Has Theoretical Linguistics Become a Cult Science?

Mo (18 Sept 2021)

1.YESTERDAY'S EMAIL

Athena, one of my mentees (real name withheld), forwarded to me yesterday an excerpt from what appears to be a textbook in theoretical syntax, requesting a zoom meeting to discuss the points made in the extract.

"Given that words (or syntactic atoms of some sort) combine with one another to form phrases, any theory of syntax must assume heads and phrases. But distinguishing between two types of phrases (intermediate projections vs. maximal projections) seems inelegant, and attempts have therefore been made to eliminate intermediate projections, along with the possibility of adjunction to them.

(13)

a. [$_{IP}$ They [$_{I'}$ never [$_{I'}$ will agree to that.]]]

b. God let [$_{VP}$ there [$_{V'}$ suddenly [$_{V'}$ be light.]]]

However, if the IP and the small clause VP in such sentences were 'split up' into two separate projections, it would be possible to eliminate the intermediate projections and to adjoin the modifiers to maximal projections instead. This is illustrated in (14), where IP has been split into Agr(eement)P and T(ense)P, and the small clause VP has been split into Pred(ication)P and a lower VP.

(14)

a. [$_{AgrP}$ They [$_{TP}$ never [$_{TP}$ will agree to that.]]]

b. God let [$_{PredP}$ there [$_{VP}$ suddenly [$_{VP}$ be light.]]]

A useful way to frame the issue is as a trade-off between two options. The first option buys a relatively small set of familiar syntactic categories at the cost of assuming intermediate projections. The second option buys an intuitively appealing two-level phrase structure scheme at the cost of a proliferating and increasingly abstract set of syntactic categories. As befits a science, current syntactic theory generally prefers the second option: generality at the cost of abstraction."

My comments on the excerpt are given below:

I am horrified at the pseudoscientific cult science norms implicit in this passage. Let me spell them out:

> **Norm 1**: "Given that words (or syntactic atoms of some sort) combine with one another to form phrases, any theory of syntax must assume heads and phrases."

Why should scientific inquiry follow this norm in the proposals for theoretical categories and subcategories? What is it that we are trying to explain-predict here, and how does the following assumption explain-predict what we seek to explain-predict?

> **Assumption 1**: The grammatical category of every whole is the same as the grammatical category of one of the daughter parts of the whole.

Without explicitly stating the theoretical assumption and providing ***evidence and arguments*** in support of it, (which requires considering alternative and supplementary assumptions) I would say that the above norm is just a cult science norm.

Also, when defending a theoretical assumption, the standard epistemic norm of scientific inquiry is

Norm 2: Theoretical assumptions are defended by their need to explain-predict observational generalisations.

What are the asymmetries of observational generalisations that the cult science represented by this passage tries to predict and explain?

The passage says:

> "But distinguishing between two types of phrases (intermediate projections vs. maximal projections) seems inelegant, and attempts have therefore been made to eliminate intermediate projections, along with the possibility of adjunction to them. "

When did "seems inelegant" become an argument against a theoretical assumption? Is this the kind of cult toxin that the 'experts' in theoretical linguistics are spreading among vulnerable grad students?

The passage says:

> "A useful way to frame the issue is as a trade-off between two options. The first option buys a relatively small set of familiar syntactic categories at the cost of assuming intermediate projections."

Trade off between what and what? What are the two advantages/disadvantages that we are trying to compare here? The relevant norms for scientific inquiry are:

> *Correct predictions of asymmetries in observational generalisations.*
>
> *Prohibition of logical contradictions*
>
> *Minimization of the theoretical premises (= definitions and axioms)*

Where is the proof that one of these options is better than the other in terms of these norms?

The passage says:

> "The second option buys an intuitively appealing two-level phrase structure scheme at the cost of a proliferating and increasingly abstract set of syntactic categories."

When did "**intuitive appeal** "become the basis for rational arguments in scientific inquiry?

When did "**increase of abstraction** become a sin in scientific inquiry?

What exactly is 'abstraction' such that one of these theoretical options commits the deadly sin of increasing abstraction?

The passage says:

> "As befits a science, current syntactic theory generally prefers the second option: generality at the cost of abstraction."

The cult science represented by the above sentence is no better than creationism dressed up as 'science'. If the passage that Athena has quoted is representative of what researchers and research students do in the name of theoretical linguistics, we have slipped right back into the dark ages.

If that is the case, it is tragic that the candle of rationality that Noam Chomsky lit in language studies more than sixty years ago has now become something that does not deserve the name 'scientific inquiry'.

If the readers are unfamiliar with the terminology of IP, I, TP, T, AgrP, AgrP, PredP, Pred, V', V etc. in (13) and (14) and the notation of square brackets that goes with it, here is a brief introduction that summarises the 'lecture' that I gave to Athena when we zoom met yesterday. I hope this summary will also be useful for those graduate students doing this kind of syntactic theory without an understanding the ***theoretical substance*** of the ***terminology*** and ***notation***.

2. YESTERDAY's ZOOM LECTURE (TO THE EXTENT I CAN RECALL)

There are two issues that we need to understand in this terminology and notation.

> What is it that the syntactic theory that this terminology and notation seeks to explain?
>
> What is the problem that formalism implicit in this terminology and notation seeks to solve?

Question 1: What any syntactic theory needs to explain

Consider the following pair of sentences:

A) 1. Lion is a noble animal.
   2. Lions are noble animals.
   3. *Lion are noble animal(s)
   4. *Lions is noble animal(s)

What I have given here is what would be called a single ***observational report*** or ***data point*** in scientific inquiry What it says is this.

> If, in an experimental set up, a member of the species of language that we call English is given sentences in (1)-(4) as the stimuli, and are asked to respond with 'acceptable' vs 'unacceptable', the experimental subject would respond with 'acceptable' for (1) and (2) but 'unacceptable' for (3) and (4).

The asterisk (*) in (3) and (4) denotes 'unacceptable to the experimental subject'. Its absence denotes 'acceptable for the experimental subject'. Now, you can replace the actual words 'lion', 'noble', 'animal' etc with other words in what is called 'English', and the contrast between (1) and (2) would remain.

What syntactic theorists seek to EXPLAIN are the ***systematic asymmetries*** in such observational reports on the ***experimentally elicited behaviour*** of the members of diverse language species:

> Why is it that members of the species of English accept (1) and (2) but reject (3) and (4)?

We can gather a ***large sample*** of such ***observational reports*** to check if the asymmetry illustrated by A is true of the sample. We can now generalise it to the population of the language species we call English.

The way the experiment is set up, in each data point we vary only what we think are the crucial variables. The crucial difference between (1) and (3) on the one hand, and (2) and (4) in the other is (i) the presence vs absence of *–s* in *lion**s***, and (ii) *is* vs *are*. That allows us to establish the following ***observational generalisation:***

> There is a ***correlation*** between the ***variable*** of the acceptability judgements in the response, and variables of (i) and (ii).

*What s **grammar** of the English species seeks to EXPLAIN-PREDICT are the systematic asymmetries of this kind. What a syntactic theory (universal grammar) seeks to explain are the homologies across the linguistic species within the human species.*

Question 2: What any syntactic formalism should allow us to do.

To explain asymmetries of this kind, we need to postulate ***premises*** (***axioms*** and ***definitions***), deduce the ***logical consequences*** of the premises (= ***theorems/predictions***), and show that the predictions agree with the observational generalisations.

Any ***theory*** that employs a logical or mathematical ***formalism*** also needs a vocabulary of algebraic expressions (like N, V, VP, P, PP, S) and a diagrammatic notation (tree diagram) to express the formal relations (e.g. arcs/edges to express relations like 'is a subcategory of'. 'is a member or', 'is composed of', 'is ordered prior to (along the temporal dimension', 'is the antecedent of' and so on. [For a discussion of trans-disciplinary relations of this kind, see 3.3 Properties and Relations in Section 3 "A theory of Circles" at http://www.thinq.education/articles/74]

The asymmetry between the acceptability of (A1) and (A2) and the unacceptability of (A3) and (A4) are found in (B1) and (B2) vs (B3) and (B4) as well.

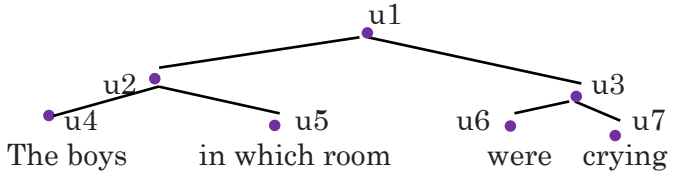
B) 1. The boy in which room was crying?
2. The boys in which room were crying?
3. *The boy in which room were crying?
4. *The boys in which room was crying?

It is intuitively 'obvious' that the asymmetry has to do with the absence and presence of *–s* in *boy* vs *boys,* and *was* vs *were*. How do we formulate a constraint/law that holds between the positions in *blue* in these expressions?

This is a classic case of what physicists call 'action at distance' (e.g., gravitational attraction in the action at distance between the sun and the earth, the entanglement between two fundamental particles that are beyond the speed of light…) To solve the problem of action at distance, linguists postulate hierarchical structure (units composed of sub-units, sub-units composed of sub-sub-units, ..) Here is a possible theoretical solution:

> *Premises (representations)*
>
> The unit "The boy in which room was crying" is composed of the units "The boy in which room" and "was crying"
>
> The unit " "was crying" is composed of "was" and "crying".
>
> The unit "The boy in which room was crying" is composed of "The boy" and " in which room was crying"
>
> and so on.

Premises of this kind are expressed in terms of the formalism of the Directed Acyclic Graph that linguists and biologists call tree diagrams.



To make it easier to refer to the units in these expressions, I have numbered the units (u1, u2, u3…). I have not represented the category status of these units (e.g., I have not indicated that boys is a noun, in is a preposition, that in the room is a prepositional phrase, and so on.)

Note 1: What I have called the ***relation*** of *compositionality* is what Chomsky calls the ***operation*** of *merge*. The distinction between "X is composed of A and B" and "Merge A and B to form X" is theoretically irrelevant (unless we set up an ordering of the operations as syntactic and phonological theories used to do till around 1980's.
Note 2: Compositionality/merge is a trans-disciplinary relation/operation relevant not only for linguistics, but also for mathematics, physics, chemistry, biology… (e,g, organs are composed of tissues, tissues are composed of cells, cells are composed of molecules, molecules are composed of atoms, atoms are composed of particles.)

Given representations of the kind illustrated above, we can correctly predict the observational asymmetries we need to explain by postulating the following constraint:

Premises (constraints/rules/laws/…)

The value of the variable for number (singular/plural) in unit 2 must be the same as the value of the variable for number in unit 3.

Linguists call this number agreement.

[Note: For our purposes, it does not matter what the category labels we assign to these units: NP or DP, VP or IP or AgrP. These are matters of differences in the terminologies of different religions/cults in linguistics that I have very little interest in.]

What is interesting, for our purposes, are the following points:

The value of the variable of number for unit 2 ***depends on*** a proper sub-unit of unit 2, namely, the unit *–s* immediately after *boy* (or its absence.)

The value of the variable of number for unit 3 ***depends on*** a proper sub-unit of unit 3, namely, *was/were* immediately.

How do we ***formally express*** the dependency relation in the atomic units and the complex units that the atomic units are part of?

The solution adopted in theories that use the terminology of IP, AgrP, VP, PP etc,. is the use of the concept of Head and Maximal Projection.

A Maximal Projecton is a complex unit whose property (the value of the variable) depends on the property of its Head ,

where 'Head' is the atomic sub-unit on whose property the complex unit depends on.

For the purposes of the constraint on number agreement stated above, the maximal projections are units 2 and unit 3, and their heads are –s or its absence in unit 2, and was/were in unit 3. In this terminology, given the units X and XP, XP is the maximal projection and X is the head.

Does this formalism of head and maximal projection to compute the dependency relation between complex units and their atomic subunits? I don't think so. Here is why:

The 'nouniness' of unit 2, the prepositionness of unit 5, and the verbiness of unit 3, depend on the nouniness of *boy,* the prepositionnes of *in*, and the

verbiness of *was/were*. So, given the formalism of heads and max projections, these are the heads unit 2, unit 5 and unit 3.

For the number of unit 2 depends on the number of –s/absence, and the number of unit 3 depends on was/were. So, given the formalism of heads and max projections, these are the heads of units 2 and 3.

The definiteness of unit 2 depends on *the*. So the is the head of unit 2.

The wh-ness of unit 2 depends on *which*. So which is the head of unit 2.

Given what I have stated above, should unit 2 be called NP, DP, NumP, or Wh-P? I leave that problem for grad students in syntax, with a sadistic glee.

[Note: The attribute-value notation and formalism used in Bay Area linguistics does not run into the kinds of difficulties that the MaxP-Head notation and formalism does.]

Mo