

MIT Press • Open Encyclopedia of Cognitive Science

Syntax

Rachel Nordlinger¹

¹University of Melbourne

MIT Press

Published on: Sep 24, 2026

URL: <https://oecs.mit.edu/pub/2c0dzjpe>

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

Syntax is the study of sentence structure—the ways in which words are combined into meaningful phrases and sentences. Syntax is what allows language users to describe and understand the details of complex events and attribute states and actions to entities in the world. It therefore provides a window into how the human mind represents and processes complex information. Sentences are not just strings of words, as shown by the fact that the English sentence “The dog chased the cat across the road” has the same words as the sentence “The cat across the road chased the dog,” yet they have very different meanings and describe different events. Furthermore, the string “Road the chased dog across the the cat” also has the same words but means nothing at all. The rules of syntax help to account for these facts, by determining which combinations of words are grammatical (or well formed) in a given language and how the arrangement of words in a sentence relates to its meaning. Central components of syntax include the linear ordering of words, the expression of grammatical relations such as subject and object, and the hierarchical grouping of words into phrases and larger constituents. Syntactic research investigates which aspects of syntactic structure are universal to all human languages, the ways in which languages vary, and the interface between syntactic structure and other elements of the linguistic system such as semantics, morphology, and discourse structure.

History

Although discussion of syntactic structure in language dates back to the works of the Sanskrit grammarian Pāṇini and the philosophers of Ancient Greece, it was Noam [Chomsky's \(1957\)](#) publication of *Syntactic Structures* that brought the study of syntax to the fore of linguistic analysis and cemented it as central to theoretical linguistics [see [Language](#)]. In this work, Chomsky argued for the independence of syntax from semantics (meaning) and introduced his idea of a transformational grammar (building on ideas of Zellig Harris and others) to provide a formal description of the rules of syntax. He used the now famous sentence, “Colorless green ideas sleep furiously,” to illustrate the fact that a sentence can be clearly grammatical according to the rules of syntax while being nonsensical in terms of meaning, thereby arguing for the importance of a generative theory of syntax that could account for these rules independently of semantics. Chomsky’s theories have evolved in many ways since 1957—for example, transformational grammar ([Chomsky, 1965](#)), government and binding theory ([Chomsky, 1981](#)), and minimalism ([Chomsky, 1995](#))—but have proven highly influential and remain a prominent theoretical perspective in syntactic research.

However, Chomsky’s theories are also polarizing (for an early critique, see [Hockett, 1970](#)), and therefore syntax is probably the most debated and divided area of modern linguistics. Diverse perspectives on syntax include alternative generative frameworks such as lexical-functional grammar ([Bresnan, 1982](#); [Dalrymple, 2023](#)) and head-driven phrase structure grammar ([Müller et al., 2024](#); [Pollard & Sag, 1987](#)), which share some of the formal principles of Chomsky’s work but differ in the detail of how syntactic structure is best modeled and how different parts of the linguistic system interact. Functional approaches to syntax—for example, role and reference grammar ([Bentley et al., 2023](#); [Foley & Van Valin, 1984](#)) and

construction grammar ([Goldberg, 1995](#))—deny the scientific validity of studying syntactic structure independently of its function as a tool for social interaction and communication and argue that structure must necessarily be analyzed with respect to its function and meaning in context. Also in this vein are theoretical approaches to syntax driven by frequency and usage (e.g., [Bybee, 2010](#); [Díaz-Campos & Balasch, 2023](#)) and linguistic processing (e.g., [Hawkins, 1994](#)).

Core concepts

Speakers of any language are able to produce an infinite number of sentences from a finite number of words; some of these will be judged grammatical by speakers, and others will not. The principles of syntax capture how this is possible in any given language and how it can vary across languages.

Constituent structure

Words can combine into larger constituents, called *phrases*, which then combine to form *clauses* and *sentences*. Thus, the syntactic structure of language is hierarchical, rather than concatenative, as in the following example:

1. [The small child] laughed at [[the dog] [having a swim]].

Although this example consists of a single English sentence, the brackets indicate internal constituency, showing that multiple words can group into larger constituents. Such constituents behave as a unit in the syntactic grammar, for example, being able to be repositioned as units in different syntactic constructions, as shown in examples two and three. This is not possible for strings of words that are not syntactic constituents, as shown in example four. A convention in linguistics is to use an asterisk to indicate ungrammatical examples, as in example four.

2. [The dog having a swim] was laughed at by [the small child].
3. It was [the dog] [the small child] laughed at [having a swim].
4. *It was [child laughed at] the small [[the dog] [having a swim]].

Constituents may also themselves consist of other constituents of the same type, which allows for recursive structures, sometimes claimed to be characteristic of human languages ([Hauser et al., 2002](#)). In example five, the single *noun phrase* (NP) “The manager of the cinema” contains another NP, “The cinema,” and the sentence (S) as a whole contains multiple sentences as well (in this case, “Pat told the manager of the cinema that the movie was terrible,” and also, “The movie was terrible”).

5. [I think that [Pat told [the manager of [the cinema]_{NP}]NP that [the movie was terrible]_S]_S].

In theory, this recursion is limitless, as the traditional nursery rhyme “This is the house that Jack built” famously illustrates, but in practice, speakers (and hearers) generally have a preference for shorter and less recursive sentences, as they are easier to process.

In some languages, there appears to be less reliance on hierarchical structure, such as in example six from Jiwarli ([Austin & Bresnan, 1996](#)), in which semantically related words are distributed throughout the sentence rather than forming a single constituent, illustrating one of the many ways in which the syntax of languages can vary. The second line provides the literal translation (the *gloss*) of each Jiwarli word. The ergative case suffix marks the subject of the clause, and the accusative case marks the object.

6. *Luru-ngku* ***ngatha-nha*** *kulypajipa-minyja* ***parna.***
 sun-ergative my-accusative make.sore-past head.accusative

“The sun made my head sore.” ([Austin, 2001](#), p. 312)

In example six, the two words translating to the object “my head” (in bold) do not appear adjacent to each other, as they must be in English, but the sentence is perfectly grammatical in Jiwarli.

Grammatical relations and argument structure

Syntactic relations do not only concern the hierarchical constituency of words in the sentence but also the roles that constituents play in the clause, known as *grammatical relations*. Common grammatical relations are *subject*, *direct object*, *indirect object*, and *adjunct*. The latter is an optional element that is not required by the verb but added to provide extra information or modification to the clause, whereas the other three are *arguments*. In examples seven and eight, the arguments are in bold and the adjuncts in italics.

7. **Sam** *slept on the bed.*

8. **The teacher** showed **the answers** *individually to the students after school.*

Languages differ in the ways that they encode grammatical relations. Languages like English use linear order and hierarchical constituency in the phrase structure to do this; subjects, such as “Sam” and “The teacher” in examples seven and eight, come before the verb, whereas objects and indirect objects follow the verb and can be shown to form a constituent with it (commonly referred to as a *verb phrase*). Other languages rely primarily on morphological means, such as *case marking* and/or *verbal agreement morphology* to encode grammatical relations; such languages have been called *nonconfigurational* ([Austin & Bresnan, 1996](#); [Hale, 1983](#); [Nordlinger, 1998](#)) [see [Morphology](#)].

Verbs select for semantic arguments, which are often referred to as *thematic roles* and are realized in the syntax as grammatical relations. A verb such as “fix” in example nine requires an *agent* (the initiator and actor, Sam) and a *patient* (the entity that is acted upon, the car); agents are typically subjects, and patients are typically objects.

9. Sam_[Agent, Subject] fixed the car_[Patient, Object].

Although there are generalizations that can be made about the relationship between grammatical relations and thematic roles cross-linguistically, the importance of distinguishing the two is shown by *valency-changing operations* such as passive constructions (example 10), in which the patient of the underived verb becomes the subject of the passive construction while remaining a patient. The agent in this case becomes an optional adjunct.

(10) The car _[Patient, Subject] was fixed (by Sam_[Agent, Adjunct]).

Questions, controversies, and new developments

A major long-standing debate in syntactic research focuses on the extent to which there is a universal syntactic structure common to all human languages and, if so, how best to model it in a theoretical framework. Some frameworks, such as Chomsky's minimalist program ([Chomsky, 1995](#)), postulate a highly articulated universal syntactic structure that incorporates phenomena that would be attributed to levels of morphology, semantics, or discourse structure in a different framework, such as lexical-functional grammar ([Dalrymple, 2023](#)).

Related to this is the question as to whether universal syntactic knowledge in humans is innate because of a language-specific endowment (*Universal Grammar*) or whether it is driven by domain-general cognitive processes (see [Cowie, 2017](#) for discussion) [see [Language Acquisition](#)].

More recent developments in syntax have focused on the use of corpora and experimental methods for syntactic analysis (e.g., [Goodall, 2021](#); [Levshina et al., 2023](#)) and increasing discussion of cross-linguistic variation in theoretical debate, which has traditionally been based on English and a small number of Indo-European languages.

Broader connections

Syntax and its hierarchical structure have been a core focus of research in the cognitive science of language, including language acquisition [see [Computational Models of Language Learning](#); [Language Acquisition](#)] and psycholinguistics [see [Neuroscience of Syntax](#); [Psycholinguistics](#)]. This research has used naturalistic, experimental, computational, and neuroscientific methods to investigate whether and how syntax is learnable or innate, how children acquire syntactic structures, and how syntactic operations are implemented in the minds and brains of language users. More recently, the advent of large language models has led to research aiming to determine how the models represent syntax and what this might mean about human representation, acquisition, and processing (e.g., [Baroni, 2022](#)) [see [Large Language Models](#)].

Further reading

- Austin, P., & Bresnan J. (1996). Non-configurationality in Australian aboriginal languages. *Natural Language & Linguistic Theory*, 14(2), 215-268. <https://doi.org/10.1007/BF00133684>

- Carnie, A. (2021). *Syntax: A generative introduction* (4th ed.). Wiley-Blackwell.
- Croft, W. (2022). *Morphosyntax: Constructions of the world's languages*. Cambridge University Press.
- Tallerman, M. (2025). *Understanding syntax* (6th ed.). Routledge.

References

- Austin, P. (2001). Word order in a free word order language: The case of Jiwari. In J. Simpson, D. Nash, M. Laughren, P. Austin, & B. Alpher (Eds.), *Forty years on: Ken Hale and Australian languages* (pp. 305-324). Pacific Linguistics.

[↩](#)

- Austin, P., & Bresnan J. (1996). Non-configurationality in Australian aboriginal languages. *Natural Language & Linguistic Theory*, 14(2), 215-268. <https://doi.org/10.1007/BF00133684>

[↩](#)

- Baroni, M. (2022). On the proper role of linguistically oriented deep net analysis in linguistic theorizing. In S. Lappin & J.-P. Bernardy (Eds.), *Algebraic structures in natural language* (pp. 1-16). CRC Press.

[↩](#)

- Bentley, D., Mairal Usón, R., Nakamura, W., & Van Valin, R. D., Jr. (Eds.). (2023). *The Cambridge handbook of role and reference grammar*. Cambridge University Press.

[↩](#)

- Bresnan, J. (Ed.). (1982). *The mental representation of grammatical relations*. MIT Press.

[↩](#)

- Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press.

[↩](#)

- Chomsky, N. (1957). *Syntactic structures*. de Gruyter.

[↩](#)

- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

[↩](#)

- Chomsky, N. (1981). *Lectures on government and binding*. Foris Publications.

[↩](#)

- Chomsky, N. (1995). *The minimalist program*. MIT Press.

[↩](#)

- Cowie, F. (2017). Innateness and language. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2017 ed.). Stanford University Press.

↩

- Dalrymple, M. (Ed.). (2023). *Handbook of lexical functional grammar*. Language Science Press.

↩

- Díaz-Campos, M., & Balasch, S. (2023). *The handbook of usage-based linguistics*. Wiley-Blackwell.

↩

- Foley, W. A., & Van Valin, R. D., Jr. (1984). *Functional syntax and universal grammar*. Cambridge University Press.

↩

- Goldberg, A. (1995). *Construction grammar: A construction grammar approach to argument structure*. University of Chicago Press.

↩

- Goodall, G. (Ed.). (2021). *The Cambridge handbook of experimental syntax*. Cambridge University Press.

↩

- Hale, K. (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory*, 1(1), 5-47. <https://doi.org/10.1007/BF00210374>

↩

- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569-1579. <https://doi.org/10.1126/science.298.5598.1569>

↩

- Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge University Press.

↩

- Hockett, C. F. (1970). *The state of the art*. Mouton.

↩

- Levshina, N., Namboodiripad, S., Allasonnière-Tang, M., Kramer, M., Talamo, L., Verkerk, A., Wilmoth, S., Rodriguez, G. G., Gupton, T. M., Kidd, E., Liu, Z., Naccarato, C., Nordlinger, R., Panova, A., & Stoyanova, N. (2023). Why we need a gradient approach to word order. *Linguistics*, 61(4), 825-883. <https://doi.org/10.1515/ling-2021-0098>

↩

- Müller, S., Abeillé, A., Borsley R. D., & Koenig, J.-P. (Eds.). (2024). *Head-driven phrase structure grammar: The handbook* (2nd ed.). Language Science Press.

[←](#)

- Nordlinger, R. (1998). *Constructive case: Evidence from Australian languages*. CSLI Publications.

[←](#)

- Pollard, C., & Sag, I. (1987). *Information-based syntax and semantics, volume 1: Fundamentals*. CSLI Publications.

[←](#)