Trends in
Cognitive Sciences

CellPress

Letter Response

Dependency grammar remains the simplest theory of grammar: a reply to Fujita

Edward Gibson [1,*]

In the original article [1], I claimed that dependency grammar is the simplest theory of grammar. All theories of grammar have rules for connecting words together to make larger units. In dependency grammar, these connections are directly between words. In phrase structure grammars, there is slightly more structure than in dependency grammar, consisting of nonterminal categories for the connected elements, such as the noun phrase and the verb phrase. I argued that these two alternatives are somewhat interchangeable. In transformational grammars, there are additional rules such that phrases or words can 'move' elsewhere within the structure of the sentence, leaving behind traces of their original position. I argued that transformational rules add unnecessary complexity to a grammar and hence are unneeded.

In his response [2], Fujita argues that grammars with constituency and transformational rules explain several phenomena better than theories without these components, and he provides five examples.

In response, first, let us observe that transformational rules make the grammar more complex, such that it becomes unlearnable [3,4], thereby forcing the need to postulate innate constraints: what Chomsky referred to as universal grammar. We now have existence proofs—in the form of large language models (LLMs)—that much of human language syntax is learnable from the input, without such language-specific innate constraints. Importantly, even LLMs trained on cognitively plausible volumes of data learn syntactic structures well [5,6]. So the fact that transformational grammars are unlearnable seems like a strike against them. Second, there is potentially some confusion in that the response suggests that dependency structures are not hierarchical: although these grammars do not have nonterminal categories such as noun phrase, the hierarchical relationships among the words are largely the same as in a phrase structure grammar.

Next, let us examine the cases provided in support of grammars that include transformations and constituents.

(i) Nested relative clauses with stranded prepositions [7]. Fujita argues that these complexity effects can be accounted for using additional parsing assumptions within a transformational grammar framework [8]. This is true, but one can always add parameters to fit the data. The phenomena that Fujita discusses [7] follow directly from the observation that long-distance dependencies are costly to process. Nothing more needs to be said within a dependency grammar approach. Parsimony favors this approach.

(ii) Subject island effects. Fujita claims that certain subject island effects are best explained by grammars with transformations [9]. However, neither Fujita nor the authors of the relevant study [9] provide a theory of these subject island effects—with or without transformations—so the evidence is not relevant to transformational rules. Subject island effects instead seem best explained by nonsyntactic (discourse) factors [10]. Specifically, substantial variability in acceptability across constructions—not just in English, but in French and Italian—is not expected under syntactic accounts.

(iii) Sluicing. Fujita, citing a recent study, suggests 'sluicing', a form of clause ellipsis, is best explained by transformational grammar. However, neither Fujita nor the authors of that study consider nontransformational alternatives (e.g., [11]). In fact, sluicing is notoriously difficult to explain within transformational theories.

(iv) Scope ambiguity effects, as in 'I did not leave because it rained.' Fujita notes that this sentence has two interpretations, whereas, when the clauses are inverted, there is only one interpretation (Figure 1). Fujita argues that (i) these effects follow from the constituent structures of the two sentences and (ii) dependency grammar cannot explain this structural asymmetry. However, exactly the same asymmetry is in fact present in dependency representations of the relevant materials (Figure 1).

(v) Empty complementizer effects. Priming evidence is presented that suggests the existence of empty complementizers (corresponding to the word 'that') in examples such as 'The director announced (that) the actor would be in the new movie' [12]. However, in spite of the fact that similar terminology is used (empty complementizers and empty elements), the cases with proposed empty complementizers do not bear on transformational accounts, because they are not assumed to be derived from some other base form. They are simply additional proposed lexical structures for verbs such as 'announce'. Hence this question has no bearing on the claims in [1].

In summary, Fujita's examples are indeed cases in which some authors have proposed transformational rules and/or constituency. But each of these cases (and all others that I have seen) can be handled by dependency rules, without any transformations or constituents. As a result, given the greater complexity of grammars with transformations, Occam's razor suggests that a simpler grammar

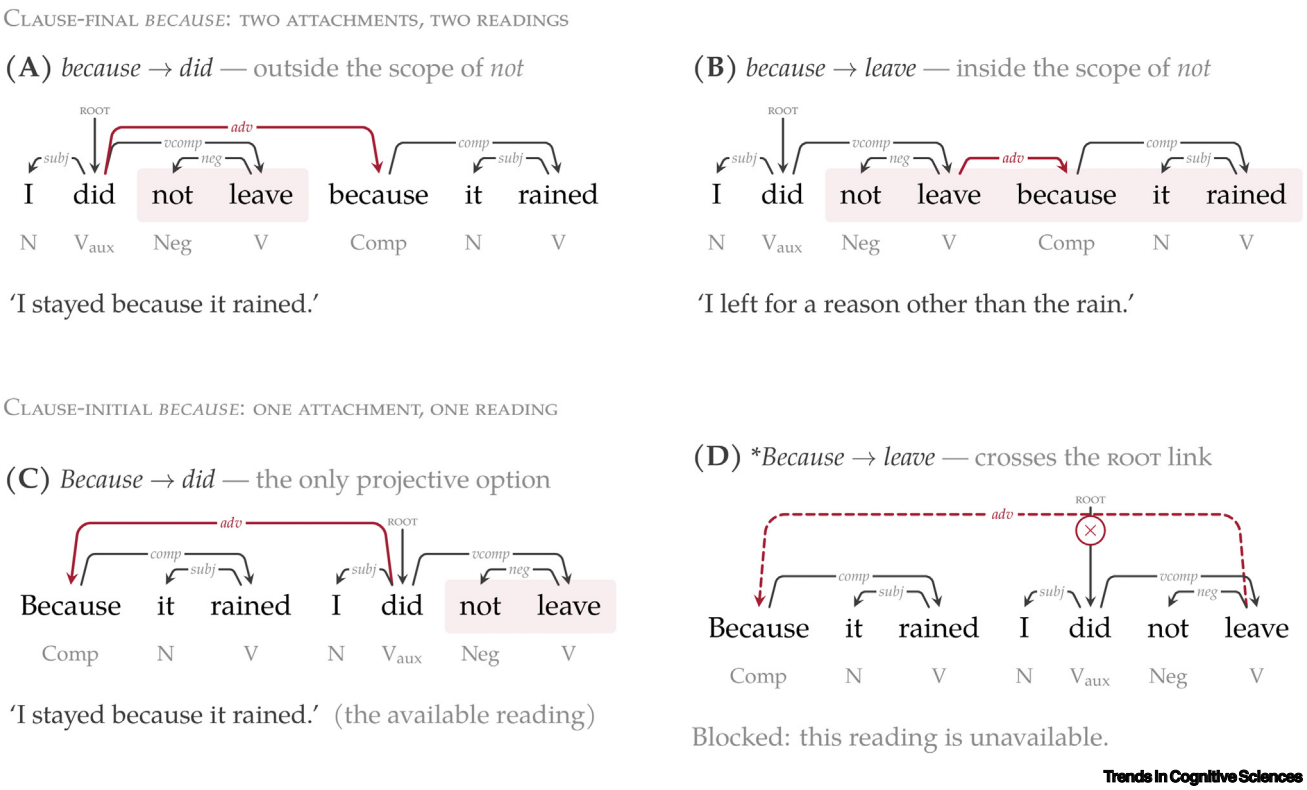


Figure 1. A dependency grammar formalism captures the ambiguity in Fujita's materials. When a clausal modifier ('because it rained') attaches to 'leave', it falls inside the scope of 'not', but when it attaches to 'did', it falls outside the scope of 'not'. In (A), the modifier attaches to 'did', resulting in the meaning 'I stayed because it rained' (outside the scope of 'leave'). In (B), the modifier attaches to 'leave', resulting in the meaning 'I left for a reason other than the rain.' In (C), the modifier attaches to 'did', resulting in the same meaning as in (A). In (D), the modifier cannot attach to 'leave' because this would require a crossed dependency (which is blocked here). Indeed, this is the same explanation for the interpretation effects within a phrase structure analysis. The existence of constituent nodes in phrase structure plays no role in the explanation, contrary to what is implied by Fujita.

formalism, such as dependency grammar, is preferable.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used Claude in order to evaluate the correctness of the claims that were made and to make the figure look better. The author did not use any text generated by Claude and takes full responsibility for the content of the publication.

Declaration of interests

The author declares no competing interests.

[1]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

*Correspondence:
egibson@mit.edu (E. Gibson).

https://doi.org/10.1016/j.tics.2026.09.012

© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

References

1. Gibson, E. (2026) Dependency syntax as the simplest theory of grammar. *Trends Cogn. Sci.* https://doi.org/10.1016/j.tics.2026.03.002
2. Fujita, H. (2026) Evidence supports constituency-based syntax, movement, and unpronounced elements. *Trends Cogn. Sci.*
3. Chomsky, N. (1971) *Problems of Knowledge and Freedom: The Russell Lectures*, Fontana
4. Gibson, E.A.F. (2025) *Syntax: A Cognitive Approach*, MIT Press
5. Warstadt, A. *et al.* (2023) Findings of the BabyLM challenge: sample-efficient pretraining on developmentally plausible corpora. In *Proceedings of the BabyLM Challenge at the 27th Conference on Computational Natural Language Learning*, pp. 1–34, Association for Computational Linguistics
6. Wilcox, E.G. *et al.* (2025) Bigger is not always better: the importance of human-scale language modeling for psycholinguistics. *J. Mem. Lang.* 144, 104650
7. da Cunha, Y. and Gibson, E. (2025) Syntactic complexity phenomena are better explained without empty elements mediating long-distance dependencies. *Cogn. Sci.* 49, e70088
8. Gibson, E. and Hickok, G. (1993) Sentence processing with empty categories. *Lang. Cogn. Process.* 8, 147–161
9. Cartner, M. *et al.* (2026) Subject islands do not reduce to construction-specific discourse function. *Cognition* 271, 106467
10. Abeillé, A. *et al.* (2020) Extraction from subjects: differences in acceptability depend on the discourse function of the construction. *Cognition* 204, 104293
11. Poppels, T. and Kehler, A. (2025) Novel cases of sluicing with mismatched antecedents: theoretical consequences. *Glossa* 10
12. Momma, S. *et al.* (2025) Speakers encode silent structures: evidence from complementizer priming in English. *J. Mem. Lang.* 145, 104671