Patterns and Linguistic Reality: a reply to Stanton

Ryan M. Nefdt [1,2]

[1]University of Cape Town
[2]University of Bristol

February 4, 2026

Abstract

In her article, Stanton (2025) argues that the real pattern analysis of language proffered by Nefdt (2023) fails as a demarcation of language versus non-language. She also makes the case that structuralism about linguistics does not characterise certain important metasemantic aspects of the broader field. In this article, I offer a rejoinder to these points by emphasising the role the linguistic community plays in determining real patterns as well as the theoretical benefits the view that grammars are multilayered compressions algorithms has for semantics and other linguistic disciplines.

Keywords: real patterns; philosophy of linguistics; ontology; philosophy of science; grammar; information theory

1 Introduction

In this article, I provide a measured response to the critique made in Stanton (2025) concerning real pattern analysis as a conduit to understanding both the ontology of language and the nature of its scientific study.

As Stanton identifies, there are two core (and interconnected) questions in Nefdt (2023) and the philosophy of linguistics more generally: (1) what is a language? and (2) what is linguistics about?

In the book, I offer two general lines as candidate answers for each question.

1. Languages are real patterns found in complex biological systems known as linguistic communities.

2. Linguistics is a science of the structure of these patterns.

In what follows, I will briefly respond to the careful and insightful critique that Stanton mounts against this view.

2 *Linguistically* real patterns

The seminal work that motivated the relevance of real patterns in philosophy (and especially the philosophy of science) is Dennett (1991), in which a 'semi-realist' view of mind is proposed. The central idea dates back further within information theory and rides on well-established results in coding theory (Kolmogorov 1963). Two simple insights guide the definition of a real pattern: (1) a random string or bit map is incompressible and (2) patterns emerge as regularities in data. Thus, for Dennett, to be a real pattern is to be a compression or more efficient representation of the data.

What makes the idea particularly rich is that it can apply directly to the concepts we inherit from folk psychology (such as propositional attitudes), as Dennett initially intended, as well as potentially explain the relationship between scientific models and their targets in the natural world (Millhouse 2021), i.e. models represent patterns in the data that indirectly map onto features of the world.

Nefdt (2023) takes both of these uses into its account of language and linguistics. I posit that grammars, akin to scientific models (Nefdt 2016), are compression algorithms that characterise formal patterns in linguistic data and associate them with compressed representations (rules and/or graphs). I also believe that this matching provides a clue as to the ontological status of languages, which are basically reified real patterns extracted from the mental and social practices of language users in a community. So to learn a language is to sort the noise from the regularities or patterns.[1] Grammars are compressed representations of this process at an abstract theoretical level.

From this general account, the book develops ontological criteria along the lines of Ladyman and Ross (2007)'s modification of Dennett's picture. Specifi-

[1]This could involve innate predispositions or strong inductive biases or massive amounts of data and statistical processing as in the case of large language models. See Futrell et al. (2020) for a real pattern analysis of the latter, partly inspired by Nefdt (2023).

cally, to be a 'linguistically real pattern' (LRP) a structure has to be (i) projectible or predictable from the language data, (ii) be modelled by a strong generative grammar, and (iii) have a unique structural mapping between strings and structures (like graphs or trees). Stanton correctly notes that I am favouring a syntactic (and in part generative) understanding of grammar and structure here. But as I point out in the book, this is only the case because I can mount a neat formal correspondence between the discrete rule systems of generative grammar and the information-theoretic properties of real patterns advocated by Dennett, Ladyman & Ross, and Millhouse among others. In fact, Nefdt (2023) shows that non-redundancy or indispensability, compressibility and importantly a similarity metric (e.g. the 'Levenshtein distance') for scientific models can be translated into generative grammars more perspicaciously. This is crucial for the picture and how it can be the launching pad for successful transfer of the theory of real patterns in the philosophy of science to the present case of formal linguistics. Importantly, it isn't really generative grammar as much as the overarching mathematical field of formal language theory (FLT) that I motivate for my theory, of which generative grammar is characterisable as a proper subset.[2]

Stanton further worries that to 'linguistically be' collapses into 'to be a language', rendering a confusion between *sub*linguistic structure and full language. I agree that this would be an issue but the correct shorthand should be that to 'linguistically be' is to be 'language'. This tracks the truism that any (proper) part of a language, i.e. linguistic feature, is *language* (i.e. the mass noun). Nevertheless, the more serious concern she has is with the possibility that the view is much too permissive. For this she mounts an creative overgeneration problem. We'll move on to this issue and its possible tonic below.

3 Train Stations and Biological Systems

To show that the formalist, structuralist approach to defining linguistically real patterns overgenerates, Stanton provides an ingenious case. She sets out to model an artificial, non-linguistic system in terms of criteria (i)-(iii). To this end she invokes a Liverpudlian toy example involving train routes between stations. She shows that for the train station map to satisfy the criteria, including mapping onto a graph or tree, it is relatively simple. For the first one she assigns probabilities to station routes (here acting like state transitions in a finite

[2]For example, many of the same formal structures and languages can be characterised by proof-theoretic and model-theoretic means (Pullum 2013; Miller 1999).

state automaton) with stations as states. The probabilities are determined by likelihood of reaching one station from another within a certain period. She then defines a nondeterministic random enumerator as a function over states and transitions between them. I think there is slight misreading of criterion (i) here which states that a linguistic pattern is real iff "it is grammatically projectible or there is a nondeterministic random enumerator which generates y from x" (Nefdt 2023, p. 78). The criterion is tolerant of alternatives to random set enumeration and specifically contains the term "grammatically projectible", which is meant to capture that we are dealing with linguistic systems. Nevertheless, let us elide that point for now. Criteria (ii) and (iii) are similarly easy to represent by further specifications on functions and a corresponding automaton.

From this case, Stanton concludes that the structuralism of the project compels a position in which we have to accept Liverpool Central station maps as languages or linguistically real patterns. Needless to say, this would be an untoward result for the theory, even if linguists might be brimming at the possibility of hitherto undiscovered languages so close to home. Fortunately, I don't think this follows. The fact that a formalism or mapping can be produced that resembles natural language structure is not a case against the real pattern view. In fact, it is a prerequisite for FLT. Furthermore, the fact that our mathematical models are multiply realisable is a general feature of scientific modelling (see Thomson-Jones (2012) for curious such case involving a pendulum and a circuit).

What Stanton really seems to be getting at is that any physical structure should be equally admissible given the structuralism of the view. She argues that "A 'language' by his admission is just a real pattern that can be described by (i-iii)" and furthermore "the sense in which the route map for Liverpool Central is physically realised in the stations and their connections is considerably clearer than the sense in which a Generative grammar is realised in the brain" (Stanton 2025, p. 6).

This produces a very narrow depiction of the view in the book. For one thing, it ignores the main arguments of Chapter 4 in which I argue that mathematical structuralism of the Shapiro and Resnik variety is insufficient to characterise linguistics precisely because the content of the structures do not matter there. There, I insist that:

> What we have above is the characterisation of pure structures in mathematics. But such a view won't be sufficient for modelling or

> representing linguistic patterns [...] In linguistics we seem to be concerned with a specific class of structures, those [...] that are produced by human linguistic competence". (Nefdt 2023, p. 63)

The entirety of Chapter 5 is devoted to the question of what or where the real patterns of language are to be found. The book offers a naturalistic picture involving linguistic communities as the nexus of biological systems from which patterns emerge. By doing so, it attempts to marry certain internalist and externalist views of language. Stanton rejects this move. She insists that ontic structual realism is not compatible with a domain restriction of this sort. She accepts that this strategy might work "if the language sciences and linguistic communities are also given a purely structural characterisation, but if they are not then the book's central claim that language is structure is not strictly true" (Stanton 2025, p. 6). But this mischaracterises the project of the book (as well as structural realism to an extent), which at various points motivates a modification of pure or ontic structuralism to capture the nature of linguistic patterns. For instance, in chapter 6, I argue that the primary means of word individuation involves identifying their structural relations within syntactic structures but I add that "I'm not saying that words are to be individuated exclusively by means of their syntactic profiles" (Nefdt 2023, p. 113). I go on to state that the reason for the emphasis is that this important element (the syntactic analysis) has been missing from the debate about the ontology of words.[3]

The view here resembles other applications of structuralism to the special sciences such as Kincaid (2008) and Ross (2008) (both cited in the book). The structuralism of Nefdt (2023) is one of the priority of structure in methodological and ontological debates on language not its exclusivity. Stanton seems to appreciate this point when she claims that my "naturalist commitments that are ontologically prior" entail that I am "not strictly a structuralist" (pg 7). This comes a lot closer to the central aim of the book, i.e. showing that structuralism is not incompatible with naturalism. For this, structuralism needs to be considered a claim about priority and indispensability, not insularity or exhaustivity. Perhaps this is a principled point of disagreement between Stanton and myself, i.e. that structuralism cannot admit the significance of non-structural elements. But I would argue that this extreme view, although represented in the literature (under the banner of 'structure is all there is', see French and Ladyman (2003)),

[3]In the more mathematical parts of the book, I also follow Parsons (1990) on 'quasi-concrete' structures as a clearer link with linguistics (pg 63).

is not the dominant account of either purely mathematical structures or those revealed by the special sciences.

I think what Stanton's example does show is that the train route mappings do have a syntax in some sense. Linguists call this 'the duality of patterning'. As Hurford (2014) notes:

> Linguists call grammar 'syntax'. Syntax, in its simplest sense, is not necessarily connected with meaning. Any system with rules for combining elements has a syntax [...] Among human activities, music has a syntactic structure. (126)

Train routes and musical notes are not biological systems in the same way that the languages that emerge from complex linguistic communities are.[4] The patterns that can be compressed to represent a piece of music or a transportation grid are not identical to those to be found in a linguistic biological system (even if there might be interesting structural overlaps). What makes the syntax of a train route map distinct from that of natural language is the underlying reality that the patterns compress, which in the case of the latter I argue are generated by linguistic communities. I do have a unique view on what a linguistic community, reconstituted in terms of systems biology, is (see chapter 5 of the book). In fact, my view is permissive in some sense since it implicates the larger environment (such as translation apps) into linguistic structure but it stops short of the station.

4 Metasemantics and structure

In Stanton's final critique she questions the extension of the structuralist picture to meaning, both semantic and pragmatic. She mounts a compelling case against what she calls 'the secret structuralists argument' or the idea that semantics is a structural enterprise as I claim. First she details my use of levels of abstraction and information theory to capture the intricate relationships between various subdisciplines of linguistics and alternative frameworks to the architecture of generative grammar (such as model-theoretic syntax and Jackendoff's parallel architecture). Although skeptical of certain practicalities she concludes that

[4] What Stanton seems to be missing at times is the distinction between formal (redundant) real patterns and the non-redundant variety at a different scale. For me, this is essential. I characterise formal patterns in terms of an analogy with groups and mathematical structuralism, but take pains to show that languages do not exist at this purely formal level (hence my opposition to Platonism about linguistics).

the view is both "novel and innovative". Her main line of contention is along a different, more metasemantic track.

Essentially, Stanton's argument is that although formal semantics is often modelled with mathematics, it is not mathematical theory. She puts the point strongly:

> To say that the meaning of 'dog' is determined by the model-theoretic representation is, on this view, a category error akin to saying that the weather is determined by the mathematical model of a weather system, instead of by the meteorological facts. (Stanton 2025, p. 10)

Metasemantics is the pursuit of the facts by which morphemes, words, phrases, and sentences get their meanings. Stanton rightfully avows that formal semantics (which is my core case in the book) might look secretly structural but it only covers a small formalised fragment of a larger unformalised picture. She is also correct to say that lexical semantics is less so characterised (although I might qualify this position based in my chapter on the ontology of words).

However, to say that I confuse the modelling (in semantics) for reality or commits a category mistake is a strong charge. But I think it stems from the same mischaracterisation of the project discussed in section 3. I am not arguing that what grounds semantic facts are all and only mathematical structures *tout court*. In fact, I explicitly states as much:

> The structures that ground our meaningful use of words and expressions might indeed be mental instructions to the interfaces or conventions grounded in linguistic practice. (Nefdt 2023, p. 166)

The point I am simply making is a holistic one, i.e. "the meaning of a term is determined by its place in a larger linguistic system". This simple dictum is what links distributional semantics to model-theoretic semantics and other views. I see the seminal principle of semantic compositionality, which takes complex meaning to be functionally determined by a mereological part-whole structure (Nefdt and Potts 2024; Pelletier 2012), as an indication of this. The 'semantic metastructuralism' simply posits that there are small clusters of structural information that ground the meanings of parts of language. Again, these could be cashed out in myriad ways such as conventional, normative, externalist or internalist (likely some amalgam). The interesting corollary is that for me, to be an English speaker is not to master or instantiate the entire English

language (whatever that could mean) but only smaller structural parts of it (based on the resources and patterns of your linguistic community *qua* biological system at time t_1). Taken to it's limits, this might entail a nonmonotonic model of natural language meaning. But I admit that it is merely a sketch of how the structuralist picture offered might be extended to metasemantics. In other words, I draw on the methodology of semantics across frameworks (which has a distinctively structural element) to push for the identification of possible ontological components which might engender. It is rather high level. But a fuller picture, as Stanton points out, will certainly need more details. Perhaps the secret structuralist argument is not that meaning is secretly grounded in contentless mathematical structure, which would be an odd metasemantic theory, but rather, as many moderate structuralist accounts of science similarly show, that meaning is secretly more structural than it often appears to be in the wild. The book as a whole aims to reveal that there is hidden truth in the structures that our sciences use to model the world.

5 Conclusion

Ultimately, Stanton's critique trades in absolutes, which the book does not. Nor can it when something so complex, messy, and intricate as natural human language is involved. Nevertheless, her arguments are well-informed and philosophically illuminating. They force further consideration and, in some parts, serious rethinking of many of the central statements and theoretical commitments of both a real pattern analysis of language and a structural realist account of linguistics. I would do well to take heed of these arguments in future work on the subject.

References

Dennett, Daniel C. (1991). "Real Patterns". In: *The Journal of Philosophy* 88.1, pp. 27–51. DOI: 10.2307/2027085.

French, Steven and James Ladyman (2003). "Remodelling Structural Realism: Quantum Physics and the Metaphysics of Structure". In: *Synthese* 136.1, pp. 31–56. DOI: 10.1023/A:1024156116636. URL: https://doi.org/10.1023/A:1024156116636.

Futrell, Richard, Roger P. Levy, and Edward Gibson (2020). "Dependency Locality as an Explanatory Principle for Word Order". In: *Language* 96.2, pp. 371–412. DOI: 10.1353/lan.2020.0024.

Hurford, James (2014). *Origins of Language: A Slim Guide*. Oxford linguistics. OUP Oxford. ISBN: 9780198701880. URL: https://books.google.co.za/books?id=R3TiAgAAQBAJ.

Kincaid, Harold (2008). "Structural Realism and the Social Sciences". In: *Philosophy of Science* 75.5, pp. 720–731. DOI: 10.1086/594517.

Kolmogorov, Andrey (1963). "On Tables of Random Numbers". In: *Sankhyā: The Indian Journal of Statistics, Series A* 25.4, pp. 369–376.

Ladyman, James and Don Ross (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Miller, Philip (1999). *Strong Generative Capacity: The Semantics of Linguistic Formalism*. CSLI Publications.

Millhouse, Tyler (2021). "Really Real Patterns". In: *Australasian Journal of Philosophy* 100.4, pp. 664–678. DOI: 10.1080/00048402.2021.1941153. URL: https://doi.org/10.1080/00048402.2021.1941153.

Nefdt, Ryan M. (2016). "Scientific Modelling in Generative Grammar and the Dynamic Turn in Syntax". In: *Linguistics and Philosophy* 39.5, pp. 357–394.

— (2023). *Language, Science, and Structure: A Journey into the Philosophy of Linguistics*. New York: Oxford University Press. ISBN: 9780197653098. DOI: 10.1093/oso/9780197653098.001.0001. URL: https://academic.oup.com/book/46518.

Nefdt, Ryan M. and Christopher Potts (Dec. 2024). "Compositionality". In: *Open Encyclopedia of Cognitive Science*. Ed. by Michael C. Frank and Asifa Majid. https://oecs.mit.edu/pub/e222wyjy. MIT Press.

Parsons, Charles (1990). "The Structuralist View of Mathematical Objects". In: *Synthese* 84.3, pp. 303–346. DOI: 10.1007/bf00485186.

Pelletier, Francis Jeffry (2012). "Holism and Compositionality". In: *The Oxford Handbook of Compositionality*. Ed. by Markus Werning, Wolfram Hinzen, and Edouard Machery. Oxford University Press.

Pullum, Geoffrey K. (2013). "The Central Question in Comparative Syntactic Metatheory". In: *Mind and Language* 28.4, pp. 492–521.

Ross, Don (2008). "Ontic Structural Realism and Economics". In: *Philosophy of Science* 75.5, pp. 732–743. DOI: 10.1086/594518.

Stanton, Kate (2025). "On Linguistically Real Patterns". In: *Analysis*. DOI: 10.1093/analys/anaf065. URL: https://doi.org/10.1093/analys/anaf065.

Thomson-Jones, Martin (2012). "Modelling without Mathematics". In: *Philosophy of Science* 79, pp. 761–772.