

MIT Press • Open Encyclopedia of Cognitive Science

Language Acquisition

Ben Ambridge¹

¹Division of Psychology, Communication and Human Neuroscience, University of Manchester and the ESRC International Centre for Language and Communicative Development (LuCiD)

MIT Press

Published on: Jul 24, 2024

DOI: <https://doi.org/10.21428/e2759450.b4a2349c>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Child language acquisition is the process by which infants and children come to speak the language of the community around them. Children must acquire a great deal of linguistic knowledge, including the set of speech sounds in the language they are learning, the ways in which their language combines them (phonetics and phonology), the individual words of their language, the ways these words are used in context, and more. In particular, language acquisition must also include the acquisition of core morphology and syntax (i.e., how languages mark who did what to whom). In very broad terms, there are two answers to the question of how children acquire language, each taking a different view of language. Under *generativist* accounts, language is a set of formal rules for putting words together; under *constructivist* accounts, children do not start out with formal rules but begin by storing phrases that they hear around them.

History

Prehistory

Until the mid 1950s, child language acquisition research consisted mainly of diaries and informal observations of children's speech. The earliest known example is Dietrich Tiedemann's 1787 diary of the linguistic development of his (German-speaking) son (see [Levelt, 2013](#)). Other diaries soon followed, including those by Hippolyte Taine (1876, French), Charles Darwin (1877, English), William Preyer (1882, German), and Clara and William Stern (1907, German).

Although the Sterns also conducted some of the earliest experimental studies of child language acquisition ([Levelt, 2013](#)), perhaps the first that remains influential and highly cited to this day is Jean Berko-Gleason's study of English morphology ([Berko, 1958](#)). Morphology is the process by which words change (or *morph*) to mark things like singular versus plural ("one cat," "two cats") or who is doing the action (e.g., "I read," "he reads") [see [Morphology](#)]. In the most famous part of the study, Berko showed children a picture of a novel, made-up creature and said, "This is a wug" (a novel, made-up word). She then showed children a picture with two of these creatures and said, "Now there is another one. There are two of them. There are two...", inviting children to fill in the blank. The headline finding is that children correctly said "wugs." Indeed, this was true for 97% of children aged five years and over. What is less commonly reported is that, for under-fives, this figure was only 76%. Furthermore, for more complex plurals (e.g., "tasses" or "gutches"), only around 30%–40% of children gave the correct answer regardless of age group.

The findings of [Berko \(1958\)](#) prefigure later debates between generativist and constructivist approaches to children's acquisition of morphology. The key distinction between these views can be illustrated with a simple example. Under a generativist account, one rule of English says that a determiner (e.g., "a," "the") always goes before the relevant noun (e.g., "cat," "dog"); we say "the cat and the dog," not "*dog the" ("*" means that the combination is ungrammatical). Generativist accounts assume that children figure out these rules very early, in some cases before they speak themselves, perhaps with the help of innate knowledge (e.g., knowledge from

birth that, generally speaking, determiners are either always before or always after nouns rather than varying randomly). They produce new sentences by analogizing across stored exemplars (e.g., “It’s *a cat* and *a dog*, so it must be *a dinosaur*, not **dinosaur a*”). In contrast, under most constructivist accounts, the adult endpoint is not a set of rules at all, but groups of exemplars (e.g., “a dog,” “a cat,” “a dinosaur”) that support generalizations known as *constructions* (e.g., a [THING]).

Returning to the findings of [Berko \(1958\)](#), generativist researchers point to children’s high level of success with some items as evidence that they have learned the general rule (e.g., plural = NOUN + “s”). Constructivist researchers point to between-item differences as evidence for analogy; i.e., that children are more easily able to find a soundlike analogy for “wug” > “wugs” (e.g., “bug” > “bugs,” “hug” > “hugs,” “mug” > “mugs”) than “gutch” > “gutches” (e.g., perhaps only the relatively rare “crutch” > “crutches”).

The development of modern theoretical perspectives (1960 – 1990)

The 1960s saw the development of the precursors of modern generativist and constructivist accounts. Given his development of the general theoretical frameworks that underpin acquisition research in the generativist framework, Noam Chomsky is considered an important figure in child language acquisition research (e.g., [Chomsky, 1965](#)). Much research in the constructivist framework takes the position that Chomsky’s theoretical frameworks have led the field of language acquisition in the wrong direction (e.g., [Tomasello, 2003](#)). Either way, it is important to be clear that Chomsky’s contributions to language acquisition—like B.F Skinner’s—are purely theoretical; neither conducted a study of children or the language input they hear.

The 1960s also saw the publication of an influential paper on *pivot schemas*. [Braine \(1963\)](#) reported the findings of one of the earliest naturalistic corpus studies, in which investigators (typically the child’s parents) try to capture everything the child says in a given period (in those days, generally by keeping handwritten notebooks). Braine reported that at the earliest stages of acquisition, children’s knowledge of language seems to consist of a toolkit of schemas that each combine a fixed word (the pivot) and a kind of slot into which suitable words can be inserted (e.g., “my [THING]”; e.g., “my Mommy,” “my Daddy,” “my milk”) [see [Word Learning](#)]. It was soon recognized that pivot schemas per se are far too coarse and simplistic to capture even early language development, but the analysis is important historically as perhaps the first to formalize a constructivist input-based approach to language acquisition as opposed to a generativist approach based on formal rules.

The 1970s saw the publication of another landmark corpus study ([Brown, 1973](#)). Although the (Biblically codenamed) children Adam, Eve, and Sarah were recorded in the early 1960s, the corpus was later computerized and made publicly available and is still used in research today (as part of the [CHILDES corpus](#)). Much of Brown’s focus was on inflectional morphology. A morphological rule (e.g., adding “-s” to a verb referring to “he/she/it”: “she likes cake,” “John likes cake,” “it goes there”) was deemed to have been mastered when a child applied it correctly on 90% of occasions (as opposed to, for example, incorrectly saying “*John

like cake”—a common error by English-speaking children). This analysis prefigured later theoretical debates around exactly why children make these kinds of errors. On a generativist analysis, errors such as “*He like cake” reflect either a stage in which children treat certain rules as optional (e.g., [Wexler, 1998](#)), or in which they are trying out different rules before settling on those that apply to their particular language (e.g., [Legate & Yang, 2007](#)). On a constructivist analysis (e.g., [Freudenthal, Gobet, & Pine, 2023](#)), these errors do reflect not children’s failure to apply a VERB + “s” rule, but the truncation (shortening) of rote-learned phrases from the input (e.g., “[Does] he like cake?”) or the fact that forms without the “-s” ending dominate in the language that children hear (e.g., “I like,” “you like,” “they like”...).

Although the 1970s saw occasional studies of the acquisition of languages other than English (e.g., [Bowerman, 1973](#), Finnish; [MacWhinney, 1976](#), Hungarian), it was only in the 1980s that researchers began to systematically compare acquisition across multiple languages. Beginning with [Slobin and Bever \(1982\)](#), a series of studies investigated how children learning different languages come to understand “who did what to whom” in the sentences they hear. English relies strictly on word order: “The dog chased the cat” can mean only that the dog is the chaser and the cat, the one chased. But in languages such as Italian and Serbian/Croatian (among those studied by Slobin and Bever), morphological *case* markers on nouns mark (for examples like this one) the SUBJECT (chaser) and OBJECT (the one chased), allowing the word order to vary. For example, “The dog + OBJECT chased the + cat + SUBJECT” means “The cat chased the dog” (or something like “As for the dog, the cat chased it”). At least, this is true for adults. For many languages, children misinterpret such sentences by assuming that they follow the more frequent word order of their language (e.g., SUBJECT VERB OBJECT for Serbian/Croatian), effectively ignoring the morphological case markers. The constructivist-oriented *competition model* (e.g., [Bates & MacWhinney, 1989](#)) argues that this pattern arises because the word-order cue is present in just about all of the relevant sentences children hear, while case marking is present and informative relatively rarely.

Many generativist accounts of the 1980s also focused on word-order acquisition. For example, the edited book *Parameter Setting* ([Roeper & Williams, 1987](#)) set out several theoretical accounts under which children learn word order by using the language they hear to set innate *switches* or parameters to the relevant setting for their language (see [Snyder, 2021](#)). Serbian/Croatian (like English) generally follows SVO word order (e.g., “the girl kicked the ball”), but between them, the world’s languages allow all logically possible word orders (from most to least common, SOV, SVO, VSO, VOS, OVS, and OSV). Under parameter-setting accounts, then, children set (broadly speaking) one switch that determines whether the VERB comes before or after the OBJECT, and another that determines whether the SUBJECT comes before or after the VERB. [Pinker \(1984\)](#) set out an alternative generativist account under which children use innate mappings between semantic and syntax (*linking rules*) to learn word order. For example, children are born knowing the following across languages:

1. The AGENT (the one doing the action) tends to be the SUBJECT (e.g., “the girl”).
2. The PATIENT (the person/thing that has the action done to it) tends to be the OBJECT (e.g., “the ball”).

3. The ACTION (e.g., “kicked”) tends to be the verb.

This means that a child who heard (and understood) “the girl kicked the ball” could read off the word order of her language (here, SVO).

Further methodological and theoretical debate (1990 – present)

Technological developments in the 1990s allowed recordings of adult–child conversations to be computerized and, crucially, automatically searched for whatever linguistic structures were being investigated (e.g., [MacWhinney & Snow, 1990](#)), leading to a resurgence in naturalistic corpus-based studies. For example, focusing on children’s acquisition of determiners (e.g., “the” and “a”), generativist-oriented researchers argued that corpus data suggest that children have a “the”/“a” + NOUN rule (e.g., “a dog,” “the cat,” “a dinosaur”) from the earliest stages of development (e.g., [Valian, 1986](#); [Yang, 2013](#)), with constructivist researchers arguing that this position is not supported by the data (e.g., [Pine & Martindale, 1996](#); [Pine, Freudenthal, Krajewski & Gobet, 2013](#)). The debate continues into the modern era (e.g., [Meylan, Frank, Roy, & Levy, 2017](#)), though some accounts (e.g., [Ambridge 2020a, 2020b](#)) have argued that these kinds of data are consistent with both generativist-style early productivity and constructivist-style early rote-learned phrases (e.g., “the” + “cup,” “the” + “ball”).

The 1990s also saw a slew of experimental studies investigating children’s acquisition of basic word order. In general terms, the generativist position (e.g., [Wexler, 1998](#)) is that children already have the relevant rules (e.g., for English, SVO) from the youngest age at which they can be tested. The constructivist position (e.g., [Tomasello, 2000](#)) is that children start out with rote-learned phrases (e.g., “I” + “want” + “it”) and slot-and-frame patterns (e.g., “I want [THING]”) and do not acquire the equivalent of a fully general SVO “rule” “in most cases until around the third birthday” ([Tomasello, 2000](#), p. 215). In the 1990s, a series of studies used novel, made-up verbs created for the purposes of the experiment (e.g., “meeking” or “taming”) to test the generativist claim of general rules that apply to *any* verb against the constructivist claim of (amongst other things) individual slot-and-frame patterns for particular verbs (e.g., “he’s eating [THING]” for the verb “eat”). A comprehensive summary of these studies can be found in [Ambridge and Lieven \(2015\)](#), but, in brief, the pattern is as follows. In production studies, which require children to actually produce new sentences with the novel verbs (e.g., “Big Bird is meeking Cookie Monster”), it is true that children do not generally succeed “until around the third birthday,” except when they have a suitable slot-and-frame pattern into which they can insert the novel verb (e.g., “he’s [ACTION]ing it”; [Dodson & Tomasello, 1998](#)). A similar pattern is found in act-out studies in which children are given familiar toys and asked to act out, for example, “Big Bird is meeking Cookie Monster” (e.g., [Akhtar & Tomasello, 1997](#)).

However, children show earlier knowledge of SVO word order in comprehension studies in which they hear (for example) “the duck is meeking the bunny” and have to choose between one screen showing a duck performing a novel action on a bunny and another screen showing a bunny performing a novel action on a duck

(by either pointing at, or simply looking at, the matching screen). The “third birthday” claim was finally put to bed in the mid-2000s with studies showing that, although the effect is fragile, children can succeed even before their second birthday ([Gertner, Fisher & Eisengart, 2006](#); [Fernandes, Marcus, Di Nubila & Vouloumanos, 2006](#)). Taken together with the word-order studies summarized above, these findings suggest the need for a theoretical account of basic word-order acquisition that can accommodate both generativist-style early knowledge of “rules” and constructivist-style early reliance on simple slot-and-frame patterns (e.g., “I’m [ACTION]ing it”).

A similar conclusion can be drawn from the debate around children’s acquisition of questions, which came to the fore in the 2000s. Questions are particularly interesting because (at least in languages such as English) a particular type of word-order error is very common: Children ask, for example, “*What he is eating?” rather than “What is he eating?” Generativist accounts (e.g., [Pozzan and Valian, 2017](#)) see such errors in terms of children’s occasional failure to apply a general (possibly innate) rule that starts out with non-question word order (e.g., “he is eating”) and moves the auxiliary verb (“is”) before the SUBJECT (“he”). Constructivist accounts (e.g., [McCauley et al., 2021](#)) argue that errors such as “*What he is eating?” instead reflect the use of stored chunks that children have heard in declarative (non-question) sentences (e.g., “He is eating”), which they then combine with a question word (“what”). Again, detailed explanation is still required to elucidate why children seem able to apply a general “rule” in some cases but produce errors in others.

Finally, no historical account of child language acquisition research would be complete without mention of the long-running and (in)famous English past-tense debate (e.g., [Rumelhart & McClelland, 1986](#); [Prasada & Pinker, 1993](#)). Generativist accounts (e.g., [Pinker, 1999](#)) argue that children show evidence for a default rule (adding “-ed”) that can be applied to any verb (including made-up verbs) regardless of its sound (e.g., “yesterday, I ploamphed”). Constructivist accounts (e.g., [Marchman, Wulfeck & Weismer, 1999](#)) argue that children show evidence for analogy between similar-sounding words (e.g., the made-up verb “wiss” has the past-tense “wissed” [pronounced “wist”]) by analogy with “miss”–“missed,” “hiss”–“hissed,” and so on. As is often the case, the debate petered out with both sides claiming victory (e.g., [Pinker & Ullman, 2002](#); [McClelland & Patterson, 2002](#)). One thing is clear: This debate, like many of those summarized above, has been hampered by an overreliance on English (e.g., [Granlund et al., 2019](#)). Any successful account of the acquisition of inflectional morphology, or of language more generally, will have to apply in principle to all the 7,000 or so languages spoken worldwide ([Kidd & Garcia, 2022](#)).

Core concepts

It is important to emphasize that the labels adopted here—*generativist* and *constructivist*—are necessarily broad and imprecise. Each approach encompasses a wide variety of different individual theories, which often differ importantly in their details. Neither would individual researchers necessarily self-identify as members of one or another theoretical camp (much as musical genres might be helpful for critics and listeners but are often eschewed or even derided by artists themselves). In the same way, each of the core concepts set out here is

broadly consistent with one or another general approach but does not necessarily form a key part of any particular theory.

Generativist concepts

Generativist theories generally adopt the assumption of the *autonomy of syntax*: that “the rules (principles, constraints, etc.) that determine the combinatorial possibilities of the formal elements of a language make no reference to constructs from meaning, discourse, or language use” (Newmeyer, 2016). Of course, syntax must *interface* with meaning, discourse, and language use, and at least one generativist approach (e.g., Grinstead, 2021) argues that childlike errors are not a result of non-adultlike knowledge of grammar (syntax) but of the interface between syntax and discourse/language use. This is an example of the famous generativist distinction between *competence* and *performance* (e.g., Chomsky, 1965): A child might have fully adultlike knowledge of syntax (perfect competence) but still make childlike errors with language (e.g., “it goed over there”)—that is, imperfect performance—due to other nonlinguistic factors such as memory failure.

Most, perhaps all, generativist accounts are also *nativist* accounts: “The prevailing opinion among generative grammarians since the 1960s has been that this system (i.e., the system that combines words and phrases into sentences) is not only situated in the human mind, but also that its fundamental principles, its inventory of combinatorial elements, and so on are innate” (Newmeyer, 2016). Precisely what is innate varies from theory to theory, but it is usually, minimally, some grammatical categories such as NOUN and VERB (but not, of course, the individual nouns and verbs of the language to be learned) and some basic rules for combining them into sentences. Some theories remain rather vague about exactly what is innate, but an excellent counterexample to this vagueness is Valian (2014): “There is good evidence for at least one innate idea – Determiners...It is impossible to have Determiners without having Nouns, because part of the definition of Determiners is that they take Noun Phrases as their complement. *Après Determiners, le deluge.*”

Chomsky famously coined the terms *language acquisition device* (LAD) and *universal grammar* (UG) as metaphors for this innate knowledge possessed by learners of all the world’s languages (i.e., “universal”). One common generativist argument for the necessity of the LAD or UG is the argument from the poverty of the stimulus: the claim that it is not possible to figure out the underlying rules of language simply by hearing examples of sentences generated using those rules. It is important to note that the term *language acquisition* itself is controversial, as it carries the implication that many crucial aspects of language are not learned—only “acquired”—since they rely on innate knowledge or mature like “the development of a second set of teeth” (Wexler, 1996, p. 117): an implication generally disputed by constructivist accounts, which eschew innate linguistic knowledge and emphasize learning.

Constructivist concepts

Constructivist accounts are often called *input-based* or *usage-based* accounts because they emphasize the importance of children learning language on the basis of the *input* (i.e., the language of parents, caregivers, and

other adults) and, importantly, by understanding how these adults are *using* language (e.g., [Tomasello, 2003](#)). For example, a child might understand that “Do you want some juice?” is a question not because it has certain grammatical properties (i.e., auxiliary-before-subject word order, as discussed above) but because it is produced with a particular intonation and because the parent is holding a cup and an open bottle of juice and raising their eyebrows in the child’s direction. [Bruner \(1983\)](#) famously coined the *language acquisition support system* as a metaphor for this real-world support, and as a kind of retort or alternative to Chomsky’s LAD.

Constructivist accounts assume that children start out by learning *frozen phrases* (also known as [rote-learned] fixed phrases or holophrases) such as “I’m kicking it” and “I’m eating it,” each paired with a meaning. Next, children abstract across these frozen phrases via a process known as *schematization* to form partially productive, lexically specific construction schemas (also known as slot-and-frame patterns) such as “I’m [ACTION]ing it” (e.g., [Lieven, Pine & Baldwin, 1997](#)). Finally, children abstract across these schemas (or individual stored utterances) to arrive at *fully abstract constructions* (e.g., [SUBJECT] [VERB] [OBJECT]) that can be used to produce or understand any relevant sentence. Constructivist accounts are not always clear as to whether an abstract construction like [SUBJECT] [VERB] [OBJECT] is actually stored in some sense, or whether a *construction* is just a kind of shorthand or metaphor for a fuzzy cluster of stored individual utterances (*exemplars*), which speakers generalize across when producing or understanding relevant utterances (e.g., [Ambridge 2020a; 2020b](#)).

Questions, controversies, and new developments

Open Science

One recent development in the field—as in cognitive science more broadly—has been an increasing adoption of open science [see [Open Science](#)]. These practices attempt to combat what [Bishop \(2019\)](#) calls the “four horsemen of irreproducibility,” referring to the fact that many key findings reported in published papers cannot successfully be reproduced by colleagues:

1. publication bias, whereby studies that fail to show an expected effect are deemed too uninteresting for publication;
2. low statistical power, whereby studies are run with low numbers of participants (or few observations per participant), making their results unreliable;
3. *p*-hacking, whereby researchers run multiple statistical analyses and report only those that meet a criterion of statistical significance; and
4. hypothesizing after results are known (HARKing), whereby unexpected (and perhaps supurious) results are retrospectively “predicted” in the Introduction section of the paper (like a would-be sharpshooter firing at a barn at random, then drawing a target around the bullet holes).

In an attempt to rein in the third and fourth horsemen, many researchers now publicly pre-register their hypotheses and analysis plans on websites such as the [Open Science Framework](#), and many journals offer a registered report format by which studies are accepted in principle on the basis of their methods and analysis plans regardless of the eventual results.

One solution to the problem of low statistical power is large, multisite replications. Here, language acquisition research has been a leader of the wider field, especially through the efforts of the [ManyBabies Consortium](#). In an attempt to combat publication bias, a new open science journal, *Language Development Research*, launched in 2020 with a commitment “to publishing any empirical or theoretical paper that is relevant to the field of language development and that meets our criteria for rigour, without regard to the perceived novelty or importance of the findings.”

Theoretical developments

Turning from methods to theory, recent years have seen the acceleration of an approach that gained traction in the 1980s: implementing theoretical proposals as computational models of language learning [see [Bayesian Models of Cognition](#)]. The advantage of computational models—as opposed to traditional descriptive verbal theories—is that they make precise quantitative predictions that can then be tested against data from child corpora or experiments. For example, the patterning of a particular error made by English-learning children (e.g., saying “*mouses” instead of “mice”) can be explained by a simple discriminative-learning model in which various meaning-based factors or “cues” (e.g., multiple items, multiple mouse items, mousiness) compete to predict the occurrence of the form “mice” versus “mouses” ([Ramscar, Dye, and McCauley, 2013](#)). While discriminative learning is broadly consistent with a constructivist approach, generativist approaches have also adopted computational modeling as a way of formal theory testing. For example, variational learning approaches (see [Pearl, 2021](#), for a review) adopt traditional generativist assumptions such as parameter setting and rules for question formation but assume that children implement these parameters or rules gradually and probabilistically on the basis of the input.

Most recently, child language acquisition—like many other fields—has seen its attention captured by large language models (LLMs) such as ChatGPT [see [Large Language Models](#)]([OpenAI, 2023](#)). Interestingly, the field’s response to LLMs has largely split along traditional party lines: Constructivist-oriented researchers (e.g., [Piantadosi, 2023](#)) argue that LLMs not only simulate many aspects of a constructivist approach to language acquisition, but even—to quote the title of Piantadosi’s controversial paper—“refute Chomsky’s approach to language.” Generativist-oriented researchers (e.g., [Kodner, Payne & Heinz, 2023](#)) are skeptical precisely because LLMs eschew the symbolic, categorical representations (e.g., VERB, NOUN) that are assumed, under this approach, to characterize language and its acquisition. On this view, as [Kodner et al. \(2023\)](#) put it, “the implications of LLMs for our understanding of the cognitive structures and mechanisms underlying language and its acquisition are like the implications of airplanes for understanding how birds fly.” Either way, with the ever-increasing processing power and flexibility of modern computer systems, what seems certain is that

computational modeling—of whatever theoretical stripe—will play an increasingly key role in understanding child language acquisition.

Broader connections

Signed languages (e.g., British Sign Language, American Sign Language) are acquired in more or less the same way as spoken languages (see [Lillo-Martin & Henner, 2021](#), for a review). Therefore, any successful theory will have to account for the acquisition of spoken and signed languages alike. Language acquisition research also includes children whose acquisition of language is (for want of a better term) “atypical” in some way: particularly the 7%–8% of children with developmental language disorder ([Norbury et al., 2016](#)).

As is usual for the field, *language acquisition* in this entry is defined as the process by which children come to speak their first (or *native*) language (or, since bi- or multilingual acquisition is very much included, *first languages* [plural]). A field that is clearly related, but that has perhaps surprisingly little overlap, investigates how older children or adults learn a second (or third, etc.) language, either via formal instruction (e.g., in school) or immersion (e.g., moving to a country where that language is spoken). Indeed, this area of research is often called *language learning* specifically to contrast with the field of language acquisition. A large study of almost 700,000 (aspiring) English speakers found that the cut-off point between *native language acquisition* and *foreign language learning* is surprisingly late: Given sufficient immersion, learners can acquire native-like accuracy in a foreign language up until around age 17, after which it tails off ([Hartshorne, Tenenbaum & Pinker, 2018](#)). The study of language representation and processing in mature adult native speakers is called psycholinguistics [see [Psycholinguistics](#)].

Another field that is clearly related to language acquisition, but that is again perhaps surprisingly distinct, is literacy acquisition (i.e., learning to read and write). The two are linked in that, as you might expect, children who experience difficulty with spoken language early in life often experience difficulties with reading and writing later (e.g., [Botting, 2020](#)). However, the two are distinct in that while almost all children come to speak the language of those around them (and without explicit instruction), it is only since the 1960s that a majority of children globally have been taught to read and write ([Roser & Ortiz-Ospina, 2016](#)).

In summary, although considerable progress has been made in the past two-thirds of a century of systematic research, many questions surrounding child language acquisition remain unanswered. Hopefully, the methodological and theoretical developments summarized here will accelerate the progress of the field towards a more complete understanding of the processes and mechanisms by which children acquire their native language.

Acknowledgments

The support of the Economic and Social Research Council [ES/L008955/1] is gratefully acknowledged.

Further reading

- Behrens, H. (2021). Constructivist approaches to first language acquisition. *Journal of Child Language*, 48(5), 959–983. <https://doi.org/10.1017/S0305000921000556>
- Pearl, L. (2021). Theory and predictions for the development of morphology and syntax: A universal grammar + statistics approach. *Journal of Child Language*, 48(5), 907–936. <https://doi.org/10.1017/S0305000920000665>
- Ramscar, M. (2021). How children learn to communicate discriminatively. *Journal of Child Language*, 48(5), 984–1022. <https://doi.org/10.1017/S0305000921000544>
- Snyder, W. (2021). A parametric approach to the acquisition of syntax. *Journal of Child Language*, 48(5), 862–887. <https://doi.org/10.1017/S0305000921000465>

References

- 10.1017/S0305000921000477
[↵](#)
- Akhtar, N., & Tomasello, M. (1997). Young children’s productivity with word order and verb morphology. *Developmental Psychology*, 33(6), 952–965. <https://doi.org/10.1037/0012-1649.33.6.952>
[↵](#)
- Ambridge, B. (2020a). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, 40(5-6), 509–559. <https://doi.org/10.1177/0142723719869731>
[↵](#)
- Ambridge, B. (2020b). Abstractions made of exemplars or You’re all right and I’ve changed my mind. Response to commentators. *First Language*, 40(5-6), 640–659. <https://doi.org/10.1177/0142723720949723>
[↵](#)
- Ambridge, B., & Lieven, E. V. M. (2015). A Constructivist account of child language acquisition. In B. MacWhinney & W. O’Grady (Eds.), *Handbook of language emergence* (pp. 478–510). Wiley Blackwell. <https://doi.org/10.1002/9781118346136.ch22>
[↵](#)
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–76). Cambridge University Press.
[↵](#)
- Berko, J. (1958). The child’s learning of English morphology. *Word*, 14(2-3), 150–177. <https://doi.org/10.1080/00437956.1958.11659661>
[↵](#)

- Bishop, D. (2019). Rein in the four horsemen of irreproducibility. *Nature*, 568(7753), 435.
<https://doi.org/10.1038/d41586-019-01307-2>

↵

- Botting, N. (2020). Language, literacy and cognitive skills of young adults with developmental language disorder (DLD). *International Journal of Language & Communication Disorders*, 55(2), 255–265.
<https://doi.org/10.1111/1460-6984.12518>

↵

- Bowerman, M. (1973). *Early syntactic development: A cross-linguistic study with special reference to Finnish*. Cambridge University Press.

↵

- Braine, M. D. (1963). The ontogeny of English phrase structure: The first phase. *Language*, 39(1), 1–13.
<https://doi.org/10.2307/410757>

↵

- Brown, R. (1973). *A first language: The early stages*. Harvard University Press.
<https://doi.org/10.4159/harvard.9780674732469>

↵

- Bruner, J. S. (1983). *Child's talk: Learning to use language*. Norton.

↵

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

↵

- Dodson, K., & Tomasello, M. (1998). Acquiring the transitive construction in English: The role of animacy and pronouns. *Journal of Child Language*, 25(3), 605–622. <https://doi.org/10.1017/S0305000998003535>

↵

- [dx.doi.org/10.48550/arXiv.2308.03228](https://doi.org/10.48550/arXiv.2308.03228)

↵

- Fernandes, K. J., Marcus, G. F., Di Nubila, J. A., & Vouloumanos, A. (2006). From semantics to syntax and back again: Argument structure in the third year of life. *Cognition*, 100(2), B10–B20.
<https://doi.org/10.1016/j.cognition.2005.08.003>

↵

- Freudenthal, D., Gobet, F., & Pine, J. M. (2023). MOSAIC+: A crosslinguistic model of verb-marking errors in typically developing children and children with developmental language disorder. *Language Learning*.
<https://doi.org/10.1111/lang.12580>

[↑](#)

- Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*, 17(8), 684–691. <https://doi.org/10.1111/j.1467-9280.2006.01767.x>

[↑](#)

- Granlund, S., Kolak, J., Vihman, V., Engelmann, F., Lieven, E. V., Pine, J. M., Theakston, A. L., & Ambridge, B. (2019). Language-general and language-specific phenomena in the acquisition of inflectional noun morphology: A cross-linguistic elicited-production study of Polish, Finnish and Estonian. *Journal of Memory and Language*, 107, 169–194. <https://doi.org/10.1016/j.jml.2019.04.004>

[↑](#)

- Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, 177, 263–277. <https://doi.org/10.1016/j.cognition.2018.04.007>

[↑](#)

- <https://doi.org/10.1017/S0305000914000336>

[↑](#)

- Kidd, E., & Garcia, R. (2022). How diverse is child language acquisition research? *First Language*, 42(6), 703–735. <https://doi.org/10.1177/01427237211066405>

[↑](#)

- Legate, J. A., & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, 14(3), 315–344. <https://doi.org/10.1080/10489220701471081>

[↑](#)

- Levelt, W. J. (2013). *A history of psycholinguistics: The pre-Chomskyan era*. Oxford University Press.

[↑](#)

- Lieven, E. V., Pine, J. M., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language*, 24(1), 187–219. <https://doi.org/10.1017/S0305000996002930>

[↑](#)

- Lillo-Martin, D., & Henner, J. (2021). Acquisition of sign languages. *Annual Review of Linguistics*, 7, 395–419. <https://doi.org/10.1146/annurev-linguistics-043020-092357>

[↑](#)

- MacWhinney, B. (1976). Hungarian research on the acquisition of morphology and syntax. *Journal of Child Language*, 3(3), 397–410. <https://doi.org/10.1017/S0305000900007261>

↵

- MacWhinney, B., & Snow, C. (1990). The child language data exchange system: An update. *Journal of Child Language*, 17(2), 457–472. <https://doi.org/10.1017/S0305000900013866>

↵

- Marchman, V. A., Wulfeck, B., & Ellis Weismer, S. (1999). Morphological productivity in children with normal language and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research: JSLHR*, 42(1), 206–219. <https://doi.org/10.1044/jslhr.4201.206>

↵

- McCauley, S., Bannard, C., Theakston, A.L., Davis, M., Cameron-Faulkner, T., & Ambridge, B. (2021). Multiword units lead to errors of commission in children's spontaneous production: 'What corpus data can tell us?' *Developmental Science*, 24, 6, e13125. <https://doi.org/10.1111/desc.13125>

↵

- McClelland, J. L., & Patterson, K. (2002). 'Words or Rules' cannot exploit the regularity in exceptions. *Trends in Cognitive Sciences*, 6(11), 464–465. [https://doi.org/10.1016/S1364-6613\(02\)02012-0](https://doi.org/10.1016/S1364-6613(02)02012-0)

↵

- Meylan, S. C., Frank, M. C., Roy, B. C., & Levy, R. (2017). The emergence of an abstract grammatical category in children's early speech. *Psychological Science*, 28(2), 181–192. <https://doi.org/10.1177/0956797616677753>

↵

- Newmeyer, F. J. (2016). Form and Function in the Evolution of Grammar. *Cognitive Science*, 41(S2), 259–276. <https://doi.org/10.1111/cogs.12333> ↵
- Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 57, 1247–1257. Advance online publication. <https://doi.org/10.1111/jcpp.12573>

↵

- OpenAI (2023). ChatGPT [Large language model]. <https://chat.openai.com/chat>

↵

- Pearl, L. (2021). Theory and predictions for the development of morphology and syntax: A universal grammar + statistics approach. *Journal of Child Language*, 48(5), 907–936. <https://doi.org/10.1017/S0305000920000665>

↵

- Piantadosi, S. (2023). Modern language models refute Chomsky’s approach to language.

<https://lingbuzz.net/lingbuzz/007180>

↵

- Pine, J. M., & Martindale, H. (1996). Syntactic categories in the speech of young children: The case of the determiner. *Journal of Child Language*, 23(2), 369–395. <https://doi.org/10.1017/S0305000900008849>

↵

- Pine, J. M., Freudenthal, D., Krajewski, G., & Gobet, F. (2013). Do young children have adult-like syntactic categories? Zipf’s law and the case of the determiner. *Cognition*, 127(3), 345–360.

<https://doi.org/10.1016/j.cognition.2013.02.006>

↵

- Pinker, S. (1984). *Language learnability and language development*. Harvard University Press.

↵

- Pinker, S. (1999). *Words and rules*. 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)

↵

- Pozzan, L., & Valian, V. (2017). Asking questions in child English: Evidence for early abstract representations. *Language Acquisition*, 24(3), 209–233. <https://doi.org/10.1080/10489223.2016.1187615>

↵

- Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8(1), 1–56. <https://doi.org/10.1080/01690969308406948>

↵

- Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of mouses in adult speech. *Language*, 89(4), 760–793. <https://doi.org/10.1353/lan.2013.0068>

↵

- Roeper, T., & Williams, E. (Eds.). (1987). *Parameter setting*. Springer. <https://doi.org/10.1007/978-94-009-3727-7>

↵

- Roser, M., & Ortiz-Ospina, E. (2016). Literacy. Published online at [OurWorldInData.org](https://ourworldindata.org/literacy). Retrieved from: <https://ourworldindata.org/literacy>

↑

- Rumelhart, D. E., & McClelland, J. L. (1986). Learning the past tenses of English verbs: implicit rules of parallel distributed processing? In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Erlbaum.

↑

- Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, 12(3), 229–265. [https://doi.org/10.1016/0010-0277\(82\)90033-6](https://doi.org/10.1016/0010-0277(82)90033-6)

↑

- Snyder, W. (2021). A parametric approach to the acquisition of syntax. *Journal of Child Language*, 48(5), 862–887. <https://doi.org/10.1017/S0305000921000465>

↑

- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74(3), 209–253. [https://doi.org/10.1016/S0010-0277\(99\)00069-4](https://doi.org/10.1016/S0010-0277(99)00069-4)

↑

- Tomasello, M. (2003). *Constructing a language*. Harvard University Press.

↑

- Valian, V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology* 22, 562–79

↑

- Wexler, K. (1996). The development of inflection in a biologically based theory of language acquisition. In M. L. Rice (Ed.), *Towards a genetics of language* (pp. 113–144). Lawrence Erlbaum Associates.

↑

- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 106(1-4), 23–79. [https://doi.org/10.1016/S0024-3841\(98\)00029-1](https://doi.org/10.1016/S0024-3841(98)00029-1)

↑

- Yang, C. (2013). Who’s afraid of George Kingsley Zipf? *Significance: The Magazine of the Royal Statistical Society and the American Statistical Society*. Dec. 29-34.

↑