Reply to comments on
Universal Morphology

Michal Starke
Masaryk University
michal.starke@mail.muni.cz

One perk of academic life is the gratification of high caliber intellectual exchanges which emerge haphazardly in personal discussions between colleagues, within and across disciplines. The hope of experiencing more such exchanges lead me early on to be an avid reader of journals such as Behavioral and Brain Sciences and others which published comments on their articles and replies to those comments. Big was my surprise as a young researcher to read senior professors calling into questions each other's basic comprehension skills and using rhetorics of the type that has now come to be associated to comment sections of social media.[1] Perhaps in an effort to preserve intellectual joy, I have come to believe in a 'let a thousand flowers bloom' approach to science. Let researchers develop the best science they can, be delighted by each other's creativity, and let posterity decide which approach led closer to truth. The field of human capacity for language is vast, it has more than enough space for the various approaches currently under consideration.

This stance has kept me away from public cross-theory comparisons and comments, while enjoying the ideas and theories carefully expounded in talks and articles that I end up disagreeing with. I'll therefore try to keep to clarifications and pointing out trade-offs – or at least that is the intended reading of the following remarks.

---

[1]   See Krausman, Morgart & Rosenstock (2003) for an academic elaboration of this topic.

1. The creation of morphemes

The essence of nanosyntax, and the first few minutes of the talk under discussion, is that morphemes are created by syntax: recursive grouping (merging) of single features creates articulated constituents corresponding to a single morpheme. To put it perhaps too plainly: there is a regular syntactic structure inside the belly of each morpheme. A happy discovery is that this allows us to entirely bypass some traditional problems of morpho-syntax — portmanteaus just happen to have a bigger constituent in their belly, allomorphy and suppletion happens when the same or similar enough concepts happen to have slightly different sizes of syntactic constituents in their belly. A corollary is that arranging morphemes into words, and words into sentences, is the wrong starting place to understand language and morphology — arranging features into morphemes (and words, sentences and paragraphs) is where the light shines.

There is a by now relatively traditional idea that syntax is responsible for the arrangement of morphemes into words. Within modern generative syntax, influential proposals along these lines include Baker's Mirror Principle (Baker 1985) and Pollock's proposal about verb movement as interpreted by Belletti (Pollock 1989, Belletti 1990), but also work by many other scholars of that time, such as Brody, Cinque, Rizzi, Roberts and others. I will refer to this syntactic school of thought as Arranging Morphemes in Syntax, or *les amis* for short.

Perhaps the deepest and most frequent misunderstanding around nanosyntax is that it does just the same as *les amis*: it just arranges morphemes into words in syntax, repeating what was done 30 years ago. This is on display in a couple of comments, for instance Manzini states that:

> Nanosyntax is meant to [...] entirely reabsorb morphology under syntax. But it is neither alone nor first in advancing this idea. Two decades ago, Manzini and Savoia (2002) suggested that phenomena generally deemed to be morphological, such as those involving the internal organisation of the clitic string, are in reality purely syntactic.

Interesting as it is (which is very), Manzini and Savoia's mentioned work belongs to what I here call *les amis*, and limits itself to arranging morphemes in syntax (here clitics) and does not take the step that makes nanosyntax nano: creating individual morphemes in syntax by mergers of single features, and thereby resolving issues such as portmanteau, suppletion, the *ABA or allomorphy. In fact nanosyntax goes further and entails that *les amis* are wrong: for *les amis*, morphemes are syntactic terminals, i.e. syntactic primitives, while nanosyntax claims that having morphemes (or words) as syntactic primitives is what has led linguistics astray.

The same is true of Thornton's claim that nanosyntax is "strikingly parallel to just items (lexicon) and arrangement (syntax)" and this "since [in both,] items (lexical items / morphemes) are put together (i.e., "arranged", in the American structuralist terminology, or "merged", in the recent syntactically oriented parlance) in syntax". Again: nanosyntax claims that this view of Items & Arrangement gives the wrong starting point, obscuring the underlying regularity of morphology and of language in general. Having morphemes arranged in syntax is a secondary side-

effect of opening a morpheme's belly and discovering regular syntax inside — a claim shared by none of the above (yet, one hopes!).

2. Where is morphology, again

The tenet that a morpheme is created by recursive groupings of single features into syntactic phrases has another corollary: the operations creating morphemes (and as a side-effect arranging morphemes into words) are regular syntactic operations, nothing else. Take for instance Thornton's "it is not clear to me [...] what the evidence for the existence of last-resort movement is". Last resort movement is a tool introduced by Chomsky (1995) into syntactic theory, used since in deriving a multitude of various syntactic phenomena. The fact that this same tool is used by nanosyntax in the derivation of apparently irregular morphology is a good illustration of how regular syntactic operations are used in the derivation of surface morpho-syntactic irregularity.

This is the polar opposite of approaches such as autonomous morphology, morphomics, post-syntactic morphological components (or parallel morphological components for that matter). Many of us agree that there is hidden regularity underlying morphology, we however locate that regularity in entirely different components of the grammar. Nano is firm in locating it entirely inside the syntax.[2]

In doing so, nanosyntax also constitutes a research program about the nature of syntax itself. Take for instance the fact that syntactic terminals are single privative features. This eliminates assumptions about valued/unvalued features, core to many versions of minimalism. Implicit assumptions such as Di Sciullo's are thus a misunderstanding: "Why would morpho-syntactic derivations differ from syntactic derivations [in Chomsky 1965, Yang 2016, Chomsky 2017], in a framework where morpho-syntax is syntax?". This implicitly assumes that all syntax is of the "Chomsky 1965, Yang 2016, Chomsky 2017" type -- which is clearly not the case.[3] In fact nanosyntax aims to simplify the rich technological

---

[2]   There seems to have been some misunderstanding based on the remark in the talk that [some facts] "have led people to give up on (irregular) morphology". This is an invited talk for a syntax slot given at one of the leading conferences in generative grammar, by a generative syntactician. The intended reading of that remark is thus that syntacticians who have tried their hand at morphology have indeed typically given up on phenomena such as English irregular verbs, French suppletive roots, etc. — and not the unlikely (and unintended) reading that morphologists have given up on (regularity in) morphology.

[3]   Di Sciullo also makes the claim that merging features in binary mergers (as in nanosyntax) is more complex than merging them in a flat structure (as in blundles): "This hypothesis further extended functional projections, see (3), and add complexity, defined in terms of length of the derivation, to the derivations". Putting aside the factual accuracy of that claim, two considerations need to be kept in mind in discussions of computational complexity. First, cherry-picking where one looks for complexity (e.g. "length of derivation") leads one to overlook complexity induced elsewhere (e.g. introduction of a different module and the concomitant cross-module coordination costs, complexity costs of accessing features hidden inside bundles or handling bundles to begin with, etc) - resulting in inaccurate complexity assessments. More importantly perhaps, it is well understood in computational discussions of complexity that results are often counter-intuitive and hence informal reasonings are of limited use. Only a thorough formalisation,

apparatus of Minimalism: no valued/unvalued features, no EPP features and accompanying postulates about movement triggering, no phases, and hence no stipulations about edge of phases, no Agree operation, etc. In place of that, nanosyntax seeks to squeeze juice out of two assumptions needed by everyone:

First, any approach needs some technology to express that the order of mergers is highly constrained. Complementisers are merged above/after the verb of the corresponding clause, never below, C > V. Some tense distinctions are always merged before complementisers, never after, C > T > V. Aspectual features are merged before tense features, C > T > Asp > V. The same is true of agentive, mood, evidential and many other features. It is traditional to ask 'why that particular order' – and I have seen no plausible answer so far – but what is usually overlooked is that the existence of this ordering allows us to drive the building of derivations and the triggering of movements, making it unnecessary to postulate technology such as the EPP, uninterpretable features, and their concomitant assumptions about movement triggers – greatly simplifying minimalism and adjacent theories.

Second, any approach needs some way of mapping syntactic representations onto gestures (ultimately yielding sounds or other physical artefacts) and concepts. Again, nanosyntax pursues the conjecture that a lot more juice can be squeezed out of that shared assumption than is commonly understood – through what we call the lexicalisation algorithm – hopefully simplifying the theoretical apparatus of other approaches.

Nanosyntax thus puts all of morpho-syntax firmly inside syntax, but also aims for that syntax to be simpler and more elegant than traditionally assumed.

As is usual in research, different starting points lead to focus on different parts of the data. Thornton for instance notes that viewing morphology as a separate domain is conducive to emphasising cross-paradigmatic patterns, as is beautifully done in Aronoff's work, or in work by Martin Maiden, Edwin Williams and others. Can that data be handled while unifying all of morpho-syntax inside a single principled generative component? Some of it is, and will be e.g. attributed to a functional element (an affix) recurring across paradigms and triggering the same effect in each one of them. Others aren't, and will typically be attributed to diachrony (perhaps the "diachronic moulding force" Maiden discusses) rather than the synchronic state we're investigating.

The synchronic aspects of this data is indeed an area of high interest to nanosyntacticians. For instance, some of us are involved in work on Latin paradigms, with an eye towards Aronoff's generalisations. Similarly for the other facts and work cited by Thornton — with more hours in a day, or more people participating, we would love to work on each and every single one of those patterns — we think they will illuminate the theory of our syntactic competence, and I have tried to illustrate in the talk under discussion why I think this is a promising bet despite a long tradition of skepticism. Our priorities however lie in other areas, that have been comparatively neglected by other morpho-syntacticians.

Which approach is correct, only time will tell.

3. The "Universal" in Universal Morphology

or thorough measurements within limited bounds, can move the discussion towards reliable conclusions.

There is a lot to be appreciated in Anderson's frankness that his world-view and experience doesn't let him "get inside Starke's Weltanschauung" (I am less sure what to think of his confession that he hasn't watched the entire talk he is commenting on). There is one misunderstanding though that may be useful to the reader. Anderson informs us that nanosyntax "reminds me of" using "rules [which] happen to be formulated in such a way that they couldn't ever apply to anything except forms [under discussion]". Nanosyntax, and the talk, repeatedly insist on the fact that the same small universal algorithm is used in all of nanosyntactic work, whether it is on Ossetic case (Caha 2020), Latin comparative suffixes (De Clercq & Vanden Wyngaerd 2017), Bantu class markers (Taraldsen, Taraldsen Medová & Langa 2018), Slavic complementisers (Wiland 2019) or many others. Even if one doesn't reach that point of the talk: it is right there in the title.[4]

      Putting these points together, the research programme under discussion is that syntax operates on individual features, recursively creating constituents large enough to be lexicalised by a morpheme, and does so using a small set of universal/invariant syntactic operations. The goal of the talk was to illustrate that this is not only feasible, but illuminates areas of morphology that have been traditionally difficult to handle in syntax, or in any regular way, such as root suppletion or portmanteaus.

4. Do you really?

A result that is dear to our hearts is that apparent irregularities are handled without adding any mechanism dedicated to them, or justified on their basis. Rather, we only use a general and universal syntax, a lexicon containing only well-formed syntax, and a small set of invariant principles relating the two (essentially the Elsewhere Principle and cyclic lexicalisation). Surprisingly, De Belder challenges this: "Do not be mistaken, the irregularity is there in the Nanosyntactic approach as well".

      This assertion rests on the idea that irregularity in nanosyntax "is masked as a lexical item". What hides behind this, is I think a common misunderstanding; let's get into it in some detail.[5]

---

[4]   Similarly for Anderson's other reminiscence that "when there are two rules, and the results are different depending on the order in which they are applied, apply the rules in such a way as to produce the correct form". The charge here is of circularity — presupposing the result and choosing the rule as a function of the presupposed result. It is easy to imagine that this may be the way it looks like when one didn't "get inside" the talk enough to follow the mechanics of the proposal. But even without understanding, watching the end will reveal a computer implementation of the algorithm, a fact clearly incompatible with Anderson's aspersion. Indeed, nothing could be further from the truth — we strive to not only have a principled but also formalised and testable approach to morpho-syntax.

[5]   De Belder adds a number of other surprising assertions, from the idea that a talk about morpho-syntax (slide 3, minute 3:00) should rather handle morpho-phonology, to mischaracterisations of this research as handling the regular, inflectional part of morphology. Handling regular morphology is for instance what *les amis* do, for the most

First, there is a reading of the above quote, under which it is correct, but not relevant: nanosyntax can handle surface irregularities, it hence has a way to represent those irregularities (as lexical items). This is for instance the reading of "irregularity" that De Clercq uses in her comment: "Irregularity (in roots or affixes) is cashed out in the system as the presence of lexically related lexical items whose structural size differs". In that sense the irregularities 'are there' — obviously if one is explaining them — but without affecting the relevant point: there are no mechanisms dedicated to them, or justified on their basis.

Seeing what it looks like to have some mechanism dedicated to irregularities (or justified by them) would perhaps help. Let's take the example of Distributed Morphology (DM). In DM, morphemes (qua syntactic bundles) "are put together ... in syntax" to reuse Thornton's description. Typically it adopts a so-called "minimalist" flavour of contemporary syntax, and hence uses tools such as Agree, phases, Merge, etc. The relevant question is then: does DM handle portmanteaus, suppletions, etc. using nothing but those syntactic tools? Or does it add something to its grammatical toolset in order to deal with those? As is well known and well commented upon, it adds a large toolset[6] — what De Belder calls "the elephant in the room", and which she describes as:

> Distributed Morphology has this post-syntactic module called Morphology which hosts a number of operations, including Fission (Noyer 1997), Impoverishment (Bonet 1991), and several Merge operations, including Morphological Merger (Marantz 1984), Syntactic Lowering (Bobaljik 1994) and Local Dislocation (Embick and Noyer 2001).

In addition to those tools, DM brings back language-specific, context-sensitive rules (in its Vocabulary Insertion mechanism), a device widely understood to be so powerful as to prohibit an explanatory theory.[7] The justification for this reintroduction is to handle irregularities.

The substantive fact then is that nanosyntax has none of that: it adds no tool dedicated to or justified by irregular morphology (and no independent morphological system), using only regular syntax and a restrictive format of lexical entries, while still deriving the apparent irregularities of morpho-syntax. I take that to be a significant breakthrough, and we will keep it that way — or as close as possible to that way.

5. Where do lexemes come from?

If lexemes contain syntactic trees, where do these trees come from? A frequent question, brought up by Borer in her comment: "the nature of lexical entries and how they are constructed is critical to the assessment of the system". We think of them as a 'garage of interesting syntactic structures', and language acquisition (of

---

part, but bears little resemblance to this talk, to nanosyntax, or to research on root suppletion and similar topics.

[6]   See Caha (2019) for article-length comparisons between DM and nanosyntax.

[7]   Whence the efforts to at least partially reign it in, e.g., Bobaljik (2000), Embick (2010), Choi & Harley (2019).

syntax) is the act of populating that garage. When the learner encounters an "interesting" structure built by syntax, she may decide to keep a (frozen, static) copy of that syntactic structure in her garage. Syntax is the dynamic component building structures, the lexicon is a static repository of such structures.

Which structures count as interesting? Here we are back to familiar territory: 'interesting' is mostly determined by predictability. If the meaning or pronunciation/gesturing of a structure is not predictable, it needs to be stored in the garage, associated with its meaning or pronunciation/gesturing. Other factors may make a structure interesting: the learner may decide to store compositional but very frequent structures — an issue more often discussed in psycholinguistics or neurolinguistics.

6. Where is the discussion of the data?

Almost none of the commentators engaged with the empirical topic of the talk: irregular verbs (in French), their suppletion patterns, their portmanteau patterns, their syncretisms. One exception is in Thornton's comment, which asks about the derivation of present tense structures:

> At minute 17:34 it is said that no lexical item has the sequence [#[T …]] in it ("none of the lexical items down here has number and T as their lower layers"); therefore, a derivation containing this sequence crashes, and must be rescued by "Last-resort movement", that moves the T head to the left of the # head. However, the lexical item ε introduced at minute 19:44 does contain exactly the sequence [#[T …]] in the lowest layers of its tree. It is unclear to me why the derivation would crash at minute 17:34 for lack of a lexical item matching it, if this lexical item does indeed exist, since it comes up a few minutes later. I can understand that this item ε was not introduced earlier in the talk for expository reasons, but it should be present in the mental lexicon of a mature French speaker at the same time in which õ, e, ə and i are present — so the derivation shown at minute 17:34 should not crash, in my view. I am very interested in knowing what I have missed: that might explain why the crash would happen.

The answer lies in the full representation of ε, which is given at minute 24:

ε: [Spk [Part [Pers [# [Past [T]]]]]]

The relevant part is in the bottom layers, i.e. [# [Past [T]]]. In the derivation Thornton asks about, syntax builds the structure [# [T]], representing a present tense with the beginning of a number representation. Since we keep mechanisms as simple as possible, a lexical entry can lexicalise a syntactic structure only when it contains the exact same constituent as that syntactic structure — and clearly [# [Past [T]]] is not the same as [# [T]] where 'Past' is missing, and hence ε cannot lexicalise [# [T]].

Beyond French irregular verbs, Di Sciullo asks how "privative aspect sub-features [would] account for simple contrasts such as *courrir* ('to run') / *acourrir*", or how such an approach would account for the different expression of directional meanings across languages:

For example, contrary to English, in Italian, the preposition a is [locative] with certain stative verbs, e.g., *stare a scuola*, 'stay at school', and [directional] with certain activity verbs, e.g., *andare a scuola*, 'go to/*at school'

One simple type of answer is that the directional feature situated between the verb and the noun is lexicalised differently in different languages: together with the verb in some languages, by a preposition in other languages, yielding the core of Talmy's distinction between satellite-framed and verb-framed languages (see Fábregas 2007, Son & Svenonius 2008, Son 2009 for discussion). Similar considerations apply in the aspectual domain brought up by Di Sciullo.

Borer brings up the interesting case of templatic verbal systems. Part of her discussion revolves around the idea that each cycle of lexical access must "constitute a well-formed phonological word". This may be so in other approaches, but it does not follow from the nanosyntax setup so far, and hence "the existence of intermediate (syntactic) derivational stages with a partial structure which does not, and indeed cannot constitute a well-formed phonological word" is prima facie not an issue. Similarly, Borer takes the "discontinuous representations" involved in templatic phenomena to be a challenge to "bottom-up lexical representations". Given the well known tier-based proposal that such "discontinuous representations" are underlyingly concatenative, it is not clear that this fact is a particular challenge.[8]

There are in fact at least three persons working on deriving the details of templatic verbal paradigms from nothing-but-fine-grained-(nano)syntax — Nico Baier on Kabyle, Gioia Cacchioli on Tigrinya, and Abdullah Nabil Almahaish on Hasawi Arabic. Similarly for the Italian gemination facts mentioned by Thornton, a phonologist and two syntacticians are currently looking into related Italian issues of root suppletion: Edo Cavirani, Maria Cortiula and myself. So as in the talk, I'll end this with an invitation to join: if the prospect of a fully principled explanation integrated with syntax appeals to you, feel free to contact me and join our online working groups. Our activities involve not only theory, but also building a comprehensive database of root suppletions and minor ("irregular") paradigms across a range of languages, as well as computer modelisation. You'll find my email at the beginning of this note.

Acknowledgments

References

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373–415.

---

8   See e.g. Zukoff (2021) for one recent example involving that line of thought.

Belletti, Adriana. 1990. *Generalized Verb Movement: Aspects of Verb Syntax*. Turin: Rosenberg & Sellier.

Bobaljik, Jonathan. 2000. The ins and outs of contextual allomorphy. In K. Grohmann, K. & C. Struijke (eds) *University of Maryland Working Papers in Linguistics* 10: 35.

Caha, Pavel. 2019. Nanosyntax: some key features. *The Cambridge Handbook of Distributed Morphology*. Cambridge University Press. Lingbuzz/004437

Caha, Pavel. 2020. On case competition in Ossetic numeral phrases. To appear. Open Generative Syntax, Language Science Press. Lingbuzz/004875

Choi, Jaehoon & Heidi Harley. 2019. Locality domains and morphological rules. *Natural Language and Linguistic Theory* 37:1319- 1365. https://doi.org/10.1007/s11049-018-09438-3

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

De Clercq, Karen & Guido Vanden Wyngaerd. 2017. *ABA revisited: Evidence from Czech and Latin degree morphology. *Glossa*, 2(1):69. Lingbuzz/003781

Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology*. Linguistic Inquiry Monograph. MIT Press, Cambridge. Lingbuzz/000791

Fábregas, Antonio. 2007. The Exhaustive Lexicalisation Principle. *Nordlyd* 34(2). Lingbuzz/000930.

Krausman, John R. Morgart & Steven S. Rosenstock. 2003. *Wildlife Society Bulletin* 31(1): 208-211.

Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar and the Structure of IP. *Linguistic Inquiry* 20: 365–424.

Son, Minjeong & Peter Svenonius. 2008. Microparameters of cross-linguistic variation: Directed motion and resultatives. *Proceedings of WCCFL.*

Son, Minjeong. 2009. Linguistic Variation and Lexical Parameter: The Case of Directed Motion. *UPenn Working Papers in Linguistics*, 15(1). https://repository.upenn.edu/pwpl/vol15/iss1/24

Starke, Michal. 2014. Cleaning up the lexicon. *Linguistic Analysis* 39(1). Lingbuzz/004309

Taraldsen, Tarald, Taraldsen Medová, Lucie & David Langa. 2018. Class prefixes as specifiers in Southern Bantu. *Natural Language and Linguistic Theory* 36:1339-1394.
https://doi.org/10.1007/s11049-017-9394-8

Wiland, Bartosz. 2019. The spell-out algorithm and lexicalization patterns: Slavic verbs and complementizers. Language Science Press.

Zukoff, Sam. 2021.The Mirror Alignment Principle.
Lingbuzz/005374.