Review of Pavel Caha's habilitation thesis "Studies in Nanosyntax"

Pavel Caha's thesis for the habilitation consists of a set of articles that taken together highlight Caha's professional interests and skills both in theoretical linguistics in general and in the study of patterns characteristic of Czech morphosyntax in particular. On the general theoretical side, these articles illustrate the explanatory power of Nanosyntax, a theory of the morphosyntax of human language that differs from standard theories in ways that are carefully laid out in the thesis. It is fair to say that there is a rapidly growing interest in Nanosyntax in the international linguistics community, and Pavel Caha, who started working on Nanosyntax early on, is possibly the most frequently cited nanosyntactician next to Michal Starke.

To illustrate the explanatory power of a linguistic theory, one needs of course to be precise about what there is to be explained. The articles in this collection satisfy this criterion admirably. Caha (and his co-authors in chapters 3 and 4) lay out the data clearly and meticulously taking care to eliminate confounds and red herrings. The data are facts about different morphological patterns across paradigms. In the normal course of events, when the facts have been clarified, Caha proceeds to a demonstration that the patterns are just as expected on the basic assumptions of Nanosyntax. In fact, one of Caha's special skills is his way to lay out the description in such a way that the choice of analysis becomes obvious even before it has been presented.

Nanosyntax is obviously not supposed to be valid for only a small subset of morphological paradigms. Consequently, the different chapters of the thesis focus on patterns in different subdomains – nominal paradigms, the patterns comprising the graded forms of adjectives and different kinds of numerals.

Nanosyntax is a theory of human language, not just Czech. Therefore, even though there is a focus on Czech, Caha and his co-authors also look at contrastive data to increase the likelihood that the analyses they propose, are viable crosslinguistically. Obviously, trying to discover all the existing patterns in all the world's languages would not be feasible, and so this part of the endeavor involves relying on available contrastive work and tracing references to studies on particular languages considered potentially relevant. Caha manages this task very competently and even sometimes unearths crucial facts invalidating claims that would be inconsistent with his own analytical proposals. For example, he points out in chapter 6 that an analysis viewing Korean honorifics as showing non-local suppletion (as argued in the literature) actually turns out to be untenable once the interaction between honorifics and negation is taken into consideration.

Nanosyntax has its own set of rock-bottom assumptions not shared by other theories such as phrasal lexicalization, the Superset Principle and the lexicalization algorithm. The beauty of Caha's analyses comes to a large extent from the fact that they derive from basic nanosyntactic assumptions without the help of new theoretical assumptions. This highlights the explanatory power of the basic theory itself.

Similarly, Nanosyntax is also characterized by theoretical parsimony. For example, there is no such thing as context-sensitive lexical insertion. Allomorphy is instead controlled exclusively by the size of the syntactic trees associated with different lexical items. In Caha's thesis, it is shown that nothing more seems to be required pending the resolution on a couple of issues mentioned in the Conclusion.

All of this is not to say that there are not also innovative proposals. To illustrate, we may look at a proposal in chapter 3 that involves the Elsewhere Principle, one of the rock-bottom assumptions of Nanosyntax: When two morphemes both match a structure S by the Superset Principle, the one with the fewest features in it wins. This plays a role in the traditional nanosyntactic account of the *ABA generalization, among other things. In a series ABA ordered by the increasing size of the underlying structures, A must be able to lexicalize both

the biggest structure and the smallest structure, but the Elsewhere Principle forbids this, since B, which lexicalizes the medium sized structure, must have fewer features than A and therefore blocks A for the smallest structure.

But there is an issue: A root like *sheep*, which contains the two features PLURAL and SINGULAR (since it occurs both in the singular and the plural) should be blocked in the singular by roots like *goat* or *book*, which don't have PLURAL in them. Thus, it should be impossible to derive *a sheep*. To circumvent this, the authors propose abandoning the Elsewhere Principle. This is one two linked innovations. But now, what excludes ABA patterns?

An answer is provided by another innovation introduced in chapter 3: A morpheme A can override another morpheme B in the course of a derivation only if A = B or the entry for A contains a pointer to B, relying on the notion "pointer" (introduced ten years ago by Caha in cooperation with Marina Pantcheva and now included in the nanosyntactic toolbox). This condition, called the Faithfulness Condition, prevents *sheep* from blocking *goats*, while allowing *mice* to block *mouses* by virtue of having a pointer to *mouse*. It also prevents A from overriding B when the derivation of ABA reaches the biggest structure, unless A contains a pointer to B, i.e. But if A has a pointer to B, A cannot lexicalize [Z], the smallest structure, where B cannot already have been introduced.

Another piece of innovative thinking concerns the shapes of lexical entries. Standard Nanosyntax routinely postulates lexical entries that associate morphemes with structures formed merely by successive external merge of syntactic features. Caha and his co-authors, following up on a proposal by Hagen Blix, suggest that the syntactic tree in a lexical entry may also reflect syntactic movement, as in [XP [YP A][XP X]] resulting from raising the complement of X in [XP X [YP A]]. This provides a way of solving some challenging problems as shown in the appendix to chapter 3, and the proposal is currently also being explored by other nanosyntacticians. It should perhaps also be emphasized that the introduction of "Movement Containing Trees" doesn't actually add a new assumption to the theory to the extent that it has always been understood that any tree that can be built by syntax can also occur in a lexical entry. What is new, is the attempt to see how much mileage can be made from exploiting this fact.

Similarly (still in chapter 3), the somewhat arcane option of achieving full lexicalization via building a specifier in a separate work space is put to good use in the account of English analytic comparatives with *more*, which are compared to Czech comparatives with *-ej-* except that *more* is a prefix. The fact that some adjectives, e.g. *intelligent*, cannot form comparatives with *-er*, but need *more*, is attributed to the size of the adjectival root plus the lack of an English morpheme that would allow the feature lexicalized by *ej* in Czech to be lexicalized all by itself in English. (One remains curious, however, about the consequence that something in the syntax of *I'm more thirsty than hungry* apparently must prevent roots that normally are big enough to lexicalize the same feature as *ej*, cf. *I'm thirstier now*, from doing so.)

With respect to the empirical side of Caha's work, it should be emphasized that he shows an exceptional flair for less than obvious connections in the data including connections that other researchers have shrugged off as insignificant. A striking example is the analysis in chapter 5 intended to lead towards an understanding of the mysterious syncretism between the genitive singular and the nominative/accusative plural found in languages like Czech – a syncretism that previously has been regarded as a case of accidental homophony. The analysis in chapter 5 somewhat surprisingly starts out from observations about syncretism patterns in the series pseudo-partitive – count – plural and proposes a structural account adequate for all the different languages looked at in this chapter. The account is quite convincing, but leaves open an issue Caha himself comments on: How to characterize the singular count forms? In many languages, these have the same morphological form as the pseudo-partitive/mass forms, e.g. in English *a rope* and *two meters of rope*, although they are countable and therefore should

come with the feature COUNT, which is lexicalized by *s* in *three ropes*. Caha indicates two ways of handling this. On one account, the singular would be created by adding MIN directly to MASS without the intermediary of COUNT, and MIN would always be lexicalized by the root or by Ø in languages like English or Dutch. But having MIN lexicalized by the root would predict some instances of root alternation in such languages between the mass form and the singular count form of a noun, quite possibly incorrectly. Moreover, as Caha points out, this doesn't seem to support a semantic analysis.

The other account would have MIN on top of COUNT. Then, English-like languages must have MIN lexicalized by a Ø footed at COUNT which overrides *s*. The question then is whether there also are languages in which singular countable nouns have a suffix in common with nouns combining with numerals plus an extra singular suffix.

In this brief review, I have tried to give the reader a sense of the qualities that make this thesis an outstanding piece of work. I have not found anything to comment on that would subtract from this impression.

Pavel Caha's qualifications actually well exceed what is required for a habilitation.

prof. dr.philos. Knut Tarald Taraldsen

[redacted]

31.5.2022