e.g., [consonant]# → [-voice]: final

consonants are devoiced); and to capture meaning, algebraic semantic rules were adopted (e.g., If all men are human and superman is a man, then superman is human).

During roughly the same period (1950–1985), the field’s understanding of word meaning was deepening in ways that are inconsistent with the notion of categorical, algebraic rules. [Wittgenstein \(1973\)](#) demonstrated that categorical boundaries for meanings were elusive, famously making his case with the word “game.” He noted that while games commonly involve multiple people, competition, skill, and fun, none of these factors are necessary, insofar as solitaire, storytelling games, tic tac toe, and mind games are games and yet lack one or more of these features. Nor is the combination of these features sufficient for labeling something a game, insofar as English speakers generally prefer to label running competitions as “races” rather than “games,” although they involve multiple people, competition, skill, and fun. Eleanor Rosch and colleagues additionally demonstrated that members of a given culture tend to agree on which instances are good examples of categories named by words ([Rosch & Mervis, 1975](#)). For instance, board games and soccer are more prototypical games than the Rubik’s cube. In fact, nearly all commonly used words, including the word “language,” are similarly challenging to define. English uses the word “language” when referring to body language, computer languages, and love languages. Similarly, an encyclopedia section about language might focus on individual languages such as Spanish, Mandarin, and Chinese; the psychological underpinnings of language; the human capacity to learn languages; or sociocultural aspects of language [see [Psycholinguistics](#)].

The observation that word meanings are not amenable to combinations of categorical features is difficult to reconcile with an assumption that language relies on algebraic rules (for discussion, see [Lakoff, 1987](#) and [Murphy, 2004](#)). Empirical observations about the dynamic, context-dependent nature of word meaning ultimately inspired a plethora of models, including exemplar-based representations ([Elman, 2009](#); [Nosofsky, 1988](#)) and knowledge-based representations ([Gardenfors, 2004](#)). Cognitive linguists also incorporated the new insights about word meaning into their theoretical perspectives on language ([Lakoff, 1987](#); [Langacker, 1987](#); see also [Goldberg, 2019](#) and [Jackendoff, 1985](#)). Other language scientists proposed a distributional approach to meaning, which relies on words’ co-occurrences with other words ([Deerwester et al., 1990](#); [Firth, 1957](#); [Harris, 1954](#); [Lund & Burgess, 1996](#)).

Today’s large language models, based on large-scale neural network implementations of distributed semantics transformed dynamically by contexts, have been enormously successful and are arguably highly relevant to human language [see [Large Language Models](#)]. One might think that their impressive success would undermine the assumption that language (as opposed to math or formal logic) relies on abstract algebraic rules, but the debate continues.

Core concepts

What counts as human language

Although definitions are bound to fall short, it is hard to deny that they can be useful. With that in mind, the following combination of factors conveys what seems unique to human languages:

- **Communicative:** language is used to intentionally communicate messages (with others or oneself), typically for cooperative goals, occasionally to intentionally deceive.
- **Abstract:** languages can convey messages about situations that are not present, including abstract ideas about language itself.
- **Conventional:** words and grammatical constructions convey messages in ways that are roughly shared within a community, which allows for successful communication.
- **Combinatorial:** units of language can be combined such that new messages are conveyed efficiently.
- **Cultural:** language learning depends on exposure so that languages implicitly convey information about the producers' cultural experiences.
- **Dynamic:** language interpretation is context dependent, and languages change over time.

Categories of language

Linguistic categories, like concepts and word meanings, are extremely challenging to define. For example, the English words “happy,” “big,” “yellow,” “round,” “loud,” and “dizzy” are all considered *adjectives*, but a definition of adjective is elusive. Prototypical adjectives meaningfully modify a noun (a happy child is a child who is happy), may appear prenominal (a happy child), and are able to appear after the verb “seem” (seemed happy). Yet there are exceptions to each of these criteria or “tests” for status as an English adjective. For instance, a subclass of adjectives resists appearing before nouns (??An asleep/alive boy vs. the boy seemed asleep/alive)¹, other adjectives are incompatible with “seem” (??The idiot seems blithering vs. the blithering idiot), and adjectives do not always semantically modify the noun (a happy painting is not a painting that is happy). The category of English adjectives requires a more nuanced, multifaceted characterization (e.g., [Goldberg, 2006](#)).

When one ventures beyond a single language such as English, the issues quickly multiply [see Linguistic Universals and Linguistic Diversity]. For instance, terms that translate into adjectives in English can be conveyed by terms that share a distribution with prototypical nouns (similar to “yellowness”) or verbs (cf. “yellowed”). Similar challenges with strict definitions arise with all linguistic categories (e.g., “subject,” “syllable,” “phoneme,” “construction”). One approach is to define *concepts* that can be usefully compared across languages ([Haspelmath, 2010](#)), although as noted, concepts, like formal categories, resist tidy definitions. Another approach is to *operationalize* terminology by using group consensus, or a particular set of features, or by stipulation, for use in particular contexts.

Lexicon

The *lexicon* is widely presumed to include words or word templates in a language. Research by phonologists, morphologists, and lexical semanticists has long made clear that each lexicon boasts a rich

tapestry of general and not-so-general phonological and morphological patterns, individual words laden with cultural meanings, borrowings, and historical artifacts. *Psycholinguists*, focusing on how people manage to retrieve and comprehend lexical units presented in a rapid-fire continuous stream, have found that ease of accessibility depends on content, context, related clusters, priming, frequency, emotional valence, and age of acquisition ([Hay & Baayen, 2005](#)).

Constructions

Most linguists today grant that the traditional lexicon needs to be expanded to include collocations and idioms, and many argue that the lexicon should be subsumed under an even more general network of constructions that includes words, partially filled words (aka morphemes), and larger patterns. Constructions can be defined as learned pairings of form and function at various levels of complexity and abstraction ([Goldberg, 2006](#)). The more general abstract patterns provide each language with a degree of compositionality by constraining the meaning and discourse function of various formal patterns. Constructions can clarify who did what to whom as well as which parts of a message are at issue and which are backgrounded ([Lambrecht, 1994](#)). Constructions exist for asking questions, issuing commands, and expressing excitement and uncertainty. The particular ways each language carves up the range of functions and the formal patterns used to do so vary in their particulars ([Croft, 2001](#)). At the same time, certain strategies for expressing particular functions tend to recur across unrelated languages ([Haspelmath, 2023](#)) [see [Linguistic Variation](#)].

Communicative pressures on language

All human languages are subject to the well-known conflicting demands of efficiency and transparency (e.g., [Haiman, 1985](#)). On the one hand, utterances should be reasonably efficient for the speaker or signer ([Gibson et al., 2019](#); [Hawkins, 2004](#); [Jespersen, 1922](#)); on the other hand, utterances need to be reasonably transparent so that they are interpretable ([Bybee, 1995](#)). If efficiency were maximized, languages might consist of a single sound /ah/, but such a language would be useless for sharing messages. If transparency were maximized, languages would look far more regular than they do ([Sapir, 1921](#)). The fact that these two factors regularly conflict sheds light on why exceptions, as well as generalizations, exist. For example, regular morphological forms are composed of meaningful parts, which makes them transparent, as in the case of the English past tense suffix /-d/. Efficiency encourages highly frequent past tense forms to be reduced (e.g., “lighted” > “lit”). The consonant cluster /-lkd/ does not occur in simple English words, and yet it is tolerated in morphologically complex terms (e.g., “walked” and “talked”) that make the forms transparent. Yet exceptions that follow the phonotactic generalization are presumably more efficient to pronounce, and this also supports irregular forms of highly frequent verbs (e.g., “made” rather than the hypothetical “maked”; [Burzio, 2002](#)).

In addition to the well-known factors of efficiency and transparency are affiliative factors of politeness and humor. Politeness gives rise to indirect expressions that may be longer than required or creative in a way that may risk miscommunication (e.g., “How would you feel about taking out the trash?”; [Brown &](#)

[Levinson, 1987](#)). Avoidance of a pithy but impolite term (e.g., “dumb”) may have led to the coinage of a less efficient expression: “Not the sharpest tool in the shed.” Finally, the desire to entertain and avoid tedium can result in new innovations (e.g., “Not the brightest crayon in the pack”; [Sanchez-Stockhammer & Uhrig, 2023](#)).

The vast majority of constructions are neither strictly predictable nor completely arbitrary; they are instead sensical or *motivated* solutions to the conflicting functional demands just described as well as the rich meanings expressed by many constructions ([Fillmore, 1989](#); [Lakoff, 1987](#); [Saussure, 1916](#)). Consider the ways various languages have coined, calqued, or borrowed terms that convey the meaning “bookworm.” This meaning presupposes a rich background frame of cultural knowledge: the concept takes for granted the idea that books can be read for a combination of information and pleasure and that some people enjoy reading books more than others. The associations of the term vary depending on the language, the context, and the user. For some people, “bookworm” is pejorative, implying a lack of engagement with the real world; others consider being a bookworm a sign of intelligence. Obviously, not every language has a conventional term for this meaning. The existence of a conventional term requires a meaning to be relevant in a given culture, and many linguistic communities do not read or have access to books. But among languages that do have a conventional term, the form of “bookworm” is not arbitrary. Each language’s term relates to world knowledge in straightforward ways: some languages characterize a voracious reader as a creature that metaphorically eats books (“book worm/moth/flea/eater”; echoed in the English phrase “voracious reader”); other terms imply that readers are difficult to dislodge from the vicinity of books (“library rat/mouse”) or difficult to move in general (“donkey reader”).

The variety of transparent ways to encode “bookworm” entails that no single one is predictable. For instance, English could have used “book eater” or “book flea” instead of “bookworm.” Rather than being completely predictable or completely arbitrary, the forms used in different languages are *motivated* ([Figure 1](#)) [see [Iconicity](#)].



Figure 1

Linguistic forms fall on a continuum from completely arbitrary to completely predictable; most are to some extent motivated by cultural and communicative pressures.

Explanations of specific phenomena often come from an understanding of historical processes, processing constraints, general functional factors (efficiency, transparency, politeness, humor), and the functions of the constructions involved [see [Sentence Processing](#)].

Questions, controversies, and new developments

Does language affect thought?

Languages differ in whether they express certain potential concepts and distinctions or whether they express them optionally or obligatorily. Does language influence thought? Undoubtedly, linguistic content influences thought in the moment and also may influence beliefs and long-term memory. A deeper question is whether the grammatical and lexical properties of a language inherently constrain or facilitate certain aspects of nonlinguistic cognition. It is clear that learning one language does not prevent humans from learning another one, which ensures that whatever impact language has on cognition is not determinant and irreversible. Many go further and argue that language can have no impact on nonlinguistic cognition whatsoever ([Fodor, 1975](#); [Li & Gleitman, 2002](#)). However, this is a difficult position to maintain, at least in the face of certain situations. For example, languages such as Tzeltal routinely express locations in terms of cardinal directions (e.g., north, south, east, west), rather than relative locations or landmarks as is common in English (e.g., left, right, front). While Tzeltal speakers keep track of absolute direction to avoid misunderstandings, speakers of languages like English are commonly unable to report cardinal directions accurately (see [Majid et al., 2004](#)). Another example comes from number terms: certain languages such as Pirahã lack exact terms for numbers, and the lack of exact numbers has implications for speakers' understanding of tasks that rely on counting ([Frank et al., 2008](#)). While these examples suggest that there are some influences of language on thought, the boundaries of these influences are still a matter of active debate.

A shift in emphasis from computation to memory

There is little controversy that language requires both memory and computation; humans have vast memory for collocations and idiomatic combinations of words. At the same time, most sentences longer than 6 to 7 words tend to be novel and so require at least some sort of computation. Whereas early on, researchers took pains to minimize reliance on memory in their models of language processing, the 21st century is witnessing an unmistakable shift in emphasis toward memory-rich representations. For instance, all agree that exceptional past tense forms of verbs in English such as “made” and “ran” need to be represented in memory, but many had argued that regular cases such as “walked,” “roamed,” and “catapulted” were instead computed on the fly ([Pinker, 2000](#)). For some, the English past tense suffix was taken to be a parade example of a rule that computes new instances as needed, since it applies to new verbs regardless of their meaning or form. That is, English speakers need not be familiar with verbs like “diagonalize” or “egregiate” to know that their past tenses are “diagonalized” and “egregiated.” The idea that regular forms were not retained in memory seemed to be supported by reports that regulars were uninfluenced by their frequencies of occurrence or by similarity to one another, unlike irregular cases, which were recognized, at least by many, to cluster in associative memory ([Pinker, 2000](#)).

On the other side of this debate were proponents of what was then called *connectionism*, the framework that paved the way for today's large language models. Connectionists argued that the processing of

regular and irregular patterns did not require separate modular systems but instead operated in parallel—across distributed networks within multilayer networks of nodes, which were loosely inspired by biological neural networks ([Elman et al., 1996](#); [Rumelhart et al., 1986](#); [Seidenberg & McClelland, 1989](#)). Proponents of what are now referred to as *neural network* models observed that the distinction between regular patterns and exceptions was not clear cut. So-called “exceptions” often constitute their own subregularities. For instance, people tend to creatively extend an irregular past tense pattern (e.g., “ring-rung,” “sing-sung”) to nonsense verbs, assigning “spling” the past tense form “splung” ([Bybee & Moder, 1983](#)). Moreover, many irregular past tense forms (e.g., “cut,” “hit,” “made,” “bent”) end in a dental phoneme (/t, d/), just as regular cases do (e.g., “walked,” “blinked,” “shushed”), and in fact, children appear to assume t/d final verbs like “bite” are past tense forms ([Bybee & Slobin, 1982](#)). This type of similarity between rule-following cases and “exceptions” is not easily captured if rules and exceptions are generated by separate systems ([Burzio, 2002](#); [Bybee, 1995](#)). In addition, frequency and similarity have been found to influence regular cases as well as irregulars, which is unexpected if regular cases were formed on the fly by abstract rules ([Albright & Hayes, 2003](#); [Alegre & Gordon, 1999](#); [Bertram et al., 2000](#); [Dąbrowska, 2004](#); [Kapatsinski, 2018](#); [Ramscar et al., 2013](#)). And in fact, even proponents of rules recognized that memory was required for examples like “blinked” and “glided” (also “tweeted”), since they block productive extensions of irregulars (“?blunk,” “?glid,” “?twote”); [Pinker & Ullman, 2002](#)).

Universal Grammar hypothesis or social cognitive prerequisites

Given the uniqueness of human-like language within the animal kingdom, two competing hypotheses present themselves, spawning another fiery debate: either there is some sort of essential biological endowment that is specific to language and unique to humans, or humans enjoy a suite of cognitive social prerequisites that allow languages to emerge as solutions to the human inclination to communicate. Those who adopted the former perspective promoted the idea that humans must biologically inherit an unlearned or “innate” universal grammar.

As the name indicates, the *Universal Grammar hypothesis* claimed the existence of domain-specific syntactic principles involving four interrelated claims:

1. Domain specificity: language acquisition is constrained by representations, principles, or cognitive mechanisms that are specific to language.
2. Universality: these representations, principles, or mechanisms are universal.
3. Innateness: these representations, principles, or mechanisms are not learned.

AND

4. Autonomous syntax: these representations, principles, or mechanisms depend on formal representations and are not explicable in terms of wholly functional correlates of any kind.

How might the uniqueness of human language be explained without appeal to the Universal Grammar hypothesis? Language scientists who embraced the complexity of word meanings and neural network

models argued for the perspective that there is a constellation of *prerequisites for human language* that facilitate the emergence of language within communities (see [Elman et al., 1996](#); [Knight & Lewis, 2017](#); [Tomasello, 2008](#)). Importantly, some of these prerequisites appear to be uniquely human. Chimps do not share humans' predilection to learn conventions for conventions' sake (see "over-imitation" in [Horner & Whiten, 2005](#)) nor do they readily understand symbols that refer to nonpresent entities (except perhaps those related to food; [Deacon, 1998](#)). While nonhuman primates are able to learn and use intentional signals to garner attention or elicit behavior from others, only humans are inspired by the cooperative goal of sharing information on the basis of common ground ([Tomasello, 2008](#)). Infants regularly begin to point, a gesture clearly motivated by a desire to share information and joint attention with other people ([Tomasello, 2008](#)), soon before they produce their first words. In contrast, chimps are flummoxed by referential pointing gestures ([Hare & Tomasello, 2005](#)), and such gestures are rarely if ever witnessed among primates in the wild ([Tomasello, 2008](#); [Wilke et al., 2022](#)). [Christiansen and Chater \(2022\)](#) observe that human language is akin to a game of charades: people learn to communicate ideas on the basis of conventionalized gestures, words, and grammatical constructions in combination with shared understanding of the context. This perspective recognizes language as a cultural construct, a conventional system that is learned by members of a group.

Broader connections

Language is connected to virtually all aspects of human cognition, social interactions, and community [see [Sign Language](#); [Developmental Language Disorder](#)].

Human language is a unique skill that affords a window into early learning, human cognition and memory, group-level interactions, and dynamic changes at time scales that vary from milliseconds to thousands of years. The understanding of language continues to evolve, on the basis of both natural conversational data, tightly controlled experiments, and comparative work. The field is making steady progress toward better understanding where and how brains process language, how children learn the language(s) they are exposed to, and how language skills are impacted by age, injury, and individual differences. However, more documentation work is needed because the absolute number of languages with copious amount of written text and readily available participants is dwarfed by thousands of understudied languages that exist around the globe, often without a writing system.

It is nearly impossible to think deeply about language in the current decade without considering the stunning advances being made in machine learning. Recall human language is communicative, combinatorial, abstract, conventional, cultural, and dynamic. While humans are widely believed to be unique among animals in learning and using a system with these properties, it is no longer clear that humans are alone in being able to produce and comprehend human languages. Today's generative pretrained transformer models are able to produce sensible (if not uniformly accurate) responses to prompts about topics that are real or fictional, concrete or abstract. They learn and use words and constructions in ways that are conventional, while also producing novel combinations of words and

constructions to convey new messages not included in their training data ([Goldberg, 2024](#); [Hu et al., 2024](#); [Misra & Mahowald, 2024](#); [Piantadosi, 2023](#)). Large language models (LMs) are clearly dynamic in that they are highly context sensitive in the short term and are also influenced by their input data, which allows for longer-term changes as well.

At the same time, LMs' use of language is not the same as humans', and neither is it created or learned in the same ways ([Frank, 2023](#)). LMs reflect world knowledge and cultural biases implicit in their input and training ([Navigli et al., 2023](#)), a problem shared with humans, but they differ from humans in having no means of developing a culture or language of their own. Humans who cannot otherwise communicate with one another create new languages, which quickly evolve into fully complex human languages in the context of a community ([Meir et al., 2010](#)). LMs on the other hand, without perception, intentions, or communities, require massive amounts of text that would take an individual thousands of years to produce or witness. LMs have been trained to be helpful ([Ouyang et al., 2022](#)), which might be considered an intention, but they do not form their own goals ([Mahowald et al., 2024](#); [van Dis et al., 2023](#)). These differences from human language may well be a feature rather than a bug, since they allow humans a measure of control over the new machines developing at lightning speed. Although Pygmalion-themed books like *Galatea 2.2* ([Powers, 2004](#)) and movies like *Her* may have seemed woefully far-fetched, and the philosophy of mind may have seemed largely irrelevant to cognitive science, the relationship between human minds and LMs may well give rise to the most fiery and impactful debates in the 21st century.

Further reading

- Christiansen, M. H., & Chater, N. (2022). *The language game: How improvisation created language and changed the world*. Random House.
- Elman, J. L., Bates, E. A., Johnson, M. M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development* (Vol. 10). MIT Press.
- Goldberg, A. E. (2019). *Explain me this*. Princeton University Press.
- Tomasello, M. (2008). *Origins of human communication*. MIT Press.

Footnotes

1. A ? before an example sentence indicates that it sounds odd (less *acceptable*) to native speakers, with ?? denoting sentences that sound quite odd (*unacceptable*). ↵

References

- Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, 90(2), 119-161. [https://doi.org/10.1016/S0010-0277\(03\)00146-X](https://doi.org/10.1016/S0010-0277(03)00146-X)

↵

- Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40(1), 41–61. <https://doi.org/10.1006/jmla.1998.2607>

↩

- Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: the role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(2), 489. <https://doi.org/10.1037//0278-7393.26.2.489>

↩

- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge University Press.

↩

- Burzio, L. (2002). Missing players: Phonology and the past-tense debate. *Lingua*, 112(3), 157–199. [https://doi.org/10.1016/S0024-3841\(01\)00041-9](https://doi.org/10.1016/S0024-3841(01)00041-9)

↩

- Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5), 425–455. <https://doi.org/10.1080/01690969508407111>

↩

- Bybee, J. L., & Moder, C. L. (1983). Morphological classes as natural categories. *Language*, 59(2), 251–270. <https://doi.org/10.2307/413574>

↩

- Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, 58(2), 265–289. <https://doi.org/10.2307/414099>

↩

- Chomsky, N. (1959). Review of verbal behavior. *Language*, 35(1), 26–58.

↩

- Christiansen, M. H., & Chater, N. (2022). *The language game: How improvisation created language and changed the world*. Random House.

↩

- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press.

↩

- Dąbrowska, E. (2004). *Language, mind and brain: Some psychological and neurological constraints on theories of grammar*. Edinburgh University Press.

↩

- Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. WW Norton & Company.

↩

- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.

[https://doi.org/10.1002/\(SICI\)1097-4571\(199009\)41:6<391::AID-ASII>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASII>3.0.CO;2-9)

↩

- Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, 33(4), 547-582. <https://doi.org/10.1111/j.1551-6709.2009.01023.x>

↩

- Elman, J. L., Bates, E. A., Johnson, M. M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development* (Vol. 10). MIT Press.

↩

- Fillmore, C. J. (1989). Grammatical construction theory and the familiar dichotomies. In R. Dietrich & C. F. Graumann (Eds.), *Language processing in social context* (Vol. 54, pp. 17-38). Elsevier.

↩

- Firth, J. R. (1957). Applications of general linguistics. *Transactions of the Philological Society*, 56(1), 1-14. <https://doi.org/10.1111/j.1467-968X.1957.tb00568.x>

↩

- Fodor, J. (1975). *The language of thought*. Harvard University Press.

↩

- Frank, M. C. (2023). Bridging the data gap between children and large language models. *Trends in Cognitive Sciences*, 27(11), 990-992. <https://doi.org/10.1016/j.tics.2023.08.007>

↩

- Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, 108(3), 819-824.

<https://doi.org/10.1016/j.cognition.2008.04.007>

↩

- Gardenfors, P. (2004). Conceptual spaces as a framework for knowledge representation. *Mind and Matter*, 2(2), 9-27. <https://doi.org/10.1017/S0140525X04280098>

↩

- Gibson, E., Futrell, R., Piantadosi, S. P., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, 23(5), 389-407. <https://doi.org/10.1016/j.tics.2019.02.003>

↩

- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

↩

- Goldberg, A. E. (2019). *Explain me this*. Princeton University Press.

↩

- Goldberg, A. E. (2024). A chat about constructionist approaches and LLMs. *arXiv*.

↩

- Haiman, J. (Ed.). (1985). *Iconicity in syntax*. John Benjamins Publishing.

↩

- Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences*, 9(9), 439-444. <https://doi.org/10.1016/j.tics.2005.07.003>

↩

- Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3), 146-162. <https://doi.org/10.1080/00437956.1954.11659520>

↩

- Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, 86(3), 663-687. <https://doi.org/10.1353/lan.2010.0021>

↩

- Haspelmath, M. (2023). Coexpression and synexpression patterns across languages: Comparative concepts and possible explanations. *Frontiers in Psychology*, 14, 1236853. <https://doi.org/10.3389/fpsyg.2023.1236853>

↩

- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford University Press.

↩

- Hay, J. B., & Baayen, R. H. (2005). Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences*, 9(7), 342-348. <https://doi.org/10.1016/j.tics.2005.04.002>

↩

- Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, 8(3), 164–181. <https://doi.org/10.1007/s10071-004-0239-6>

↩

- Hu, J., Mahowald, K., Lupyan, G., Ivanova, A., & Levy, R. (2024). Language models align with human judgments on key grammatical constructions. *Proceedings of the National Academy of Sciences, USA*, 121(36):e2400917121. <https://doi.org/10.1073/pnas.2400917121>

↩

- Jackendoff, R. S. (1985). *Semantics and cognition* (Vol. 8). MIT Press.

↩

- Jespersen, O. (1922). *Language, its nature, development & origin*. Allen & Unwin.

↩

- Kapatsinski, V. (2018). *Changing minds changing tools: From learning theory to language acquisition to language change*. MIT Press.

↩

- Knight, C., & Lewis, J. (2017). Wild voices: Mimicry, reversal, metaphor, and the emergence of language. *Current Anthropology*, 58(4), 435-453. <https://doi.org/10.1086/692905>

↩

- Lakoff, G. (1987). *Women, fire, and dangerous things : What categories reveal about the mind*. University of Chicago Press.

↩

- Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents* (Vol. 71). Cambridge University Press.

↩

- Langacker, R. W. (1987). *Foundations of cognitive grammar. Volume 1: Theoretical prerequisites*. Stanford University Press.

↩

- Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition*, 83(3), 265-294. [https://doi.org/10.1016/s0010-0277\(02\)00009-4](https://doi.org/10.1016/s0010-0277(02)00009-4)

↩

- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, 28(2), 203-208. <https://doi.org/10.3758/BF03204766>

↩

- Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2024). Dissociating language and thought in large language models: A cognitive perspective. *Trends in Cognitive Sciences*, 28(6), 517-540. <https://doi.org/10.1016/j.tics.2024.01.011>

↩

- Majid, A., Bowerman, M., Kita, S., Haun, D. B., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, 8(3), 108-114. <https://doi.org/10.1016/j.tics.2004.01.003>

↩

- Meir, I., Sandler, W., Padden, C., & Aronoff, M. (2010). Emerging sign languages. In M. Marschark & P. E. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education* (Vol. 2, pp. 267-280). Oxford University Press.

↩

- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97. <https://doi.org/10.1037/h0043158>

↩

- Misra, K., & Mahowald, K. (2024). Language models learn rare phenomena from less rare phenomena: The case of the missing AANNs. *arXiv*. <https://doi.org/10.48550/arXiv.2403.19827>

↩

- Murphy, G. (2004). *The big book of concepts*. MIT Press.

↩

- Navigli, R., Conia, S., & Ross, B. (2023). Biases in large language models: Origins, inventory, and discussion. *ACM Journal of Data and Information Quality*, 15(2), 1-21. <https://doi.org/10.1145/3597307>

↩

- Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(4), 700-708. <https://doi.org/10.1037/0278-7393.14.4.700>

↩

- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Aspell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *arXiv*. <https://doi.org/10.48550/arXiv.2203.02155>

↩

- Piantadosi, S. (2023). Modern language models refute Chomsky's approach to language. In E. Gibson & M. Poliak (Eds.), *From fieldwork to linguistic theory: A tribute to Dan Everett* (pp. 353-414). Language Science Press.

↩

- Pinker, S. (2000). *Words and rules: The ingredients of language*. Basic Books.

↩

- Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6(11), 456-463. [https://doi.org/10.1016/S1364-6613\(02\)01990-3](https://doi.org/10.1016/S1364-6613(02)01990-3)

↩

- Powers, R. (2004). *Galatea 2.2*. Picador.

↩

- Ramscar, M., Dye, M., & Hübner, M. (2013). When the fly flied and when the fly flew: How semantics affect the processing of inflected verbs. *Language and Cognitive Processes*, 28(4), 468-497. <https://doi.org/10.1080/01690965.2011.649041>

↩

- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573-605. [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9)

↩

- Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, 1, 45-76.

↩

- Sanchez-Stockhammer, C., & Uhrig, P. (2023). “I’m gonna get totally and utterly X-ed.” Constructing drunkenness. *Yearbook of the German Cognitive Linguistics Association*, 11(1), 121-150.
<https://doi.org/10.1515/gcla-2023-0007>

[↩](#)

- Sapir, E. (1921). *Language: An introduction to the study of speech*. Harcourt.

[↩](#)

- Saussure, F. D. (1916). *Course in general linguistics*. Fontana/Collins.

[↩](#)

- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523. <https://doi.org/10.1037/0033-295X.96.4.523>

[↩](#)

- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189–208.
<https://doi.org/10.1037/h0061626>

[↩](#)

- Tomasello, M. (2008). *Origins of human communication*. MIT Press.

[↩](#)

- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224–226. <https://doi.org/10.1038/d41586-023-00288-7>

[↩](#)

- Wilke, C., Lahiff, N. J., Sabbi, K. H., Watts, D. P., Townsend, S. W., & Slocombe, K. E. (2022). Declarative referential gesturing in a wild chimpanzee (*Pan troglodytes*). *Proceedings of the National Academy of Sciences, USA*, 119(47), e2206486119. <https://doi.org/10.1073/pnas.2206486119>

[↩](#)

- Wittgenstein, L. (1973). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Pearson (Original work published 1953)

[↩](#)