Conceptual diversity of linguistic structure: a mixed-methods corpus study

Wiktor Rorot (w.rorot@uw.edu.pl)[1,2] **& Joanna Rączaszek-Leonardi**[1]
[1]Human Interacticity and Language Lab, Faculty of Psychology, University of Warsaw
[2]Institute of Philosophy and Sociology, Polish Academy of Sciences

Abstract

The goal of this paper is to uncover how conceptualizations of linguistic structure differ between the practice of research on language emergence and standard, formalized approaches of philosophy of language and theoretical linguistics. We analyze a corpus of 5000 scientific articles within the field of language emergence. Adopting a mixed-methods approach, we combine computational analysis—semantic search, topic modelling, and collocation analysis—with qualitative close reading. Results indicate that researchers in language emergence view linguistic structure as a complex phenomenon, which cuts across some existing theoretical distinctions, most importantly between "compositional" and "contextual" accounts. These results indicate the need for a new view of the systematicity of linguistic structure.

Introduction

One of the key questions in the field of language emergence concerns the sources and (cognitive and social) mechanisms underpinning the structure of language. Here, we present the results of a corpus study on the scope and diversity of conceptualizations of linguistic structure.

In our earlier work (Rorot & Rączaszek-Leonardi, 2023), we pointed out that the domain of language emergence seems to focus on only selected aspects of language structure. We argued that the terms used in the domain often have specific, narrowed meaning, contrasting with other domains of research on language and not giving justice to the multiple sources of language structure (Zubek et al., 2024).

For example, we noted that the majority of works target the productivity and open-endedness of language (e.g., Chomsky, 1965/1970; Jackendoff, 1997; Planer & Sterelny, 2021), i.e., the property of language that allows infinite productions from a finite set of constituent elements (e.g., Collier et al., 2014; Engesser & Townsend, 2019). This property is commonly conceptualized as *compositionality* (see Gontier et al., 2024) and defined as the ability of discrete, meaningful elements—morphemes, words—to form larger constituents—words, sentences—with meaning derived from that of the elements:

> The meaning of a compound expression is a function of the meanings of its parts and of the way they are syntactically combined. (Partee, 2004, p.152)

In our prior analyses of the theoretical literature we noted that this focus is often further narrowed: while linguistics and philosophy of language have focused on viewing compositionality on the "level of meaning", in the study of language emergence the use of the term targets the "level of form". The meaning of the constituents and compound expressions is assumed to be stable, or bracketed, and the focus is placed on the complexity of the *structures* that emerge in communication (Rorot & Rączaszek-Leonardi, 2023) at the cost of the work on the emergence of meaning of compound expressions.

Due to this focus and limitations, the aspect of the flexibility of meaning within linguistic structure has not been treated with an equal attention. This concerns both the principle of contextuality, as formulated by Frege (1960, p. xxii):

> [N]ever ask for the meaning of a word in isolation, but only in the context of a proposition.

and other, more extreme, cases of flexibility of meaning exemplified by holism (Pelletier, 2012), where meanings can only be attributed to complete utterances (Arbib, 2012; Smith & Kirby, 2012) and where knowing the standard meaning of the elements will not be helpful in deriving the meaning of an expression (e.g., as in the idiom "to kick the bucket").

Such aspects of linguistic structure thus largely fall out of the purview of the work on language emergence, even though the non-compositional—holistic or contextually dependent—structures are numerous and have a central role in meaning generation. The most common examples are idioms and irregular verbs, but any construction which is "non-trivially" compositional (Steinert-Threlkeld, 2020) manifests this property. Arguably, this is the case for deriving semantics of almost any compound expression: consider what "red" means in "red apple" compared to "red brick".

These considerations led some researchers to formulate an anti-realist, instrumentalist view of compositionality as a superficial property, a "theory-internal [construct] of formally-inclined branches of the philosophy of language" (Steffensen & Harvey, 2018, p. 12) which should be abandoned as an explanatory target. To an extent, we agree with those criticisms (see Rorot & Rączaszek-Leonardi, 2023). In this context, we propose to view formal principles of compositionality, contextuality, and holism as idealizations of co-existing "forces" which help explain the generation of meaning within complex communication systems. Effectively, we put the *systematicity* of language as the proper explanatory target for understanding the consistency and predictability of complex expressions in natural language.

However, both Steffensen and Harvey (2018) and Rorot and Rączaszek-Leonardi (2023) based their arguments on the review and close reading of primarily theoretical literature. In the current paper, we set out to confirm our claims by novel methods of text analysis and expand the scope by including literature on modelling and empirical studies in the domain. To realize these aims, we conducted a computational analysis of a corpus of approximately 5000 full text articles related to language emergence research and supplemented it by close reading of relevant samples from the corpus, identified with quantitative methods.

Based on the above arguments, we hypothesized that "compositionality" will be a central conceptualization of linguistic structure in debates on language emergence, while the alternatives—in particular, "contextuality" and "holism"—would be mentioned sparsely and in separation from compositionality. Results indicate that compositionality plays a significantly less central role than hypothesized, and that compositional and contextual properties of meaning are often considered in tandem. The quantitative methods allowed us to explore the conceptual map of linguistic structure in greater detail, recognizing different areas the various conceptualizations of structure address. We conclude that the existing theoretical accounts trail behind the actual conceptual practice in language emergence research, and we point to the directions in which they can be advantageously developed.

Methods

The methodology of the study follows the approach of digital philosophy of science (see Lean et al., 2023; Miłkowski & Nowakowski, 2025; Pence, 2022, 2025; Pence & Ramsey, 2018; Rorot, 2025). Drawing on tools advanced in digital humanities and computational linguistics, proponents of this approach construct mixed-methods or multi-level methodologies (e.g., Murdock et al., 2017; Rorot, 2025; Rorot & Miłkowski, 2024), which connect large scale quantitative analysis of corpora of scientific texts with qualitative close reading using standard philosophical methodologies. The research design is summarized in Figure 1, each step is detailed below.

We use the Semantic Scholar Open Research Corpus (S2ORC, version from August 8, 2024, see Kinney et al., 2025; Lo et al., 2020) which features over 12 million full-text, open access articles and is currently the largest openly available corpus of academic papers. Additionally, the corpus provides pretrained vector embeddings (using the SPECTERv2 LLM, Cohan et al., 2020) for full texts of the articles (Singh et al., 2023, we will refer to this dataset as SPECTER Embeddings).

Step 1. Subcorpus selection and preparation To select a corpus of papers related to language emergence we identified a subset of 5000 articles using semantic search based on a seed list (we will refer to it as the COMP5000 subcorpus). The seeed list consisted of 200 papers that we judged to belong to the field of language emergence, broadly construed. The construction of this seed list was based on the expert knowledge

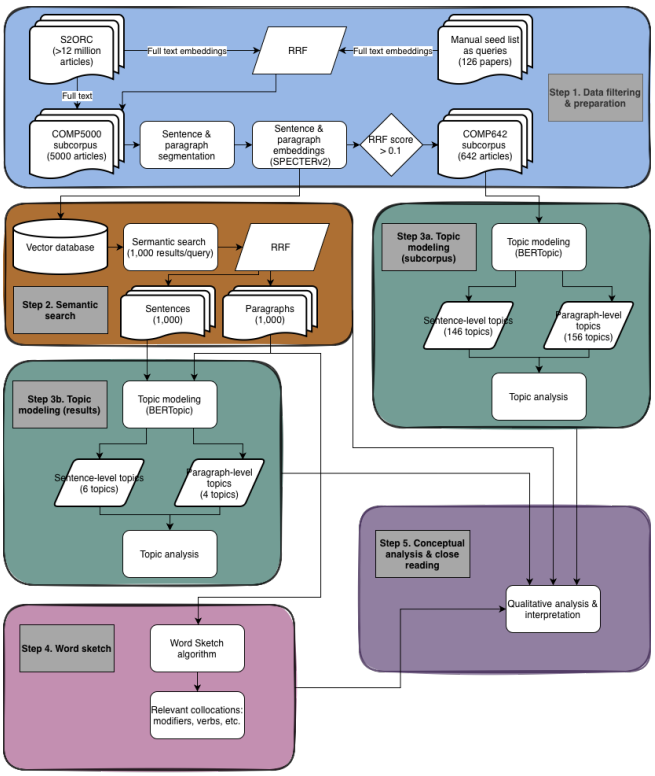

Figure 1: Flowchart summarizing the research design. Individual steps are discussed in detail in the main text.

of the Authors, drawing on papers that were relevant in our previous research and teaching, and expanded with keyword search on Google Scholar. This list was reduced to 126 articles available in the SPECTER Embeddings dataset.[1] Using full-text embeddings, for each seed article we selected 10,000 articles from S2ORC (ranked by cosine distance). This resulted in a total of 211,649 unique articles (due to overlap between the results of individual queries). We used Reciprocal Rank Fusion algorithm (henceforth RRF, Cormack et al., 2009), which combines multiple ranked lists into a unified ranking by summing inverse ranks (shifted by a constant value) across the lists. We selected 5,000 articles which ranked best across all queries (among these, 321 were not English language and were excluded from further analysis, as the model used is not validated for multilingual datasets).

The articles were automatically segmented into sentences (with the spaCy Python package) and split into paragraphs (using annotations from S2ORC where available, with an arbitrary fallback of 10 sentence-long chunks). A total of 1,166,815 sentences and 289,208 paragraphs were identified.

[1] The online supplementary materials available at: https://doi.org/10.17605/OSF.IO/TFW5D include the full list of seed articles, subcorpus metadata, detailed results and (interactive) visualizations of all analysis reported in the article. The code used for the analyses and visualizations conducted for the study is available on GitHub: https://github.com/wikror/digital-philsci-tools (accessed on May 4, 2026).

For each paragraph and for each sentence, a vector representation was calculated using the SPECTERv2 LLM.

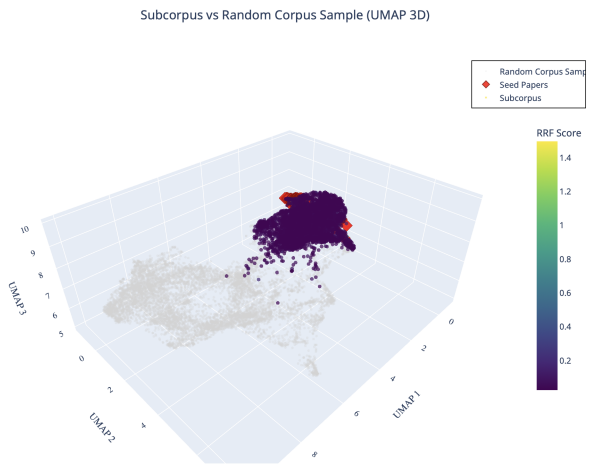

Figure 2: A 2D view on the 3D UMAP visualization of article embeddings from the SPECTER Embeddings dataset. Red diamonds indicate the papers from the seed list. Documents from the subcorpus are marked according to the RRF value (see the scale). Grey points indicate a pseudo-random sample of 5,000 articles from the whole dataset. The distribution of articles selected parallels the distribution of the seed list.

Step 2. Semantic search Using a vector database, we have queried the sentence- and paragraph-level vector embeddings of the subcorpus using queries based on theoretical definitions of structural concepts (see next section). For each query, the search produced 1,000 results (arbitrarily selected parameter). Results shorter than 16 characters were omitted to limit the appearance of short headers and similar fragments containing 2–3 words. Using RRF, 1,000 sentences and 1,000 paragraphs which ranked best across all queries were selected for further analysis.

Step 3a. Topic modelling of the subcorpus To establish how significant are discussions of linguistic structure within the corpus, we conducted topic modelling of a subsample of the corpus. We used full text of articles which achieved an arbitrary threshold RRF score ≥ 0.1, calculating topic models at both sentence and paragraph granularity. This reduced the subcorpus to 642 highest ranking articles, with 60 excluded as non-English. Summarily, these articles contained 171,351 sentences and 39,835 paragraphs (we will call this corpus COMP642).

We used the BERTopic algorithm (Grootendorst, 2022), which relies on vector embeddings, in our case SPECTERv2 embeddings calculated in Step 1. Subcorpus selection and preparation. BERTopic uses dimensionality reduction and clustering algorithms in an unsupervised procedure which identifies semantic clusters in the dataset which are treated as "topics". For each such topic, BERTopic constructs a representation from most characteristic words (using a version of term frequency—inverse document frequency algorithm, see Spärck Jones, 1972) or vector embeddings (using the KeyBERT algorithm, see Grootendorst, 2020). The number of topics is selected by the algorithm, but the topics can be merged based on similarity of the vector embeddings of representations.

Step 3b. Topic modelling of semantic search results To highlight large-scale patterns in the results of semantic search produced in Step 2. Semantic search we conducted topic modelling of both sentence- and paragraph-level results. The methodology of the topic modelling is analogous to the analysis conducted in Step 3a. Topic modelling of the subcorpus.

Step 4. Word Sketch Fine-grained semantic structure of relevant terms (e.g., "compositionality", "contextuality") is explored through the Word Sketch algorithm (Kilgarriff et al., 2014). Word Sketch uses lemmatized representation of the corpus with syntactic tags which allows for an analysis of collocations of target words, retaining the syntactic structure (e.g., separating modifiers or verbs for which the target is a subject or an object). The Word Sketch algorithm produced lists of top 20 collocates (using a logDice statistic to establish typicality of co-occurrence, see Rychlý, 2008) from different syntactic categories.

Step 5. Conceptual analysis and close reading The quantitative data produced in the course of computational analysis, as well as the sentences and paragraphs selected in the semantic search, are submitted to close reading and conceptual analysis. This step follows the standard methods of philosophy of science (e.g., see Braddon-Mitchell & Nola, 2009) and conceptual engineering (Cappelen, 2018; cf. Rorot and Miłkowski, 2024; Rorot, 2025). There are important limitations to the evidence provided by computational analysis (see Pence, 2025) which these qualitative, interpretative procedures can overcome.

Results

Semantic search results (Step 2) We start from the analysis of semantic search results. The queries used included sentences such as (see the online supplement for a full list):

1. The meaning of an expression is a function of its parts and of their syntactical structure (compositionality query)
2. The meaning of an expression is a result of the context in which it is used (contextuality query)
3. Meaning can only be attributed to complete utterances (holism query)

We focus on paragraph-level results since they provide greater interpretability (compared to sentence-level results). Consider the following example (all analyses reported use the full results):

> The meanings of linguistic expressions are grounded neither in the world, nor in an internal idea of the world, but

> in the linguistic corpus as a whole. The holistic role of language implies that the meanings of individual linguistic expressions are grounded in multiple relations with respect to other linguistic expressions and their combinations. […] Thus, the meaning of an expression is not something essential that lies behind the expression or to which the expression refers. […] If an expression is to represent something, then it represents its place in the complexity of relations to all other expressions. (corpus ID: 264490710, paragraph number: 49, rank: 9)

As expected, compositionality is acknowledged as a principle guiding the interaction of syntactic and semantic properties. For instance one sentence reads: "[Proponents of minimalism and contextualism] both acknowledge the importance of lexical meaning, syntactic structures (the semantic glue that combines lexical constituents), and the role of context in adding levels of meaning." (229170128, par. 53; emphasis added). Nevertheless, as in this example, it is often accompanied by reference to contextual or (more rarely) holistic properties of language. As this is a common pattern across many paragraphs, this suggests a less central role of "compositionality" than initially hypothesized.

One significant recurring aspect in the results is reference to the debate on the origins or the primary function of language: either as a medium of thought (e.g., Fodor, 1975; Sellars, 1974) or a communicative tool (e.g., Deacon, 1998; Fedorenko et al., 2024). We will return to this point in the discussion, see Linguistic and cognitive structures.

Full subcorpus topic models (step 3a) The paragraph-level topic model for COMP642 allows for placing concepts of linguistic structure on the broader map of the debates in language emergence. BERTopic automatically created a fine-grained model with 156 (including outliers) topics. Several topics referenced "structure", and two specifically center conceptualizations of structure (in this case, "compositionality"):

- topic 140 (139 paragraphs), represented by "syntactic prose", "idioms", "sentences"; discusses studies comparing the neurological underpinnings of sentence structure processing, and its relation to meaning;
- topic 145 (112 paragraphs), represented by "compositionality", "iterated", "generation"; uses the "iterated learning" paradigm to model the emergence of compositional structure.

The remaining topics that mention structure highlight the different experimental and modeling paradigms used to investigate structure: for instance, different forms of priming (topics 80, 82), graph models (topics 143, 148) or information-theoretic measures (topic 110).

Semantic search results topic models (step 3b) Crucial insights are provided by the topic models for the semantic search results. At both sentence- and paragraph-level, BERTopic automatically chose small models with 5 and 3 topics respectively. These model sizes were robust across a range of hyperparameters tested. This is likely due to the small size of semantic search results, but also indicates their semantic coherence. The topics (with KeyBERT representations) at the sentence level are the following:

1. **MEANING** (meaning, linguistic meaning, lexicon, interpretation, semantic)
2. **LINGUISTIC THEORY** (linguistics, linguistic theory, theoretical linguistics, linguistic forms, different linguistic)
3. **LANGUAGE** (language user, natural language, compositional language, syntax language, formal linguistics)
4. **SYNTAX** (syntactic structure, syntactic semantic, phrase structure, lexical item, semantic)
5. **SENTENCE** (utterance, consider sentence, meaning utterance, utterance level, linguistic)

At the paragraph level, the topics are following:

1. **LANGUAGE** (language change, linguistics, human language, lexical, grammar)
2. **SENTENCE** (syntactic, lexical, parsing, semantic, grammar)
3. **MEANING** (linguistic expressions, linguistic, linguistic reality, meanings, linguistics)

The topics are relatively abstract and indicate that the discussion of linguistic structure is related to key areas of research within the field of language emergence. Indeed, they underpin the debates on foundational questions regarding language in the analyzed corpus.

The review of representative documents (see online supplementary materials) assigned to the different topics shows that the structuring principles outlined in the introduction are commonly used. These principles are often mentioned together. In that context an important research question is how the different structural properties interact and whether they can be disentangled. For instance, references to "syntactic information" indicate that the information provided by syntax can be isolated, while some other examples point to the unique position of one of the structural properties (e.g., see the paragraph above in Semantic search results (Step 2)).

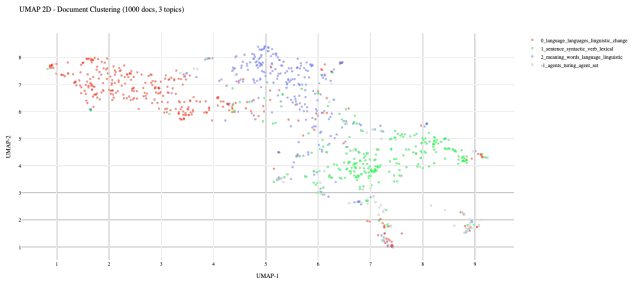

Figure 3: A 2D UMAP visualization of paragraph-level semantic search results. Colours indicate topics ascribed to individual documents by BERTopic. The topic numbered "-1" collects the outliers.

Word Sketch collocations (step 4) Word Sketch found relevant collocations only for "compositionality". This suggests that "contextuality" and "holism" appear too infrequently to identify any consistent collocation patterns. For adjectives, a similar number of collocations was found for "contextual" and "compositional", while "holistic" or "holophrastic" had only a few consistent collocates ("moderately holistic", holistic "property", "role", or "approach", and only "holophrastic construction").

For "compositionality" theoretically interesting results include its relation to "productivity" (its modifier in both noun and adjective forms), and its links to "idiomaticity" or "context" (an "and/or" relation according to Word Sketch). This indicates the previously signalled tendency towards a more complex view on language structure, where compositionality does not exclude other structural properties. The collocates for "compositional" are more informative, as it is used as a modifier for words such as "truth", "referent", "prosody", "semantics", "behavior", "mechanism", "structure", "pattern", "meaning", "grammar" or "language", among others. "Contextual", on the other hand, modifies "assumption", "information", "meaning", "category", or "property", while itself being characteristically modified by "inherently". This largely tracks the expected behavior of these terms.

Interestingly, the collocates of "structure" refer to several levels or domains: from "syntactic" or "sentence" structure, through "logical" or "symbolic" to "conceptual", "semantic", or "information". In verb form, "to structure" is considered with objects such as "cognition", "language", and "meaning", as well as with the subject "mind". This relation—of either "mind structuring language" or "language structuring cognition"—highlights that the discussion of structure often takes into account the function of language, as we've already noted in the qualitative survey of the semantic search results, see Semantic search results (Step 2).

Discussion

Explanation and description Building on existing reviews of theoretical literature on linguistic structure (in particular Rorot & Rączaszek-Leonardi, 2023; Steffensen & Harvey, 2018), we initially assumed that "compositionality" will act as a central conceptualization of linguistic structure. Our mixed-methods study of the broader field of language emergence, including empirical and modeling papers, does not immediately support this conclusion.

The semantic search (step 2) of the most relevant references to linguistic structure and the close reading of selected fragments (step 5) reveals that the notion of "compositionality" is frequently analyzed together with notions of "contextuality" or "holism" and highlights the theoretical connection between the function of language and its structure. While the analysis of the full subcorpus (step 3a) indicates that explicit problematization of linguistic structure is not a central topic for the field, it shows that the discussion of structure recurs at different levels of linguistic phenomena. The detailed topic model of semantic search results (step 3b) links discussions of structure to key problem areas in language research. In these contexts, researchers deal simultaneously with multiple aspects of linguistic structure, showing their entanglement. This final point is further confirmed by Word Sketch collocations (step 4).

These results point to a complex theoretical landscape underpinning research on language emergence, in line with what we suggested previously on theoretical grounds (Rorot & Rączaszek-Leonardi, 2023). The adopted methodology allows for identifying main lines of existing and required work.

Compositional, contextual, and holistic properties of language are important explanatory targets, commonly regarded as mutually interdependent. At the "level of form" (Rorot & Rączaszek-Leonardi, 2023), as indicated by topic models and collocations, the properties refer to word order relationships as well as to morphology, or more generally to an abstracted, logical structure of complex linguistic expressions. Most of the results clearly refer to only one of those levels: either to "form" or to "meaning", and those that mention both seem to commonly assume that they are at least partially independent. This connects to the broader debate on the relation of syntax, semantics, and pragmatics.

Within each of these levels, mentions of at least one of the three principles (compositionality, contextuality, holism) almost inevitably require a reference to another, as a way to highlight the limits of any one of those conceptualizations. As the results indicate, it is impossible to capture language in neither fully compositional nor fully contextual or holistic terms. This impacts more "practical" aspects, such as discussions of polysemy or semantic ambiguity, e.g., in the case of pronouns (visible in results of semantic search).

At the same time, compositionality clearly has an epistemically privileged position. It is directly associated with the framework of generative grammar, and can be modelled and simulated with iterated learning paradigms, graph or network models, as well as studied with neuroscientific approaches (drawing on the topics identified in the COMP642 subcorpus, see Full subcorpus topic models (step 3a)). Importantly, generative grammar and iterated learning paradigm are motivated by opposed theoretical positions, hence the fact that they share the formal apparatus of "compositionality" indicates the robustness of that apparatus. Analogous robust research toolbox is still lacking for contextual or holistic accounts of language—at least in our dataset. Resultingly, these concepts are not considered as explanatorily powerful as "compositionality".

Linguistic and cognitive structures As we already noted in the discussion of the results of semantic search, the relation between cognitive and linguistic structure is a recurring consideration in the analyzed dataset. An analogous relation between linguistic and cognitive structure seems to be a common practical assumption in the study of emergence of compositional languages. For instance, in the iterated learning paradigm (Kirby et al., 2008) it is assumed that the inputs the model receives have a compositional structure (see also An-

dreas, 2019; Korbak et al., 2020)

However, this is a strong assumption in the context of biological systems and their relation to the environment, which are likely to be structured in complex, nonlinear fashion. The assumption that cognitive processes impose a compositional structure onto percepts, with its "atomic" view on word meanings, requires a substantivist view on the object of perception, reified by the classical concept of stimulus as physical energies. Gibson (1960), drawing on Stevens (1951, p. 32), suggested characterizing psychological stimuli instead in terms of "all the transformations of the environment, both external and internal, that leave the response invariant." This ecological viewpoint escapes the requirements of compositionality (see Rorot & Rączaszek-Leonardi, 2023): the individual "objects" and their "properties" are entangled bundles of (relational) features which interact in complex, nonlinear fashion. Simply, nothing "composed" is *given* in perception.

Instead, such ecological views often underscore the role of context in perception, and ecological psychology in particular highlights the direct availability of information in the environment (Gibson, 1986). This translates into a possible explanation of contextuality in language processing: cognitive processes geared towards picking out ecological information might be reused for processing information in linguistic behaviors. Considering there's a growing number of models which aim to describe mechanism capable of handling ecological information in perception (Falandays et al., 2021, 2023; Raja, 2018, 2019, 2021, 2024), this is a promising avenue for future research, as these models can be (and in some cases, have been) adapted to the study of linguistic capacities.

In the described theoretical landscape, the affinity between linguistic and cognitive structures is not a decisive argument in favor of either compositionality or contextuality, as both aspects of structure relate to existing models of cognition.

Diversity of structure What the results clearly indicate is that there are multiple different kinds of conceptual diversity intersecting in the considerations of linguistic structure. All our corpus analyses indicate that the notion of compositionality extends across multiple levels of linguistic behavior, impacting syntax, semantics, and likely pragmatics, but also beyond, playing an important role in neural network research (e.g., Hupkes et al., 2020). However, across these different levels, the references to compositionality appear to be targeting the same (kind of) phenomena. Instead, the analyses of contextuality exhibit a more explicit diversity.[2] In this case, the notion of context is considered largely separately on the level of syntax, semantics, or pragmatics (e.g., Sowa, 1995).

Griffiths (2015) distinguishes diversity in terms of the range of concepts and theoretical frameworks available to researchers for characterizing phenomena of interest (external diversity), and the co-existence of multiple different definitions of key notions (internal diversity). The former seems to characterize how compositionality is approached, as despite the disciplinary differences, there seems to be an agreement among the researchers regarding its characterization as an explanatory target. The latter, in turn, seems to apply to how contextuality is considered, as the variety of different contexts refers to distinct phenomena.

Both types of diversity have been variously considered either as a roadblock for the progress of science, or its condition of possibility (e.g., Griffiths, 2015). In perspectival views on scientific realism it is commonly accepted that the robustness of properties targeted in scientific research, i.e., the convergence of various independent methods, is a marker of their reality (e.g., Massimi, 2022; Mitchell, 2002; Wimsatt, 2007). This form of integrative pluralism seems well-suited to explain the epistemic value assigned to compositionality in the study of linguistic structure. The various methodologies and contexts in which compositionality is approached converge on the same phenomenon, providing a multilayered understanding of that explanatory target. Instead, the current lack of unity in the study of contextuality, exemplified by the variety of "contexts", prevents a meaningful exchange of results across the different perspectives.

Summary

In this paper we attempted to scrutinize the conceptual landscape of language emergence studies to explore how researchers conceive of linguistic structure and what are the practices of the use of key terms describing it. To this end, we employed quantitative analyses to help us trace the diversity, epistemic roles, and identify possible gaps in concepts used. Compared to standard theoretical review of literature, such a large-scale corpus study based on quantitative methods extends the available evidential horizon, allowing for the exploration of a greater variety of examples. When paired with close readings of computationally selected examples, the mixed-methods approach adopted here identifies a preference towards using multiple different conceptualizations of structure in literature on language emergence. While this diversity is clearly necessary to capture full complexity of language structure, the lack of an overarching theoretical treatment which explicitly considers the intertwinement of compositional, contextual, and holistic properties hinders the understanding of how language structure emerges. It also impedes the synthesis of language emergence work, in which all the above concepts do appear, although with differing roles and significance. These results motivate the direction of the future theoretical projects directed at understanding linguistic structure in its many aspects and levels.

Acknowledgments

The authors wish to thank the members of the HILL Lab at the Faculty of Psychology, University of Warsaw, and the partic-

[2]For instance, drawing on Word Sketch results, "contextual information" refers to information available in the environment, and "contextual property" to the fact that properties are significant only in a given situation. In the first case context is reified into "environment", whereas in the second instance it is considered in counterfactual terms.

ipants of the 14th Peripatetic Conference on Dynamical Systems Modeling, Zakopane, Poland (October 2025), where an early version of this work was presented.

The work on the paper has been funded by an internal grant at the University of Warsaw as part of the "Excellence Initiative—Research University", action II.3.9 "Knowledge Lab: Scientometrics for scientific excellence", awarded to JRL.

References

Andreas, J. (2019). Measuring Compositionality in Representation Learning. *arXiv:1902.07181 [cs, stat]*.

Arbib, M. A. (2012). Compositionality and Beyond: Embodied Meaning in Language and Protolanguage. In W. Hinzen, E. Machery, & M. Werning (Eds.), *Oxford Handbook of Compositionality* (pp. 475–492). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199541072.013.0024

Braddon-Mitchell, D., & Nola, R. (Eds.). (2009). *Conceptual analysis and philosophical naturalism*. MIT Press.

Cappelen, H. (2018). *Fixing language: An essay on conceptual engineering* (First edition). Oxford University Press.

Chomsky, N. (1970). *Aspects of the theory of syntax*. The MIT Press. (Original work published 1965)

Cohan, A., Feldman, S., Beltagy, I., Downey, D., & Weld, D. S. (2020, May). SPECTER: Document-level Representation Learning using Citation-informed Transformers. https://doi.org/10.48550/arXiv.2004.07180

Collier, K., Bickel, B., Van Schaik, C. P., Manser, M. B., & Townsend, S. W. (2014). Language evolution: Syntax before phonology? *Proceedings of the Royal Society B: Biological Sciences*, *281*(1788), 20140263. https://doi.org/10.1098/rspb.2014.0263

Cormack, G. V., Clarke, C. L. A., & Buettcher, S. (2009). Reciprocal rank fusion outperforms condorcet and individual rank learning methods. *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 758–759. https://doi.org/10.1145/1571941.1572114

Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. Norton.

Engesser, S., & Townsend, S. W. (2019). Combinatoriality in the vocal systems of nonhuman animals. *WIREs Cognitive Science*, *10*(4), e1493. https://doi.org/10.1002/wcs.1493

Falandays, J. B., Nguyen, B., & Spivey, M. J. (2021). Is prediction nothing more than multi-scale pattern completion of the future? *Brain Research*, *1768*, 147578. https://doi.org/10.1016/j.brainres.2021.147578

Falandays, J. B., Yoshimi, J., Warren, W. H., & Spivey, M. J. (2023). A potential mechanism for Gibsonian resonance: Behavioral entrainment emerges from local homeostasis in an unsupervised reservoir network. *Cognitive Neurodynamics*. https://doi.org/10.1007/s11571-023-09988-2

Fedorenko, E., Piantadosi, S. T., & Gibson, E. A. F. (2024). Language is primarily a tool for communication rather than thought. *Nature*, *630*(8017), 575–586. https://doi.org/10.1038/s41586-024-07522-w

Fodor, J. A. (1975). *The language of thought*. Crowell.

Frege, G. (1960). *The Foundations of Arithmetic* (J. Austin, Trans.). Harper & Brothers. (Original work published 1884)

Gibson, J. J. (1960). The concept of the stimulus in psychology. *American Psychologist*, *15*(11), 694–703. https://doi.org/10.1037/h0047037

Gibson, J. J. (1986). *The ecological approach to visual perception*. Erlbaum.

Gontier, N., Hartmann, S., Pleyer, M., & Rodrigues, E. D. (2024). Combinatoriality and Compositionality in Communication, Skills, Tool Use, and Language. *International Journal of Primatology*, *45*(3), 473–481. https://doi.org/10.1007/s10764-024-00416-8

Griffiths, P. E. (2015). Conceptual change and conceptual diversity contribute to progress in science. In G. Bammer (Ed.), *Change! combining analytic approaches with street wisdom* (pp. 163–176). Australian National University Press.

Grootendorst, M. (2020). KeyBERT: Minimal keyword extraction with BERT. https://doi.org/10.5281/zenodo.4461265

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.

Hupkes, D., Dankers, V., Mul, M., & Bruni, E. (2020). Compositionality Decomposed: How do Neural Networks Generalise? *Journal of Artificial Intelligence Research*, *67*, 757–795. https://doi.org/10.1613/jair.1.11674

Jackendoff, R. (1997). *The architecture of the language faculty*. MIT Press.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, *1*(1), 7–36. https://doi.org/10.1007/s40607-014-0009-9

Kinney, R., Anastasiades, C., Authur, R., Beltagy, I., Bragg, J., Buraczynski, A., Cachola, I., Candra, S., Chandrasekhar, Y., Cohan, A., Crawford, M., Downey, D., Dunkelberger, J., Etzioni, O., Evans, R., Feldman, S., Gorney, J., Graham, D., Hu, F., … Weld, D. S. (2025, April). The Semantic Scholar Open Data Platform. https://doi.org/10.48550/arXiv.2301.10140

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686. https://doi.org/10.1073/pnas.0707835105

Korbak, T., Zubek, J., & Rączaszek-Leonardi, J. (2020). Measuring non-trivial compositionality in emergent communication. *arXiv:2010.15058 [cs]*.

Lean, O. M., Rivelli, L., & Pence, C. H. (2023). Digital Literature Analysis for Empirical Philosophy of Science. *The British Journal for the Philosophy of Science*, *74*(4), 875–898. https://doi.org/10.1086/715049

Lo, K., Wang, L. L., Neumann, M., Kinney, R., & Weld, D. S. (2020, July). S2ORC: The Semantic Scholar Open Research Corpus.

Massimi, M. (2022). *Perspectival realism*. Oxford University Press.

Miłkowski, M., & Nowakowski, P. (2025). Tracking norms: A plea for a normative digital philosophy of science. *Synthese*, *206*(3), 120. https://doi.org/10.1007/s11229-025-05200-6

Mitchell, S. D. (2002). Integrative Pluralism. *Biology and Philosophy*, *17*(1), 55–70. https://doi.org/10.1023/A:1012990030867

Murdock, J., Allen, C., Börner, K., Light, R., McAlister, S., Ravenscroft, A., Rose, R., Rose, D., Otsuka, J., Bourget, D., Lawrence, J., & Reed, C. (2017). Multi-level computational methods for interdisciplinary research in the HathiTrust Digital Library. *PLOS ONE*, *12*(9), e0184188. https://doi.org/10.1371/journal.pone.0184188

Partee, B. H. (2004). *Compositionality in formal semantics: Selected papers of Barbara H. Partee*. Blackwell Pub.

Pelletier, F. J. (2012). Holism and Compositionality. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 149–174). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199541072.013.0007

Pence, C. H. (2022). Testing and discovery: Responding to challenges to digital philosophy of science. *Metaphilosophy*, *53*(2-3), 238–253. https://doi.org/10.1111/meta.12549

Pence, C. H. (2025). Textual Analysis and Conceptual Cartography. In A. Currie & S. Veigl (Eds.), *Methods in Philosophy of Science: A User's Guide*. MIT Press.

Pence, C. H., & Ramsey, G. (2018). How to Do Digital Philosophy of Science. *Philosophy of Science*, *85*(5), 930–941. https://doi.org/10.1086/699697

Planer, R. J., & Sterelny, K. (2021). *From signal to symbol: The evolution of language*. The MIT Press.

Raja, V. (2018). A Theory of Resonance: Towards an Ecological Cognitive Architecture. *Minds and Machines*, *28*(1), 29–51. https://doi.org/10.1007/s11023-017-9431-8

Raja, V. (2019). From metaphor to theory: The role of resonance in perceptual learning. *Adaptive Behavior*, *27*(6), 405–421. https://doi.org/10.1177/1059712319854350

Raja, V. (2021). Resonance and radical embodiment. *Synthese*, *199*(1), 113–141. https://doi.org/10.1007/s11229-020-02610-6

Raja, V. (2024). The motifs of radical embodied neuroscience. *European Journal of Neuroscience*, *60*(5), 4738–4755. https://doi.org/10.1111/ejn.16434

Rorot, W. (2025, December). *Scale-Free Communication? An investigation of the use of the concept "communication" in biology and cognitive sciences* [Doctoral dissertation, University of Warsaw].

Rorot, W., & Miłkowski, M. (2024). Empirical Evidence in Conceptual Engineering, or the Defense of 'Predictive Understanding'. In P. Stalmaszczyk (Ed.), *Conceptual Engineering* (pp. 255–280). Brill | mentis. https://doi.org/10.30965/9783969753026_013

Rorot, W., & Rączaszek-Leonardi, J. (2023). Understanding "Compositionality" in Research on Language Emergence. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*.

Rychlý, P. (2008). A lexicographer-friendly association score. In P. Sojka & A. Horák (Eds.), *Proceedings of Recent Advances in Slavonic Natural Language Processing* (pp. 6–9).

Sellars, W. (1974). Meaning as functional classification: A perspective on the relation of syntax to semantics. *Synthese*, *27*(3-4), 417–437. https://doi.org/10.1007/BF00484606

Singh, A., D'Arcy, M., Cohan, A., Downey, D., & Feldman, S. (2023, November). SciRepEval: A Multi-Format Benchmark for Scientific Document Representations. https://doi.org/10.48550/arXiv.2211.13308

Smith, K., & Kirby, S. (2012). Compositionality and Linguistic Evolution. In W. Hinzen, E. Machery, & M. Werning (Eds.), *Oxford Handbook of Compositionality* (pp. 493–509). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199541072.013.0024

Sowa, J. F. (1995). Syntax, semantics, and pragmatics of contexts. In G. Goos, J. Hartmanis, J. Leeuwen, J. G. Carbonell, J. Siekmann, G. Ellis, R. Levinson, W. Rich, & J. F. Sowa (Eds.), *Conceptual Structures: Applications, Implementation and Theory* (pp. 1–15, Vol. 954). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-60161-9_25

Spärck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, *28*(1), 11–21. https://doi.org/10.1108/eb026526

Steffensen, S. V., & Harvey, M. I. (2018). Ecological meaning, linguistic meaning, and interactivity. *Cognitive Semiotics*, *11*(1), 20180005. https://doi.org/10.1515/cogsem-2018-0005

Steinert-Threlkeld, S. (2020). Toward the Emergence of Nontrivial Compositionality. *Philosophy of Science*, *87*(5), 897–909. https://doi.org/10.1086/710628

Stevens, S. S. (1951). *Handbook of experimental psychology*. Wiley.

Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: Piecewise approximations to reality*. Harvard University Press.

Zubek, J., Korbak, T., & Rączaszek-Leonardi, J. (2024). Models of symbol emergence in communication: A conceptual review and a guide for avoiding local minima. *Artificial Intelligence Review*, *58*(2), 65. https://doi.org/10.1007/s10462-024-11048-y