Ling 566
Dec 5, 2017

Variation in the English Auxiliary System

1

Overview

- AAVE copula absence

- Why it's not phonological deletion

- Alternative syntactic analyses

- The winner: An empty element (!)

- Reflection on syntactic argumentation

- Reading questions

- Course evals

Linguistic Argumentation

- 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

3

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 Vernacular English

- aka 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 AAVE

- Some AAVE sentences:

   Chris at home

   We angry with you

   You a genius

   They askin for help

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

- Analogous sentences occur in many languages

AAVE 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*

7

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?

8

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

9

Our Challenge

- Provide a precise analysis of AAVE 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:

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

11

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
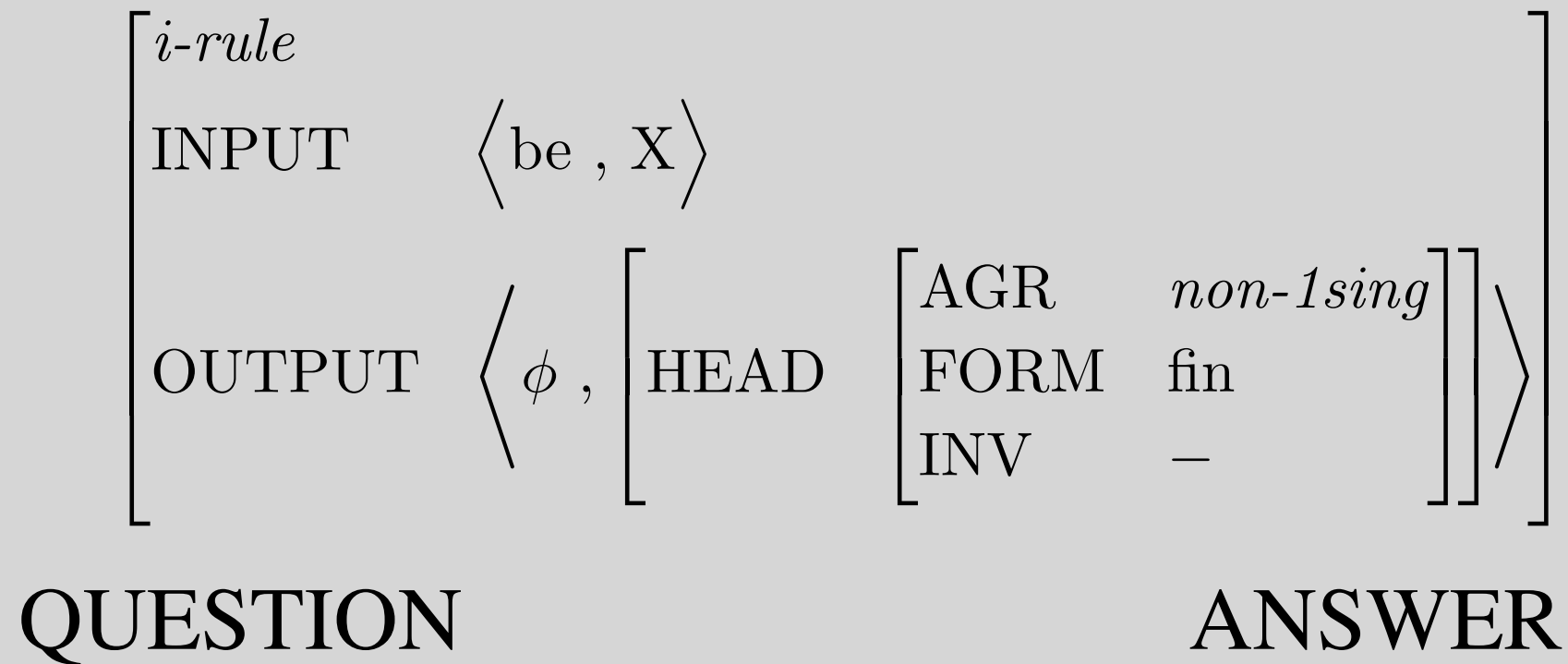
The Silent *be* Analysis

Silent *be* Lexical Rule

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

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

Some Questions About This Rule

Silent *be* Lexical Rule

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

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?*

14

How does this account for LDDs?

Silent *be* Lexical Rule

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

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

- 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

- AAVE copula absence

- Why it's not phonological deletion

- Alternative syntactic analyses

- The winner: An empty element (!)

- Reflection on syntactic argumentation

- Reading questions

- Course evaluations

Reading Questions

- For different variations of language, how do people decide which one is "standard" and which ones are "dialect"? Is it based on the number of people who speak that variation of language?

- How do we define the grammaticality of English? I agree that native speaker's intuition is one important kind of empirical confirmation, but is there any statistical support or evidence on how to define grammaticality?

Reading Questions

- I am interested to know what variety of English is analyzed by our Ch 1-14 grammar? And if the answer is Standard American English, how we do make the determination on which variety is the "standard"? :)

- How did you and the other developers of HPSG decide on the dialect of English to map the grammar to?

Reading Questions

- In dealing with variations of English, or perhaps of other languages, do we create different 'sets' of grammar fragments to handle them independently?

- We now have a rule that is dialect-dependent, which both makes total sense and seems problematic. In a judgment of "is this sentence grammatical" do we then output a set of things it's grammatical in (AAVE, SAE, BrE) etc? Since we would need to block that construction being considered grammatical in some environments but not in others, and we aren't given that environment.

Reading Questions

- It is interesting to see how non-standard English can be analyzed with HPSG. Can we also analyze language used on the internet (like tweets) in a similar fashion?

- I understand AAVE is used in North America by African Americans and some black Canadians. But are there also variations in other English speaking countries with a history of African influence? For example, how similar would a South African or British African Vernacular English resemble that of AAVE?

Reading Questions

- I see how the textbook's analysis of the silent be differ's from Labov's deletion account in that it isn't based on contractions, but in general, how is making be silent any different from deleting it entirely?

- With all the discussion of be-deletion, I'm surprised to see no mention of "ain't" on p 458. Does the grammar not consider that a form of "be," and can it be deleted?

- What about other zero copula languages?

Reading Questions

- A few of the reading questions above asked how we could model synchronic variation across the dialects of a language. I'd like to know more about this too, and also if there is a way to model diachronic language variation. Can we account for language change over time? And maybe hope to answer one of Historical Linguistics' great questions: why do languages change?

Reading Questions

- Is there an addition of probabilistic modeling for HPSG to allow auxiliary verbs like have to vary its lexical entries based on its context ?

- Will there be rules or are there works that allow conversion from one dialect to another (e.g. AAVE <-> SAE) by treating it as a monolingual machine translation problem?

Reading Questions

- Is anything being done in industry to work with dialect variations, or is it just a "speak SAE or don't use our system" type of approach?

- I know treebanks seem to be used a lot, but it also seems like the web is scraped for data as well - that seems like there would be huge variation in data. Do the models just go with what seems most common and ignore that ones that don't work, or decide that multiple versions of something are acceptable?