Ling 566
Dec 8, 2020

Variation in the English Auxiliary System

Overview

Reminder:
Course evals!

- AAL copula absence
- Why it's not phonological deletion
- Alternative syntactic analyses
- The winner: An empty element (!)
- Reflection on syntactic argumentation
- Reading questions
- If time: More "untangle this"

Linguistic Argumentation

Poll!

- The available data usually underdetermines the analysis (cf *to*)
- Sometimes appeals to naturalness can help
- Further constraints come into play when we try to make interacting analyses consistent
- Still, just about everything could be done differently if we're willing to change assumptions
- Data underdetermines the theory; difficult to argue that something must be analyzed a certain way

An Unusual Case

- The verbless sentences in Chapter 15 provide a rare example where the data seem to force a particular kind of analysis
- Specifically: an empty element
- And we tried **very** hard to avoid it

Notes on African American Language

- aka AAE, AAVE, Ebonics, Black English, and various other things
- All natural languages are systematic
- This is just as true of stigmatized varieties as of prestige dialects
- The claim that AAVE has “no discernible rules” (columnist William Raspberry) is blatantly false
- This is not to deny the social and economic value of using a prestige dialect
- But prestige is not correlated with systematicity

Missing *be* in AAL

- Some AAL sentences:

  Chris at home

  We angry with you

  You a genius

  They askin for help

- Like GAE sentences with a form of *be* missing

- Analogous sentences occur in many languages

AAL Also Allows Sentences With *be*

Chris at home	*Chris is at home*
We angry with you	*We're angry with you*
You a genius	*You are a genius*
They askin for help	*They're askin for help*

Labov's Deletion Account

- Copula absence comes about when contracted auxiliaries (*'s* and it *'re*) are deleted altogether
- Predicts that copula absence is only possible where contraction is: (strong claim)

You got to be good, Rednall!
**You got to ∅ good, Rednall!*

Be nice to your mother!
**∅ Nice to your mother!*

It ain't a flower show, is it?
**It ain't a flower show, 's it?*
**It ain't a flower show, ∅ it?*

Counterexamples to Labov's Account

How old you think his baby is
**How old you think his baby 's*
How old you think his baby ∅

Tha's the man they say is in love
**Tha's the man they say 's in love*
Tha's the man they say ∅ in love

- The relevant examples here are with fully contracted *'s*
- These examples show that copula absence can't depend on copula contraction

Our Challenge

- Provide a precise analysis of AAL copula absence within our theory
- Account for all of the facts covered by the deletion account
- Deal with the counterexamples to the deletion account

Two Possible Analyses

1. Add another initial symbol which is [HEAD [PRED +]], not [HEAD *verb*]:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} pos \\ \text{PRED} & + \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \langle\,\rangle \end{bmatrix} \end{bmatrix}$$

2. Write a special grammar rule for verbless clauses:

$$\begin{bmatrix} phrase \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} verb \\ \text{FORM} & \text{fin} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \text{prop} \\ \text{INDEX} & \boxed{2} \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{1}\text{NP} \\ \text{CASE} & \text{nom} \\ \text{AGR} & non\text{-}1sing \end{bmatrix} \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{PRED} & + \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\,\boxed{1}\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & \boxed{2} \end{bmatrix} \end{bmatrix}$$

A Counterexample to Both:

How old you think his baby ∅

- LDDs require that a non-empty GAP list be licensed by a lexical head that is missing an argument
- Neither the initial symbol analysis nor the grammar rule analysis posits a lexical head corresponding to *is* that would license the gap
- If we posit a silent variant of finite forms of *be*, we solve this problem

The Silent *be* Analysis

Silent *be* Lexical Rule

$$\left[\begin{array}{l} \textit{i-rule} \\ \text{INPUT} \quad \left\langle \text{be} \ , \ \text{X} \right\rangle \\ \\ \text{OUTPUT} \quad \left\langle \phi \ , \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \text{AGR} & \textit{non-1sing} \\ \text{FORM} & \text{fin} \\ \text{INV} & - \end{array}\right] \end{array}\right] \right\rangle \end{array}\right]$$

- This is a highly specialized lexeme-to-word rule (i-rule)

Some Questions About This Rule

Silent *be* Lexical Rule

$$\left[\begin{array}{ll} \textit{i-rule} & \\ \text{INPUT} & \left\langle \text{be} \text{ , } \text{X} \right\rangle \\ \text{OUTPUT} & \left\langle \phi \text{ , } \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \text{AGR} & \textit{non-1sing} \\ \text{FORM} & \text{fin} \\ \text{INV} & - \end{array}\right] \end{array}\right] \right\rangle \end{array}\right]$$

QUESTION	ANSWER
Which lexemes does it apply to?	Those spelled *be*
Why is the output [FORM fin]?	**You got to ∅ good*
Why is the output AGR *non-1sing*?	**I ∅ hungry.*
Why is the output [INV –]?	**It ain't a flower show, ∅ it?*

How does this account for LDDs?

Silent *be* Lexical Rule

$$\left[\begin{array}{ll} \textit{i-rule} & \\ \text{INPUT} & \left\langle \text{be} \, , \, \text{X} \right\rangle \\ \text{OUTPUT} & \left\langle \phi \, , \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \text{AGR} & \textit{non-1sing} \\ \text{FORM} & \text{fin} \\ \text{INV} & - \end{array}\right] \end{array}\right] \right\rangle \end{array}\right]$$

Answer: The usual way. That is, the output of this rule (silent *be*) can have a non-empty GAP list. The fact that the verb is not pronounced doesn't matter.

A Possible Objection

Poll!

- Earlier, we touted the WYSIWYG character of our theory: everything justified by something observable.
- Doesn't positing an inaudible verb undermine that claim?
- Response
  - A word with no phonology is just the shortest possible word
  - Positing one such word, with restricted distribution is qualitatively different from allowing multiple "empty categories" that can appear in many places

Conclusions

- Studying a variety of languages and dialects is important to discovering what formal devices are necessary to account for natural language
- Formulating a precise theory of grammar allows us to investigate in detail the differences between dialects and between languages
- We were able to make the argument for a silent verb because our analyses were precise, and the consequences could be worked through

Overview

- AAL copula absence
- Why it's not phonological deletion
- Alternative syntactic analyses
- The winner: An empty element (!)
- Reflection on syntactic argumentation
- Reading questions
- If time: More "untangle this"

Reading questions

- The AUX feature allows us to account for non-auxiliary verbs that exhibit the NICE properties, but are there verbs that exhibit some of the NICE properties and not others?

- *I don't want no business with you*, *Ain't no need in being greedy*, *ain't nobody got time for dat*. As well as negative inversion *can't nobody go out tonight*. This feels to me asks for a modification of the construct of NICE properties to accurately licensing these sentences?

- Can we think of "defective" (using Jackendoff's terminology) lexical entries in terms of defeasibility?

Reading questions

- As a non-American reader, the fact that it is a variety of English with systematic rules would have sufficed, and I found it interesting that there was more than a page allocated to it. Is there any other reason that I am missing here, or is the legitimacy of AAVE under question because linguistically (and descriptively) it could lack something to make it worth studying here?

- One of the awesome things about language is how many different ways it can be used and it was very cool to dive into an analysis on AAVE. Prescriptive grammar is very instilled in the US school system. How can we encourage those who have been taught to be prescriptive grammarians to be accepting of other dialects rather the usual judgmental reaction to non-SAE dialects?

Reading questions

- What is a partial contraction of the copula? We get two syllables without it just being "is"? Like Christina, 12c sounds fine for me, and I was wondering if I just think the vowel is dropped but it isn't entirely. But something like "That's the man they think's in love" is actually fine for me as well, and the vowel is definitely gone there, right?

- Why is it necessary to have full contraction as a prerequisite for copula absence, as opposed to also allowing contractions that still have two syllables as in 12(c)?

Reading questions

- A common refrain in this chapter was, "Our theory of grammar does not allow any operations that destructively modify feature structures."

- Are there any major theories which do? I ask because Labov's deletion account is compelling from an "evolutionary" perspective, e.g.,

  (i) I do not know what you are up to.

  (ii) I don't know what you're up to.

  (iii) I dunno what you up to.

- It seems to me we could account for the examples which allegedly contradict this theory (e.g., "That's the man they say in love") by suggesting these constructions evolved from people who learned the construction in (iii) first and applied it to other constructions.

- Relatedly, how valuable is such speculation on the evolution of a grammar to linguistic analysis?

Reading questions

- 15.3.6 says "a silent word could appear anywhere in the string, and a parser (human or computer) must always allow for that possibility". First, what evidence is there that humans "check" for silent words. Second, is the possibility of silent words a problem? It doesn't seem any more of a problem than the possibility that have could be either AUX + or -?

- This chapter mentions that our grammar does not have operations that destructively modify representations. The analysis of ellipsis from Chapter 13 seemed somewhat destructive to me, is that a different kind of destructive than is being talked about here?

Reading questions

- I am just curious where the three possible analyses (revised initial symbol, a new phrase structure rule, and silent forms) originated from. Do linguists generally agree with Labov's treatment of zero copula and that the silent form is the best method?

- If we spoke a language that was identical to AAVE but allowed only zero-copula in places where zero-copula is permitted in AAVE, would we still have come up with this analysis? I guess the underlying reason this analysis is useful is because the phenomenon of null-copulas is very closely linked to the usage of copula itself. It still seems strange to me though to have an analysis that assumes the existence of an empty word.

Reading questions

- I thought the silent copula rule and using an empty phonological form to avoid having to delete something in the syntax was another classic sneaky-subtle HPSG moment. I feel as though this rule makes somewhat of an assumption that dropping the copula "marked" and that there must be an underlying unexpressed syntactical element. If our analysis is that it would be more "marked" for a word/phonological form to be present than for it to be "empty string", is it possible to have an i-rule than essentially does the opposite and has empty string as an input and a non-empty phonological form as the output?

Reading questions

- One of the arguments for HPSG was that it is surface-orientated. Doesn't allowing a silent form of a lexeme to exist contradict this? Doesn't this lead to underlying representations?

Reading questions

- Is the inability to account for LLDs the main motivator for not using the modified initial symbol for handling the silent copula, or just one example?
- Is there any particular psycholinguistic work supporting the idea that phonetically absent words still exist in cognition? It feels like it would be pretty taxing for processing if that were the case.
- What does it mean for the silent copula treatment to be “strongly lexicalist”?

Reading questions

- In reference to “We therefore depart from strict surface-orientation only when faced with data that admit no other analysis”, I’m a little unclear on what exactly makes the Initial Symbol and Phrase Structure Rule analyses surface-orientated. Is it that a silent copula analysis is considered “abstract” whereas the other two analyses posit actual changes in the feature-structures? If so, what about having silent copula is considered “abstract”?

Reading questions

- Do we use the silent analysis in languages that allow pro-drop?
- I'm curious if a similar silent lexical rule is used for complementizer "that".

Reading questions

- I know we don't really talk about tense, but if we did would the Zero Copula Rule include a constraint for present tense somewhere in the phrase structure?

- Can we revisit what [PRED +] means? The impression I had from revisiting that recently that it boils down to "can follow the copula" but we can have "(Cassius Clay | That man | The greatest boxer ever) is Mohammad Ali", so it has to be more than that.

- If the Semantic Compositionality Principle doesn't apply to the Zero Copula Rule, how do we know in example (21) which values on S's RESTR list belong to which daughter?

Reading questions

- Do we use the same silent copula analysis for other languages that omit the copula or allow for optional copulas (Hungarian, Mandarin, Russian, etc.)?

Reading questions

- As someone who hasn't heard much AAVE, I could still understand the meanings of the AAVE sentences without much problem. More generally, I see how variants of English even extend to language use on social media where the bounds of syntax are weak at best, and still many people would have little trouble understanding the semantics of such sentences. My overarching question is why is it that we are able to understand the semantics of dialectic English (with little prior exposure) and how might the syntactic representation of English dialects help computers to do the same?

Reading questions

- Is there a mechanism in HPSG for restricting certain phrase structure rules, lexical rules, lexical entries etc. to one particular variety (or register) of a language? I know that there are certain syntactic phenomena that I use very commonly in casual speech (like quotative like) that probably wouldn't be permissible in the SAE that the textbook describes.

- When I see the Silent Be Lexical Rule, it makes me think that there could be maybe not an infinite but a really huge amount of different lexical rules to encompass all the different patterns across English dialects, especially if one is taking into account different countries' Englishes. How do syntacticians manage the seemingly endless list of possibilities Lexical Rules when working on grammars?

Reading questions

- How do we detect whether or not a sentence is an AAVE in the first place, and not marking it ungrammatical?
- For a less technical question, how are language varieties handled in the application of grammars? For example, an information retrieval system built with SAE in mind will reject/misinterpret the query "Where the Eiffel Tower at?". But an IR system built with AAVE in mind could reject/misinterpret SAE queries. Language production will also depend on the language variety chosen. Should the the system be built with two completely separate grammars?

Reading questions

- In practice, for similar dialects and languages (or also for representing language changes over time), is it common to use the approach in the chapter of handling variation with rules that go from one to the other? How different would they have to be to just be easier implementing two grammars? Does this have to do with, "operations that destructively modifies feature structures"?

- If the standard variety of a language requires to obey the SHAC principle, but other varieties of the same language do not have SHAC constraints, how does the grammar incorporate that variation?

Reading questions

- I'm thinking about the role NLP tools can play in reinforcing prescriptive attitudes, and I'm wondering what some best practices are for trying to lessen this. For example, if you create a grammar checker for SAE, what's the best way to make it clear to users of the grammar checker that it is designed for only one dialect and that sentences it marks as ungrammatical could be grammatical in other dialects?

- By incorporating more vernacular into NLP systems beyond just prescriptive grammars of a language could we not also run into the issue of making it very hard for our systems to detect ungrammatical sentences?

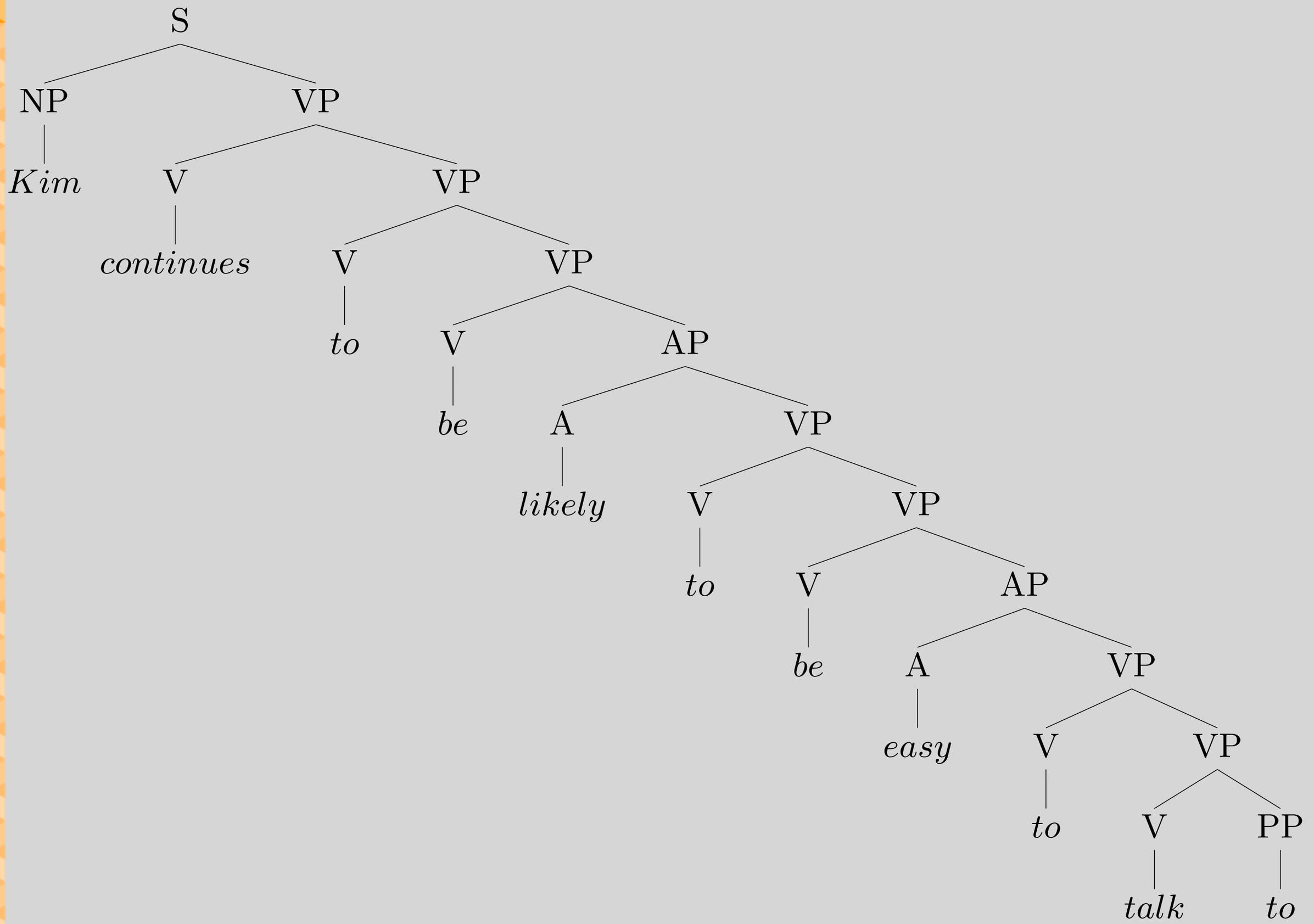
Complicated example #6

Kim continues to be likely to be easy to talk to.

**Kim continue to be likely to be easy to talk to.*

**Kim continues to be likely to is easy to talk to.*

**Kim continues to Kim be likely to be easy to talk to.*



Complicated example #7

That cake, Kim thought would be easy to eat.

**That cake, Kim thought would be easy to eat pie.*

**That cake, Kim thought would be easy to eaten.*

**Cupcake, Kim thought would be easy to eat.*

**That cake, Kim thought that would be easy to eat.*