

## HABILITATION THESIS REVIEWER'S REPORT

### Masaryk University

**Applicant**

Mgr. Pavel Caha

**Habilitation thesis**

Studies in Nanosyntax

**Reviewer**

Prof. Jonathan David Bobaljik

**Reviewer's home unit,  
institution**

Harvard University

I am pleased to offer the following report on Dr. Pavel Caha's habilitation thesis *Studies in Nanosyntax*. The thesis marks an outstanding contribution to the field, and is of the highest quality, meeting the highest standards in all respects. I elaborate on these points in the following report, in response to the items you enumerated in the request for a report dated 7 April:

### **How relevant is the topic within the field?**

The theoretical framework Nanosyntax has emerged over the last 15 years or so as a prominent approach to morphology in large part due to the work of Dr. Caha and colleagues. The work in this thesis is timely and exciting, and extremely relevant as one of the most carefully formulated current morphological frameworks. The work in the thesis offers a compelling new perspective on venerable questions within the field, notably regarding explanations for morphological patterns seen in Czech and cross-linguistically. The relevance of the papers taken individually is shown by their publication in top-tier journals, and the thesis lays out their relevance when taken together as a package. The proposals in this thesis are gaining international attention and are, for example, on assigned reading lists for recent seminars in the field at my own institution (Harvard) as well as at other leading departments including at least MIT, NYU, and the Université du Québec à Montréal.

### **How precisely, clearly and understandably is the topic formulated?**

Dr. Caha's work is an admirable model of clarity and serves as the main point of entry for most scholars into the Nanosyntax framework. Caha's writing starts with clear descriptions of the topic, and concise discussions of the motivations for the approach he is pursuing. The technical details can become complex, and Caha does an excellent job of providing scaffolding for the reader, starting with simple parts of the analysis, and then building to the more complex pieces. The mode of presentation is extremely accessible.

### **What contributions does the thesis make to the field and how does it compare to other works in the field?**

The primary contribution of the thesis lies in advancing our understanding of morphology, and thereby our understanding of how Language works in more general terms, specifically through the lens of a rigorous and constrained theory of possible languages (Nanosyntax). The thesis achieves this through case studies, which, in particular in the latter two chapters,

in addition adduce novel empirical generalizations which constitute contributions to knowledge in the descriptive sphere.

Chapters 2 and 3 specifically offer important contributions to a theory of possible morphemes, a fundamental area of inquiry in linguistics. From observation, it appears that language variation is, at some level of abstraction, quite limited: abstract patterns recur among unrelated languages, and the known languages represent only a limited subset of what should be logically possible.

The Nanosyntax approach to lexical entries aims to tackle one aspect of this venerable problem head-on. Chapter 2 on the Russian declension illustrates this particularly well. An established view in morphology is that inflection (declension, conjugation) classes are quintessentially morphological in nature, not reducible to phonology, semantics, or syntax, and that they are synchronically fundamentally arbitrary—the detritus of accumulated historical change. Caha's proposal on Russian declension argues persuasively that there are nevertheless tight patterns in the way declension classes are constituted, and that the tenets of the Nanosyntax framework provide an explanatory account of these patterns, by requiring specific, and limited, devices to encode classes of roots that show different combinatorial possibilities with inflectional affixes. This is an exciting new proposal whose full ramifications across a wider range of data are already inspiring further investigation.

Despite the focus on Russian, the patterns seen in Russian are common across Slavic (and beyond), and the demonstration that the variation can be reduced in a principled way to the spell-out mechanism that is independently motivated within the framework is an important proof-of-principle with implications then for other languages, such as Czech, being considered, as Caha notes, in ongoing work by Caha and his students.

Chapter 3 examines the same abstract topic from the same Nanosyntax perspective, but over a different data set, namely the composition of comparative morphology in Czech (and beyond). Czech is of particular interest here, since it seems, as Caha and colleagues argue, to have morphological patterns that clearly express the internal complexity of the comparative affixes, which are predicted to exist on theoretical grounds by the Nanosyntax framework, but less readily discernable in many other languages. Here too the contributions extend across many levels: as a direct contribution to Czech linguistics, the proposals explain why Czech morphology has the specific patterns it does, as opposed to others which would be readily conceivable but happen not to exist. Conversely, Czech morphology provides crucial evidence for theories of general linguistics regarding the internal complexity of morphological features, and, according to Caha's approach, about the nature of the grammatical (cognitive) representations of morphemes at an abstract level.

Chapters 4 and 5 make contributions regarding a different topic area, namely generalizations across patterns of lexicalization that have come to be known as \*ABA generalizations. Both chapters 4 and 5 contribute new types of evidence that propose to broaden the scope of patterns that fall under this general heading and draw conclusions about the structure of the grammatical representation of number categories. Chapter 4 makes the argument that expressions used for object counting ('five' in five books) are universally more complex than, and derived from, expressions that denote the number itself, as in mathematical expressions. This is widely suspected, and Wağiel and Caha provide new empirical evidence for this view from Czech and other languages, along with a detailed semantic and morphological analysis. Chapter 5 is more ambitious, and makes the bold, novel proposal that the use of genitive with lower numerals in languages like Czech and Russian should be seen as a part of a structural trio that includes the partitive genitive and the nominative plural, proposing an observed \*ABA generalization as intriguing evidence for this.

Caha's work in all respects compares favourably with other work in this area: it is carefully and explicitly argued, provides new empirical evidence and new perspectives on old puzzles, and argues at length for a fresh and exciting new perspective, while engaging with the existing literature. One way (among others) in which it holds more promise than much of the alternative literature lies in the nature of theoretical restrictiveness. For example, chapters 4 and 5 propose to derive \*ABA effects from the workings of a tightly constrained "spell out algorithm." Abstracting away from the details, one effect of this perspective is that strict locality of conditioning (contiguity) is more or less a theorem, rather than an axiom, derived, rather than stipulated. Much current literature is concerned with debates about the right formulation of locality (Embick, Moskal, Thornton, myself, and others have argued for a varying array of subtly different proposals). But from the perspective of Nanosyntax, the choice among these is essentially forced by deeper principles. As noted in the conclusions to the thesis, there is work to be done to address empirical challenges known from the literature, but the work presented here makes important headway in two quite different areas. It is an extremely exciting research program.

### **Methodology and formal standards**

The methodology is perfectly appropriate for the field and the thesis meets the highest formal standards. Theoretical analysis, including explicit generation of hypotheses from formal theory and testing of such predictions against new data are the gold standard for theoretical linguistics. The level of analysis is on the one hand the highest there is in the field: very explicit, intricately presented, and with internal consistency rigorously checked, and on the other, as noted above, Caha takes great care to make sure that even the most intricate and complex parts of the analysis are clearly presented with sufficient aids to the reader to be able to understand the work. All formal aspects (language, citations, etc.) meet the highest standards (again, as further evidenced by the fact that the articles have been published in prestigious international journals).

### **Summary**

This thesis makes an outstanding contribution to the area of Czech Linguistics narrowly and to General Linguistics more broadly. Specific to Czech Linguistics, it can quite centrally be read as asking why Czech has the particular morphological patterns that it has (the organization of the declension classes, the comparative morphology, the relations among the number words, etc). While an answer to the broad version of this question is in part contingent—languages are products of their history and context—Caha makes a very compelling case that significant and surprising aspects of the answer to this question have a universalist (and thus ultimately cognitive) basis—abstract patterns in Czech are the way they are because grammars are limited in the way that lexical entries (roots, morphemes) are represented, and the way in which the spell-out algorithm regulates choices among competing lexical items. This vastly limits the space of possible variation. Conversely, Czech provides a particularly compelling argument for this in the area of comparative grades (Chapter 3) and provides confirmatory evidence for this perspective in other domains where the patterns in Czech are similar to those in other languages.

For the reasons noted above, I am extremely pleased to offer my highest evaluation of Dr. Caha's habilitation thesis, *Studies in Nanosyntax*. Caha's work represents a significant accomplishment and sets a benchmark in the field—an important contribution that subsequent work will have to engage with. He offers well thought-out, provocative and stimulating theoretical proposals with empirical support, and clearly identifies both the achievements of the approach and where there is work to be done. The work sets out an exciting research program which I look forward to following with much interest over the coming years. This is outstanding work, which would unquestionably be of the level

corresponding to tenure and promotion in a North American university (the context with which I am most familiar) and fully deserving of the award of a Habilitation.

**Reviewer's questions for the habilitation thesis defence** (number of questions up to the reviewer)

1. The proposal on Russian declension classes seems to make an intriguing general prediction, namely, that if there are (sub)classes whose differences are neutralized in some cases, then the cases that are distinct should be lower on the Case Hierarchy (contiguous with the root) than those that are neutralized: treating masculine and neuter as distinct only in nominative/accusative seems to illustrate this well. Likewise, I take it that this requires that the genitive /i/ in Classes II and III must be treated as homophonous since syncretism exclusively in the middle of the case hierarchy is not possible. This strikes me as a fascinating prediction, unique (if I have understood correctly) to this approach. Is it borne out? What about the small subclasses in Russian declension, such as the 100 or so masculine nouns that have a distinct partitive genitive or a locative (in -u), but otherwise take the regular masculine endings? This appears to be root-specific declension information, thus must be encoded as part of the size of the root, but differences that are a factor of root size need to play out over root-contiguous structures, do they not?
2. Kiparsky's classic 1968 paper *How abstract is phonology?* was a critical response to attempts to analyse root-controlled class-like behaviour without diacritics via appeal to abstract phonological representations – part of his allegation was that the program merely masked what was fundamentally an appeal to diacritics by coding these with as hoc uses of phonological notation. How would you respond to the claim that Nanosyntax is doing on the syntactic side what Kiparsky objected to on the phonological side, i.e., using syntactic notation for diacritics in ways that have no independent syntactic justification?
3. How does the Russian partitive genitive (čáj-u) fit into the model described in Chapter 5? About 100 mass nouns in Russian have a partitive genitive distinct from the regular genitive. As far as I know, the partitive genitive (which is what the semantics in this chapter appears to apply to) is never part of the syncretism with counted N and nom.pl. (this is hard to see directly, since it only occurs with mass nouns, but when one looks at noun classes as a whole, it is the -a, not the -u, that enters into the syncretic patterns described here). Does this present a challenge?
4. Chapter 2 provides an argument against having a null morpheme to mark the masculine singular nominative (102) because that null morpheme and the neuter -o would have the same sub-tree/features. This leads to the interesting treatment of the masculine root *zavod* as bigger than the neuter root *mest*, and requires or motivates backtracking. How do yers fit into this proposal? A common proposal is that the nominative singular masculine ending is yer, not zero, in order to account for vowel-zero alternations (lev~lva) etc. But if yer is a suffix, then does the problem in (102) arise again? Or is the yer part of the root, deleted before other (case) affixes?

## Conclusion

The habilitation thesis entitled *Studies in Nanosyntax* by Mgr. Pavel Caha **fulfils the** requirements expected of a habilitation thesis in the field of Czech Language.

Date: 2 June 2022 Signature:

