

Is Chat-GPT a grammatically competent informant?

Hubert Haider

Dept. of Linguistics & Centre for Cognitive Neuroscience, Univ. Salzburg

Abstract

This is a report on Chat-GPT's qualities as a 'native speaker' based on grammar competence tasks on English and partly on German. The chatbot passes the tests in style, be it restrictions on wh-movement such as wh-in-situ restrictions or movement of indirect objects, That-trace-effect, or the Subject condition, as well as other grammatical constraints like the LLC or exceptional subcategorization restrictions (ECM infinitives). The bot's ability to resolve ambiguities is surprisingly good but not perfect. The results not only indicate that properties of grammars previously attributed to a specific human disposition (aka 'Universal Grammar') are amenable to machine learning. Such outcomes are relevant for the "*no-negative-evidence*" issue often appealed to in nativist argumentations.

1. The range of the examined grammaticality assessments

The 'chats' mainly consisted of queries plus answers of the following form: "*Is the following sentence grammatically correct?*" The inquiry started with the following question "*What is the difference between grammaticality and acceptability in linguistics?*" Here is the bot's answer:

"Grammaticality and acceptability refer to different aspects of language evaluation. Grammaticality refers to whether a sentence or utterance follows the rules of grammar in a particular language. [...] Acceptability, on the other hand, refers to whether a sentence or utterance sounds natural or makes sense to a native speaker of the language." (May 1st, 2023).

The above statement sounds like a solid basis for questioning an informant who knows to distinguish between formal well-formedness and pragmatic (in-)appropriateness. Such a background of knowledge would be more than one can usually expect from a willing informant. The following range of linguistic competence areas has been examined:

- ⊗ *That-trace effect in English
- ⊗ *Wh-subject in situ in multi-wh-constructions in English and German
- ⊗ *Wh-in-situ of 'why' or 'how' in English
- ⊗ *Wh-fronting of the indirect object in English double-object constructions
- ⊗ Subject condition in English and German
- ⊗ Left-left constraint in English and German
- ⊗ Idiosyncratic subcategorization requirements (*believe* vs. *expect*)
- ⊗ Identification of ambiguities

As linguists, we are used to refer to rules and principles when we describe the grammatical well-formedness of expressions. This is not what native speakers do when they are supposed to judge a sentence. They have to rely on their gut feeling since their implicit linguistic know-how is not consciously accessible. Native speakers' acceptance ratings do not directly reflect grammatical status but only a general sense of acceptance. This shows, for instance, in the phenomenon of acceptable ungrammaticality and grammatical illusions (Haider 2011, Phillips et al. 2011).

As for the bot, its grammatical competence is not programmed in. It is a result of having been trained on and exposed to vast text corpora. The internal setup of the bot that makes up its grammatical competence is a black box. Chat-GPT is based on a “*foundational large language model*.” The crucial term is *language model*. Put simply, this means that the bot has been trained to write meaningful text on just about anything, without requiring an in-depth understanding of the subject matter. Here is the bot's self-disclosure on its grammatical competence.

Query 1: “How does Chat-GPT decide whether a given sentence is grammatically correct?”



As an AI language model, ChatGPT does not have a fixed set of rules to determine whether a sentence is grammatically correct or not. Instead, it has been trained on a vast corpus of text data, and it uses statistical patterns and probabilities to make predictions about the grammaticality of a given sentence.

In other words, ChatGPT uses statistical analysis and machine learning algorithms to determine the grammaticality of a sentence, based on its training data and the patterns it has learned from that data.

Assessing an utterance as well-formed or not, or ambiguous or not is a specifically human ability and could therefore serve as a Turing test. A generation ago, the typical data-mining technique for a grammarian was either introspection or interviewing native informants and collecting their judgements. Today, we have access to huge corpora and techniques for searching them. But corpora do not tell us whether an utterance is grammatical or not. The absence of a sentence type in a corpus is only an indication but no proof of the ungrammaticality of that sentence type. It cannot be completely ruled out that the absence in a corpus is only due to the low frequency of the respective pattern. This is the well-known dilemma of a linguist studying corpora of a dead language and it is also a challenge for an AI system. Here is a series of queries that may be useful not only for clarifying this problem.

2. The queries

2.1 That-trace effect

The *that-trace* effect is a constraint on the filler-gap relation of wh-constructions with a gap in a special structural position of a clause. It is the very position that matters rather than the particular grammatical function associated with this position, i.e. ‘subject’. This is the reason why the effect is observed in [S[VO]] languages like English or French and in other, structurally similar languages outside of the IE-family, such as Nupe or Wolof (see Pesetsky 2017). These are languages with an obligatory structural position for the subject. In OV languages such as Dutch (1a) or German (1b), a trace following a complementizer does not affect grammaticality,¹ as the following book-corpus excerpts illustrate (see Haider 2010: 88, 128):

- (1) a. Wie_i denk je [dat [e_i hem gestuurd heeft]]?
 who think you [that him sent has]?

¹ This does not exclude that there is a preference for the complementizer-less variant in general in German (and also in English, as (Coward 1997: 19) observed), but this preference does not differentiate between subject- versus object-extraction. A set of corpus data has been compiled by Paul (1919: 321f.) in a subsection named “Satzverschlingung” (sentence convolution).

- b. Wer_i, glauben Sie denn, [dass [e_i den Ausführenden sagt, was sie tun müssen]]?
 who think you PARTICLE [that the performers_{Acc} tells what they do must]?
 ‘Who do you think tells the performers what to do?’

Depending on the theoretical background, the that-trace effect tends to be viewed either as a grammatical restriction or as a by-product of sentence processing.² “*The core pattern, though easily stated, remains a conundrum. It is evident in numbers of unrelated languages, but also appears to vary even in closely related languages.*” Cowart & McDaniel (2021: 258).

The “easily stated” core pattern in English (2a) is a constraint on wh-subjects extracted from a C-introduced clause, with two apparent exceptions noted already by Bresnan (1977: 194 fn.), namely the intervention-effect of adverbs (2b) on the one hand, and the total absence in relative clause formation (2c).

- (2) a. *What_i does he think [*that* [e_i affects everyone]]?
 b. What_i does he think [*that* [in this situation e_i affects everyone]]?
 c. This is an incident [0_i [*that* [e_i affects everyone]]

The unifying property of (2b,c) is the fact that the complementizer is either obligatory (3a) or at least strongly preferred (3b), as in sentences with a fronted PP (see Doherty 2000: 15). So, this is likely to override the that-t-restriction, especially if the restriction is the reflex of a processing effect.

- (3) a. This is an incident *(*that*) affects everyone.
 b. He thinks ?(*that*) in this situation such an incident affects everyone.

Another factor that tends to be underestimated is the fact that extraction from a *that*-clause reduces acceptability in general, regardless of the extraction site (Cowart 1997: 19). In other words, extraction out of a *that*-clause is dispreferred if there is a variant available without *that*. The fact that nevertheless, there is a difference between subject and object extraction reduces to the fact that objects are not C-adjacent in a VO clause structure. Hence, (2b) should not be contrasted with (2a) but rather with object extractions as in (4). In each case, there is material intervening between C° and the extraction site.

- (4) What_i does he think [*that* they need e_i]?


Finally, the *that*-trace restriction is not categorical but gradual, which is typical for processing effects. Sobin (2002: 542) reports the following acceptability scores for (5a) and (5b), respectively. In each case, the “impossible” rating is far away from a 95% level of unacceptability due to ungrammaticality.

- (5) a. Who did you say that saw Elvis last night. 64% good, 27% possible, 9% impossible
 b. Who do they think that might visit the pope. 22% good, 22% possible, 56% impossible

² Schippers et als. (2020: 128) conclude: “*Although COMP-trace violations in German are not categorically ruled out, they are associated with decreased acceptability. [...] We argue that our data is best explained by assuming COMP-trace effects are in essence processing related, which has led to the exclusion of COMP-trace configurations in certain languages, including English.*”

Chat GPT seems to apply the stricter measure. It evaluates the sentences in (5) as grammatically incorrect. The bot identifies the complementizer as the critical item but misidentifies its grammatical category.

Query 2: questioning the subject


 I apologize for the confusion in my previous response. These two sentences are not grammatically correct in English due to the placement of the relative pronoun "that."

Here are the corrected versions of these sentences:

- Who did you say saw Elvis last night?
- Who do they think might visit the pope?


The reason for the apology seems to be that the same question had been asked already earlier in the session.³ The confusion about the grammatical status of "that" shows that the bot is really naïve when it comes to grammar.


As for German, the bot correctly does not reject the test sentence:⁴

 Yes, the sentence "Wer glaubst Du denn, dass die Lösung finden könnte?" is grammatically correct in German. It translates to "Who do you think could find the solution?"

Query 3: questioning the object

The bot confirms the preference for the omission of the complementizer. This is what Cowart (1997: 19) observed with human test subjects, who rate the variants without complementizer generally better irrespective of the grammatical function, that is, subject vs. object.

 H Are the following two sentences grammatically correct in English?
Who did you say that Elvis saw last night?
Who do they think that the pope might visit?

 The first sentence "Who did you say that Elvis saw last night?" is grammatically correct in English, although it could be rephrased for clarity. A better option would be: "Who did you say saw Elvis last night?" This removes the unnecessary relative pronoun "that" and makes the sentence more concise.

The second sentence "Who do they think that the pope might visit?" is also grammatically correct in English, although it could be rephrased for clarity. A better option would be: "Who do they think the pope might visit?" This removes the unnecessary relative pronoun "that" and makes the sentence more concise.


Query 4: self-corrected that-trace query


Interestingly, in some cases, the bot reacts in a way that is familiar from human informants. It

³ At the first time, the bot answered: "The first sentence "Who did you say that saw Elvis last night?" is grammatically correct in English, but it could be rephrased for clarity. For example: Who did you say saw Elvis last night?"

⁴ "Wer glaubst Du denn, dass die Lösung finden könnte"
who think you *particle* that the solution find could? ('Who do you think could find the solution?')

does not react precisely to the test item but to a self-corrected variant of it, with the complementizer removed, as in the following case.


 H Is the following sentence grammatically correct in English? What does she think that cured her disease?


 Yes, the sentence "What does she think cured her disease?" is grammatically correct in English.

It is a complex question that asks about what the person believes was the cure for her disease, where the pronoun "she" refers to the person who holds this belief.

Query 5: acceptance of that-trace chains


The bot does not reject every instance of a that-trace configuration. The following specimens taken from web corpora contain a that-trace wh-chain. The bot rates them as grammatically correct.

 H Are the following two sentences grammatically correct in English?
Sentence 1: Who does he think that will primarily benefit?
Sentence 2: What does he think that makes him what he is?

 Both sentences are grammatically correct in English.

In the following query, the bot correctly identifies the intervention effect of adverbials:


Query 6: That-trace with an intervening adverbial phrase:

 The first sentence, "Who do you think that will answer the question correctly?" is grammatically incorrect. A correct version of this sentence would be: "Who do you think will answer the question correctly?" or "Who do you think is going to answer the question correctly?"

The second sentence, "Who do you think that under these circumstances will answer the question correctly?" is grammatically correct. It is a question asking for an opinion on who would answer the question correctly given certain circumstances.

Finally, and ironically, the bot is steadfastly egalitarian with typical minimal pairs used in textbooks. This behaviour did not change under retesting, without any apologies for being confused as in query 2.

Query 7: minimal pair judgement

 H Are the following two sentences grammatically correct in English?
Who do you think that they invited?
Who do you think that invited them?

 Both of the sentences you provided are grammatically correct in English.


2.2 Wh-subjects in situ in English and German


In languages like English, a *wh*-subject is deviant in-situ if it depends on a preceding *wh*-item (Haider 2010, sect. 3.4). This is what rules out the examples in (6). In Generative Grammar, Chomsky (1973) ruled out such constructions by means of the Superiority condition⁵. In the Minimalist Program, the constraint has been replaced by a constraint on derivations, viz. a Minimal-Link or Shortest-Move condition. *Wh*-subjects in situ as in (6) are not deviant because another *wh*-item has been fronted. They are ungrammatical because a *wh*-operator in a specification position cannot be dependent on a *c*-commanding *wh*-operator, independent of movement (see the discussion of example (9) below).

- (6) a. *What did *who* say?⁶
 b. *I do not remember [what *who* said].

Query 8: what – who

Although the introductory sentence in the bot’s answer to the following query promises something else, the bot correctly identifies the deviant sentence:

 Are the following two sentences grammatically correct in English?
 Sentence 1: I cannot remember what who said?
 Sentence 2: I cannot remember who said what?


 Both sentences have some grammatical errors.

Sentence 1: "I cannot remember what who said?" is grammatically incorrect. It should be "I cannot remember what was said by whom?" or "I cannot remember who said what?"

Sentence 2: "I cannot remember who said what?" is grammatically correct in English.

Query 9: where – who

 Is the following sentence grammatical in English: I cannot remember where who was hiding.

 The sentence "I cannot remember where who was hiding" is not grammatically correct in English. It appears to be a combination of two questions or statements, and the phrasing is unclear. A more grammatically correct version of this sentence could be "I cannot remember where someone was hiding," or "I cannot remember who was hiding and where."

In English, there is a difference between *wh*-pronouns as subjects and phrasal *wh*-subjects, as in (7a,b), compared to (6a,b). This difference tends to be – wrongly – attributed to a pragmatic effect, namely “discourse-linking”. “*Which witness*” presupposes a discourse in which there are witnesses in the common ground. But, as (8) demonstrates, discourse linking is not the crucial factor since “*how many witnesses*” also presupposes a common ground with witnesses.

⁵ “If a transformation can in principle be applied to two constituents in the structure, it has to be applied to the one that is superior.” http://www.glottopedia.org/index.php/Superiority_condition


⁶ For the sake of completeness, let’s note that Bolinger (1978: 130) does not rule out the following examples:
 i. Where did *who* go? ii. Why would *who* accept? iii. When did *who* die?


- (7) a. What did [*which witness*] say?
 b. I do not remember [what [*which witness*] said].
- (8) a. *What did [*how many witnesses*] say?
 b. *I do not remember [what [*how many witnesses*] said].

The relevant difference is this: In (6a,b), the wh-subject is a wh-operator, viz. *who*. In (7a,b), the wh-pronoun is part of an NP and not the subject itself. This is a structural difference. The contrast between (7) and (8) is a contrast in semantic type (see section 2.3). (8a) and (8b) are ruled out by the very same constraint that rules out *how* in-situ: A semantically higher-type wh-expression cannot be interpreted as dependent. In (7), the wh-operators range over sets of <e>-type elements (i.e. individuals) while “how many witnesses” asks for a *property of a set*, viz. its cardinality. This is semantically a higher-order type (Haider 2010: 123-124).

Query 10: Discourse linking?

The bot does not discriminate between ‘which’ and ‘how many’ and the paraphrases it appends shows that the interpretation of ‘how many’ is an individual-level one, namely a “certain number”.

 Are the following two sentences grammatically correct in English?
 Sentence 1: When did which witness arrive?
 Sentence 2: When did how many witnesses arrive?

 Both sentences are grammatically correct in English.

The first sentence, "When did which witness arrive?" is asking for information about the time that a particular witness arrived.


The second sentence, "When did how many witnesses arrive?" is asking for information about the time that a certain number of witnesses arrived.

The constraint against a dependent wh-item in a spec-position (Haider 2010: 113-114; 117) covers also structures that are not covered by the Superiority condition or its successors, such as the Minimal-link condition or a Shortest-move constraint, since in patterns like (9) there are no alternative movement options involved. Such patterns have been recognized as a problem for Superiority already by Chomsky (1981: 236).

(9) I know perfectly well *who* thinks (that) she/**who* is in love with him.

However, this is where the bot has reached its limits, it seems. It doesn't fault Chomsky's example (9), and it does not correctly identify it as a multiple-wh clause, with an interpretation⁷ with the second ‘who’ depending on the preceding one.

Query 11: wh-subject in situ without crossing

 The following sentence is grammatically correct in English:
 "I know who thinks that who is in love with him."

⁷ For which x, for which y: x thinks (that) y is in love with him.


The bot, however, excels when it is asked to compare English with another language that differs from English in this respect. In German and other OV languages, *wh*-elements may be fronted across *wh*-subjects. There are even minimal pairs such as in the following corpus excerpt (10):

- (10) Es ist ja auch zweitrangig *wer was* oder *was wer gesagt hat!*
it is also secondary *who what* or *what who* said has

Here is what the bot returns on the grammatical status of the sentence in German and its translation into English. It differentiates accurately between the situation in English and German in the translation of the German sentence:

Query 12: Wh-subject in situ in German

H Is the following sentence grammatically correct in German? Es ist zweitrangig wer was oder was wer gesagt hat!

 Yes, the sentence "Es ist zweitrangig wer was oder was wer gesagt hat!" is grammatically correct in German. It translates to "It is secondary who said what or what was said by whom!" in English.

2.3 Wh-in-situ with ‘*why*’ or ‘*how*’ in English


There seems to be consensus that ‘*how*’ and ‘*why*’ are unacceptable in situ in multiple *wh*-questions⁸ in English (see Bolinger 1978: 137) and that there is a contrast (Reinhart 1998: 31) between the bare pronouns (11a,b) and their individuated⁹ counterparts (11c,d).

- (11) a. *Who lost *why*?
b. *Who whistled *how*?
c. Who lost for *which reason*?
d. Who whistled *what way*?

The bot fails to identify (11a,b) as deviant and accordingly does not discriminate between (11a,b) and (11c,d), respectively, but interprets their content accurately. For the algorithm, the absence of a pattern in a corpus apparently is not a sufficient reason to classify it as ungrammatical.

Query 13: *why* or *how* in situ (intransitive)

H Are the following two sentences grammatically correct in English?
Who lost why?
Who whistled how?

 Yes, both of these sentences are grammatically correct in English. However, without context, it is difficult to determine their intended meaning. "Who lost why?" could be asking who lost for what reason or purpose, while "Who whistled how?" could be asking who whistled and in what manner or style.

⁸ ‘*How*’ and ‘*why*’ may be used in situ only in echo-questions: You did it why? You fixed it how?

⁹ In (9c,d), the *wh*-item ranges over individual-type variables (= type *e*). *Why* and *how*, however, range over higher than type-*e* variables (corresponding to properties of events and propositions).

On the other hand, the bot rejects *why* in situ in the following example:

Query 14: *why* or *how* in situ (transitive)

H Are the following two sentences grammatical in English? Sentence 1: Who has fixed it how?
Sentence 2: Who has returned it why?



Sentence 1, "Who has fixed it how?" is grammatically correct in English. It is a question in the active voice asking who has performed an action ("fixed it") and how they did it.

Sentence 2, "Who has returned it why?" is not grammatically correct. It appears to be missing a verb between "returned it" and "why." A corrected version of this sentence could be "Why has it been returned?" or "Who returned it and why?"

Retesting shows that the bot's judgement is not grammar-driven, since it accepts the following example in spite of *why* in situ.

Query 15: *why* or *how* in situ (transitive)

H Are the following two sentences grammatically correct:
Who would decorate the table how?
Who would deny the offer why?



Yes, both sentences are grammatically correct in English. The first sentence is asking about who would decorate the table and in what manner or style. The second sentence is asking about who would deny the offer and for what reason or purpose.

In sum, the subtleties of the distribution of '*why*' and '*how*' in English (and other VO languages) seem to escape the bot.

2.4 Wh-fronting in English double-object constructions

Wh-movement of indirect objects in English double object constructions displays a puzzling picture. In *multiple-wh-constructions*, wh-movement of the indirect object (12a) is judged as fully acceptable, cf. Bolinger (1978: 138), Larson (1988: 136). As for single-wh questions, the pattern (12b) is deviant. Erteschik-Shir (1986: 122) or Holmberg et al. (2019, ex. 8a) rate it as ungrammatical. Barss & Lasnik (1986: 348): characterize it as "*less than fully grammatical*".

- (12)a. Who did you give what?
b. *Who did you give the book?¹⁰

In relative clauses, the situation is parallel to the difference w.r.t. the That-trace effect. What is unacceptable for wh-questions is acceptable for wh-relatives. Wh-movement of an indirect object has no negative effect on relative clause formation. The following examples (13) are corpus excerpts.

¹⁰ Corpus search produced a single example:

i. Who did he give a voice? <https://quizlet.com/339167205/john-steinbeck-flash-cards/> [May 1st, 2023]

- (13) a. Shirleen, *who* she gave a small smile before she moved away
b. the only guy *who* she gave a chance to know her on a personal level
c. Mars will stand on a hilltop with others *who* he shows the way.
d. The one *who* he showed the clock first knew it wasn't a bomb.

Query 16: wh-moved indirect object ('give')

The bot signals deviance by proposing a switch from IO to prepositional object.



The sentence "Who did you give the book?" is grammatically correct, but it would be better to include a preposition to clarify the sentence's meaning. For example, "To whom did you give the book?" or "Who did you give the book to?"

Query 17: wh-moved indirect object ('bequeath')

In this query, the bot does not accept wh-movement of the IO.

H

Is the following sentence grammatically correct? Who did he bequeath his fortune?



The sentence is not grammatically correct. A better way to phrase this question would be "To whom did he bequeath his fortune?" or "Who was the recipient of his bequeathed fortune?"

Query 18: wh-moved indirect object ('return')

H

Is the following sentence grammatically correct in English? Who did he forget to return the book?



Yes, the sentence "Who did he forget to return the book?" is grammatically correct in English. It is a question in the past tense, asking about a person who was not returned a book. However, the sentence might sound a bit awkward and it might be clearer to say "Who did he forget to return the book to?" or "To whom did he forget to return the book?"

Query 19: wh-moved indirect object in the multiple-wh question (12a):

H

Is the following sentence grammatically correct? Who did you give what?



The sentence is grammatically correct, but it is a question fragment that requires more context to be fully meaningful. It could be a question asked by someone who did not understand a previous conversation, or it could be a rhetorical question that implies frustration or disbelief.

Query 20: Example (12a) in the variant with 'whom'

The bot is consistent in its preference for "whom" over "who" whenever there is a choice.

H

Is the following sentence grammatically correct? Whom did you give what?



Yes, the sentence is grammatically correct. It is a question that uses "whom" correctly as the object of the verb "give". This question is asking for specific information about who received what, and is a grammatically correct way to ask for this information in formal or informal situations.

Query 21: Indirect object wh-moved in a relative clause (give so. a smile)

The bot is stricter than necessary and insists on “whom” in place of who, but admits it in Query 22 as a grammatical option.



Is the following sentence grammatically correct? The girl who she gave a small smile before moving away is Shirleen.



The sentence is not grammatically correct. A corrected version of the sentence would be:

"The girl whom she gave a small smile before moving away is Shirleen."

Query 22: Indirect object wh-moved in a relative clause (give so. a chance)



Is the following sentence grammatically correct? He is the only guy who she gave a chance to know her on a personal level.



The sentence is grammatically correct, but it could be improved for clarity.

A better way to phrase it would be: "He is the only guy to whom she gave a chance to know her on a personal level."

2.5 Subject condition

The obligatory structural subject position of [S[VO]] languages such as English is known to be an extraction ‘island’. The Subject Condition (Chomsky 1973) rules out the extraction from subjects, and the bot correctly identifies the ungrammatical pattern. Much later, Chomsky (2008, ex. 19) notes that subjects of infinitival constructions such as in Query 24 are not opaque for extraction and judges¹¹ them as acceptable and grammatical.

Query 23: Wh-movement out of the subject of a tensed clause



Are the following two sentences grammatically correct?

Who did he buy a picture of?

Who was a picture of on sale?





The first sentence, "Who did he buy a picture of?" is grammatically correct.

The second sentence, "Who was a picture of on sale?" is not grammatically correct. It seems that the sentence is trying to ask for the subject of a picture that was on sale, in which case a better way to phrase it would be: "Whose picture was on sale?"


¹¹ In Generative grammar, introspection is still an accepted way of data assessment. In psychology, the intuition-istic assessment has been discarded with good reasons already at the beginning of the 20th century.


Query 24: Wh-movement out of an ECM-subject of an infinitival clause:

 H Is the following sentence grammatically correct? Who did he believe a picture of to be on sale?

 The sentence is grammatically correct, but it sounds a bit awkward. A more natural way to express the same idea might be: "Whose picture did he believe was on sale?" or "Which picture did he believe was being sold?"

Query 25: Wh-movement out of the subject of a subject-raising construction


 H Is the following sentence grammatically correct? Of which car is the driver likely to cause a scandal?


 The sentence is grammatically correct. It is a complex interrogative sentence with an embedded question. The main clause is "the driver is likely to cause a scandal" and the embedded question is "of which car." The phrase "of which car" functions as the object of the preposition "of," and it refers to a specific car.

In German, the subject remains in its VP-internal base position (Haider 2010: 24). Therefore, it does not block extraction. The bot correctly identifies (14) as grammatical in German.

(14) Wen_i hätte [e_i damit zu konfrontieren] gereicht?
whom would-have [with-that to confront] sufficed

Query 26: Wh-movement out of an infinitival subject clause (German)

 H Is the following German sentence grammatically correct? Wen hätte damit zu konfrontieren gereicht?


 The sentence "Wen hätte damit zu konfrontieren gereicht?" is grammatically correct in German. It translates to "Whom would it have been enough to confront with that?".

2.6 Left-left constraint in English and German

The Left-Left-constraint (LLC) says that a phrase that is left-adjoined to a left-headed phrase must be head-adjacent to its host phrase (Haider 2022). In English, it applies to adjuncts of NPs as well as VPs since both are head-initial. In German, the NP is head-initial but the VP is head-final. So, the LLC applies only to left adjuncts of NPs, that is, attributive APs. The bot correctly identifies the effect in the queries 27 to 30.

Query 27: LLC with a prenominal attributive AP

 H Is the following sentence grammatically correct? He is a capable of it expert.

 The sentence is not grammatically correct.

A possible correction would be: "He is an expert capable of it." or "He is a capable expert of it."

Query 28: LLC with a prenominal attributive AP in German

H Is the following sentence grammatically correct in German? Er ist ein geeigneter dafür Kandidat.



The sentence "Er ist ein geeigneter dafür Kandidat" is not grammatically correct in German. It seems to be attempting to say "He is a suitable candidate for it," but the word order and choice of preposition are incorrect.

Here are a few possible ways to rephrase the sentence to make it grammatically correct:

- Er ist ein geeigneter Kandidat dafür. (He is a suitable candidate for it.)
- Er ist dafür ein geeigneter Kandidat. (He is a suitable candidate for it.)
- Er ist ein dafür geeigneter Kandidat. (He is a suitable candidate for it.)

Query 29: LLC with a preverbal comparative adverbial AP

H Is the following sentence grammatically correct? He has faster than Bill solved the problem.



The sentence "He has faster than Bill solved the problem" is not grammatically correct. It seems to be attempting to compare the speed at which "he" solved the problem with the speed at which Bill solved the problem, but the comparison is not structured properly.

Here are a few possible ways to rephrase the sentence to make it grammatically correct:

- He solved the problem faster than Bill did.
- He solved the problem faster than Bill.
- He solved the problem more quickly than Bill did.
- He solved the problem more quickly than Bill.

Query 30: LLC with a preverbal comparative adverbial AP

H Is the following sentence grammatically correct in German? Er hat das Problem schneller als Bill gelöst.



Yes, the sentence "Er hat das Problem schneller als Bill gelöst" is grammatically correct in German.

The sentence translates to "He solved the problem faster than Bill" in English, and it follows the correct word order and grammar rules for comparative statements in German.

Query 31: LLC with a prenominal infinitival

In Haider (2020: 201-202), sentence 1 in the following query is regarded as a case of acceptable ungrammaticality. The bot does not differentiate.

H Are the following two sentences grammatically correct?
 Sentence 1: This is an easy to answer question.
 Sentence 2: This is an easy question to answer.



Yes, both sentences are grammatically correct. They have similar meanings, but with a slight difference in emphasis.

Query 32: LLC with a prenominal infinitival attribute with a final adverbial

It is easy to see that the attribute is rejected once an adverbial is added. This shows that the verb in the ‘easy-to-V’ constituent is subject to the LLC. Details in Haider (2022: 202).



Are the following two sentences grammatical in English? Sentence 1: This is an easy to answer correctly question. Sentence 2: This is an easy question to answer correctly.



The second sentence, "This is an easy question to answer correctly," is grammatically correct and the more commonly used structure.

The first sentence, "This is an easy to answer correctly question," is not grammatically correct. It may be rephrased as "This is an easy question to answer correctly" or "This is an easy-to-answer question, correctly."

Query 33: Grammatical illusion with a ‘fake’ phrase-final adjectival head.



Are the following sentences grammatically correct?
 Sentence 1: This is a higher than average percentage.
 Sentence 2: This is a higher than last year percentage.



Sentence 1: "This is a higher than average percentage" is grammatically correct.

Sentence 2: "This is a higher than last year percentage" is not grammatically correct.

In the examples of the query, the head of the AP, viz. “higher” is not adjacent. Nevertheless, the first sentence is accepted by informants, some linguists and the bot. This is a case of acceptable ungrammaticality. In German, agreement reveals that the wrong head is inflected. In (15a,c), the V-adjacent head is inflected, but this is not the head of the attributive AP. The contrast between the two query sentences is clear. ‘Average’ as the V-adjacent adjective can be mistaken as the head, but ‘last year’ cannot.¹² The same is true for ‘expected’ in (15b).

- (15) a. ein [höher als durchschnittlicher] Prozentsatz
 a [higher than average_{Ag}] percentage
 b. an earlier than expected intervention


¹² Another way of avoiding the structural dilemma would be this: “*This was a higher than last year's team finish.*” Here, the illicit attribute is replaced by a licit pre-nominal genitive construction, which is not subject to the LLC since it is in a spec position, viz. the spec of the D° head associated with the Saxon genitive “s”.

- c. ein [früher als erwartetes] Eingreifen
 an [earlier than expected_{Agr}] intervention

Query 34: Attribute with inflection on the wrong, but V-adjacent head

The following test sentence is a corpus excerpt. The bot answers like many native informants.

H Is the following German sentence grammatically correct? Ein früher als erwartetes Eingreifen der Federal Reserve verhinderte Schlimmeres.


 Yes, the sentence "Ein früher als erwartetes Eingreifen der Federal Reserve verhinderte Schlimmeres" is grammatically correct in German. It translates to "An earlier than expected intervention by the Federal Reserve prevented worse things from happening."

Query 35: Attribute with inflection both on the AP head and on the wrong, but V-adjacent head.

Examples like in the following query can be found in corpora of spontaneous speech. When some speakers realize that they are in a grammatical conflict, they try to obey each of the two conflicting rules. First, they standardly put the agreement on the head of the AP and second, they feel obliged to have a head-final head, which, if it were the head of AP, would have to agree with the noun. So, they put agreement on both adjectives, and the bot accepts it:

- (16) *ein [besseres_{Agr} als erwartetes_{Agr}] Ergebnis
 a better than expected result

H Is the following sentence grammatically correct in German? Das ist ein besseres als erwartetes Ergebnis.

 Yes, the German sentence "Das ist ein besseres als erwartetes Ergebnis" is grammatically correct. It translates to "This is a better than expected result" in English.


2.7 Idiosyncratic subcategorization requirements (*believe* vs. *expect*)

In English, the verb ‘believe’ is an outlier. It cannot form a control construction with an infinitival clause (17a), which is the default option for other verbs of this semantic class (cf. *assume*, *imagine*, *suppose*). ‘Believe’ as a verb that selects an ECM infinitive (17b) is exceptional since in all other Germanic languages its equivalents are control verbs. ‘Expect’, on the other hand, is both, a control verb (17c) and a verb that alternatively, takes an ECM-infinitival construction as complement (17d). The bot correctly identifies the different complementation requirements.

- (17)a. *He believed [to be in charge].
 b. He believed [himself to be in charge].
 c. He expected [to be in charge].
 d. He expected [himself to be in charge].

Query 36: Control infinitivals with *believe* and *expect*.

H Are the following two sentences grammatical in English?
Sentence 1: He believed to be in charge.
Sentence 2: He expected to be in charge.


 Only the second sentence is grammatically correct.

Sentence 1 ("He believed to be in charge") is not grammatically correct because "believed" is a transitive verb that requires an object, but in this sentence, there is no object for "believed". Instead, you could rephrase the sentence to say "He believed he was in charge."

Sentence 2 ("He expected to be in charge") is grammatically correct. The verb "expected" is followed by an infinitive phrase ("to be in charge") that functions as the direct object of the verb.

Query 37: ECM infinitival complements of *believe* and *expect*

H Are the following two sentences grammatical in English?
Sentence 1: He believed himself to be in charge.
Sentence 2: He expected himself to be in charge.


 Yes, both sentences are grammatically correct in English.

2.8 Identification of ambiguities

The ability to identify ambiguities is part of the tacit linguistic knowledge, that is, the linguistic competence of native speakers. The bot masters various types of structural ambiguity. The only structures that it fails to identify is comparative ellipsis (query 42 and 43) and the ambiguity between a relative clause and a complement clause, both introduced by 'that' (query 44).

Query 38: ambiguous PP attachment

H What are the two different meanings of the following sentence which is structurally ambiguous. He is attacking a man with a golf umbrella.

 The sentence "He is attacking a man with a golf umbrella" is structurally ambiguous because it can be interpreted in two different ways, depending on which word is emphasized or how the sentence is parsed:

1. The first interpretation is that "he" is using a golf umbrella as a weapon to attack a man. In this case, the emphasis would be on "golf umbrella," indicating that the type of umbrella being used is significant to the action. This interpretation suggests that the golf umbrella is being used as a makeshift weapon.
2. The second interpretation is that "he" is attacking a man who happens to be holding a golf umbrella. In this case, the emphasis would be on "a man," indicating that the identity of the victim is significant to the action. This interpretation suggests that the golf umbrella is incidental to the attack and might not be used as a weapon.

Query 39: Bound pronoun vs. deictically interpreted pronoun

H What are the two different interpretations of the following sentence? Nobody doubts that he is able to win.

The sentence "Nobody doubts that he is able to win" can also be interpreted in two different ways, depending on whether "he" is being used as a gender-specific pronoun or a generic pronoun.

The first interpretation is that "he" refers to a specific male person, and nobody has any doubts that this person is able to win. In this case, "he" is being used as a gender-specific pronoun, and the sentence is referring to a particular male individual.

The second interpretation is that "he" is being used as a generic pronoun to refer to anyone, regardless of gender, who is able to win. In this case, the sentence means that nobody has any doubts that whoever is being referred to is able to win. This interpretation assumes that "he" is being used as a non-gendered pronoun.

Query 40: Sloppy identity

H Is the following sentence ambiguous? Bill hugs his wife and John does too.

Yes, the sentence "Bill hugs his wife and John does too" is ambiguous because it is not clear whether John is hugging Bill's wife or his own.

Query 41: Comparative construction

H Is the following sentence grammatically correct? He spends more money than he earns.

Yes, the sentence "He spends more money than he earns" is grammatically correct. It is a simple sentence with a subject "he" and two verb phrases "spends" and "earns" linked by the comparative conjunction "more". The sentence expresses the idea that the subject spends more money than he earns, which could indicate that he has financial problems.

Query 42: Ambiguous comparative construction (with or without ellipsis)

The bot does not identify the possible but implausible reading¹³ that the speed of aging is compared with the speed of thinking.

H Is the following sentence ambiguous? People age faster than they think.


The sentence "People age faster than they think" is not inherently ambiguous, as there is only one possible interpretation of the sentence.

¹³ The German version of the following test sentence was the slogan of a life insurance company. Someone had added a handwritten comment on the advertisement in a train compartment: "*Erstaunlich, wie langsam die Leute denken!*" (Amazing how slowly people think!).

Query 43: Comparative (with VP ellipsis)

The bot also resists being forced to identify the reading based on an ellipsis, as the following query shows. The paraphrase does not match the meaning of the sentence.


H What does the following sentence mean? He speaks faster than he thinks he does.

 The sentence means that the person being referred to speaks quickly, at a faster pace than the speed at which he can form his thoughts. In other words, he may have difficulty keeping up with his own speech or he may find himself making mistakes or saying things that he did not intend to say because he is not able to think through his words quickly enough.

Query 44: complement clause vs. relative clause

The bot fails to identify the ambiguous grammatical function of the *that*-clause, as object clause and relative clause, respectively. It focuses on the interpretation variants of the pronoun instead.

H Is the following sentence ambiguous? Nobody informed the team that he has cheated.

 Yes, the sentence "Nobody informed the team that he has cheated" is ambiguous. It is unclear who has cheated and who was not informed.

The sentence could have two different interpretations, depending on the intended meaning:

1. Nobody informed the team that someone (other than the team) has cheated: In this interpretation, "he" refers to a third-party individual who has cheated, and the team was not informed of this fact.
2. Nobody informed the team that a member of the team has cheated: In this interpretation, "he" refers to a member of the team who has cheated, and the team was not informed of this fact.

3. Final assessment

Considering that one mustn't compare the bot to a professional linguist, but to a native-speaker informant, the final judgment verges on perfection. The appropriate benchmark would be the mean score on a standardized test calibrated to native-speaker informants. Given the performance in the above queries, chances are the bot will at least perform no worse than the average human informant.

In all, the bot has a good chance to pass the Turing test in the category grammatical competence of a native speaker. The answers are generally within the range of what we would expect from a human native speaker. The bot is never false negative, that is, it does not judge clauses as ungrammatical that are grammatical. This shows also in some cases in which the bot is less restrictive than it could or should be, given the recognized classification of the respective data. But even then the answers of the bot are within the range of variation that one might be willing to concede to a native speaker. Here is a summary of the queries:

Summary of the queries:

	False positive	False negative
That-trace	partially	no
Wh-subject	no	no
‘Why’ or ‘how’	yes	no
Wh-double-objects	no	no
Subject condition	no	no
Left-left constraint	partially	no
Believe vs. expect	no	no
Ambiguities	no	partially ¹⁴

False positive: Bot judges a sentence as grammatical that is in fact ungrammatical.

False negative: Bot judges a sentence as ungrammatical that is in fact grammatical.

The final word has the bot:

Query 45:



What is a reasonable test design for finding out whether Chat-GPT masters English as well as a native speaker?



To test whether Chat-GPT masters English as well as a native speaker, a reasonable test design would involve several steps:

1. Develop a set of test tasks that cover a range of language skills, such as grammar, vocabulary, comprehension, and writing. The tasks should be challenging but not so difficult that they are impossible to complete.
2. Gather a group of native English speakers to serve as a control group. Ideally, the control group should be matched to the Chat-GPT in terms of age, education, and other relevant factors.
3. Administer the test tasks to both the Chat-GPT and the control group. The test should be administered under the same conditions for both groups, with the same time limits and instructions.
4. Score the tests using a standardized rubric or set of criteria. The rubric should be designed to assess the language skills being tested in a consistent and objective manner.
5. Compare the scores of the Chat-GPT and the control group. If the Chat-GPT scores as well as the control group, this would provide evidence that it has mastered English to a similar degree as native speakers.

¹⁴ In some cases, the bot does not recognize the ambiguity of an ambiguous sentence.

Bibliography

- Barrs, Andrew and Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17: 347-354.
- Bolinger, Dwight. 1978. Asking more than one thing at a time. In Henry Hiz (ed.) *Questions*. 107-150. Dordrecht: Reidel.
- Bresnan, Joan. 1977. Variables in the theory of transformations. In Peter W. Culicover, Thomas Wasow and Adrian Akmajian (ed.) *Formal syntax*. 157-196. New York: Academic Press.
- Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson and Paul Kiparsky (eds.) *A festschrift for Morris Halle*. 232-286. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero and María-Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory*. 133-166. Cambridge: MIT Press.
- Cowart, Wayne. 1997. *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks, CA: Sage.
- Cowart, Wayne & Dana McDaniel. 2021. *The That-Trace Effect*. In Grant Goodall (ed.), *The Cambridge Handbook of Experimental Syntax*. 258-277. Cambridge: Cambridge University Press.
- Doherty, Cathal. 2000. *Clauses without “that”: The case for bare sentential complementation in English*. New York: Garland.
- Erteschik-Shir, Nomi. 1986. Wh-questions and focus. *Linguistics and Philosophy* 9: 117-149.
- Haider, Hubert. 2010. *The syntax of German*. Cambridge: Cambridge University Press.
- Haider, Hubert. 2011. Grammatische Illusionen. *Zeitschrift für Sprachwissenschaft* 30: 223-257.
- Haider, Hubert. 2022. The Left-Left Constraint – A structural constraint on adjuncts. In Freywald, Ulrike & Simon, Horst J. (eds.) *Headedness and/or grammatical anarchy?* 177-210. Berlin: Language Science Press.
- Holmberg, Anders, Michelle Sheehan, Anglia Ruskin and Jenneke van der Wal. 2019. Movement from the double object construction is not fully symmetrical. *Linguistic Inquiry* 50(4): 677-721.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19: 335-391
- McDaniel, Dana & Cecile McKee & Wayne Cowart & Merrill F. Garrett. 2015. The role of the language production system in shaping grammars. *Language* 91(2): 415 - 441.
- Paul, Hermann. 1919. *Deutsche Grammatik*. Vol. III. *Syntax*. Halle/Saale: Niemeyer.
- Pesetsky, David. 2017. Complementizer-Trace effects. In Martin Everaert and Henk van Riemsdijk (eds.) *The Wiley Blackwell Companion to Syntax*. 2nd ed. Oxford: Wiley-Blackwell.
- Phillips, Colin, Matthew W. Wagers and Ellen F. Lau 2011. Grammatical illusions and selective fallibility in real-time language comprehension. In Jeffrey Runner (ed.) *Experiments at the interfaces*. 147-180. Bingley: Emerald Publications. (Syntax & Semantics 37).
- Reinhart, Tanya. 1998. Wh-in-situ in the framework of the Minimalist Program. *Natural Language Semantics* 6: 29–56.
- Sobin, Nicholas. 2002. The Comp-trace effect, the adverb effect and minimal CP. *Journal of Linguistics*. 38: 527 - 560.
- Schippers, Ankeliën & Margreet Vogelzang & David Öwerdieck. 2020. COMP-trace effects in German: the role of processing. *Nordlyd* 44(1): 117-132 (Tromsø: The Arctic University of Norway).