Structure and Component: Factors Influencing the Compositionality Development of Pre-trained Language Models During Fine-tuning

Anonymous ACL submission

Abstract

In this work, we investigate factors influencing the compositionality development of pre-trained language models during fine-tuning from the perspective of two key concepts in linguistic compositionality: structure and component. Our key finding is that even fine-tuning of identical symbolic essences can lead to different compositionality developments of the same pre-trained language model, driven by the following factors: (1) structures uncommon in natural language lead to high compositionality, (2) components uncommon in natural language lead to high compositionality, and (3) component combinations uncommon in natural language lead to low compositionality. This phenomenon is intrinsically linked to linguistic compositionality, offering a new perspective for compositionality research on language models.

1 Introduction

Compositionality is a property that languages possess to some degree, expressed as "the meaning of a complex expression is determined by its structure and the meanings of its components" (Szabó, 2004; Pagin and Westerståhl, 2010). This concept extends from the expression-meaning mapping in language to the input-output mapping of any task. The compositionality of a language model on a task refers to the extent to which it leverages compositionality to process that task. High model compositionality aids generalization on unseen combinations, but may also introduce hallucinations in parts of the task that are inherently non-compositional.

Although there has been extensive research on the evaluation of the compositionality of language models in specific downstream tasks requiring fine-tuning (Hu et al., 2023; Xu and Wang, 2024), the factors stemming from pre-training on natural language that influence the compositionality development of pre-trained language models during fine-tuning remain an unexplored area. Similar to debates surrounding language learning capabilities (Kallini et al., 2024), a contentious issue is whether the compositionality development of pre-trained language models during fine-tuning is entirely symbolic (i.e., whether fine-tuning the same model with training data that provides linguistically distinct mappings of identical symbolic essence would lead the model to develop the same compositionality).

In this work, we investigate the two key concepts of linguistic compositionality—structure and component—to identify and explain factors influencing the compositionality development of language models during fine-tuning. We find that the compositionality development of pre-trained language models during fine-tuning is not entirely symbolic, and our investigation into structure and component reveals two distinct pathways that exert influence. We further discuss the possible causes of this phenomenon and its connection to linguistic compositionality.

2 Compositionality Quantification

We construct a conversion task where the input is natural language text generated based on a template (Lee et al., 2025) with 12 components:

`In the [place.ADJ] [place.N], I [action.VI] beside the [animal.N] and the [plant.N], while the [person.ADJ] [job.N] [action.ADV] [action.VT] the [object.ADJ] [texture.ADJ] [object.N].`

Each component enclosed by [] has 50 distinct candidate words (common and consistent with the component role; details are provided in the appendix A), with each candidate word corresponding to a unique 5-digit random number code. The task output is the sequence formed by concatenating the corresponding number codes for each component in order (separated by spaces).

By selecting one candidate word for each component, we can obtain 50^{12} distinct instances. How-

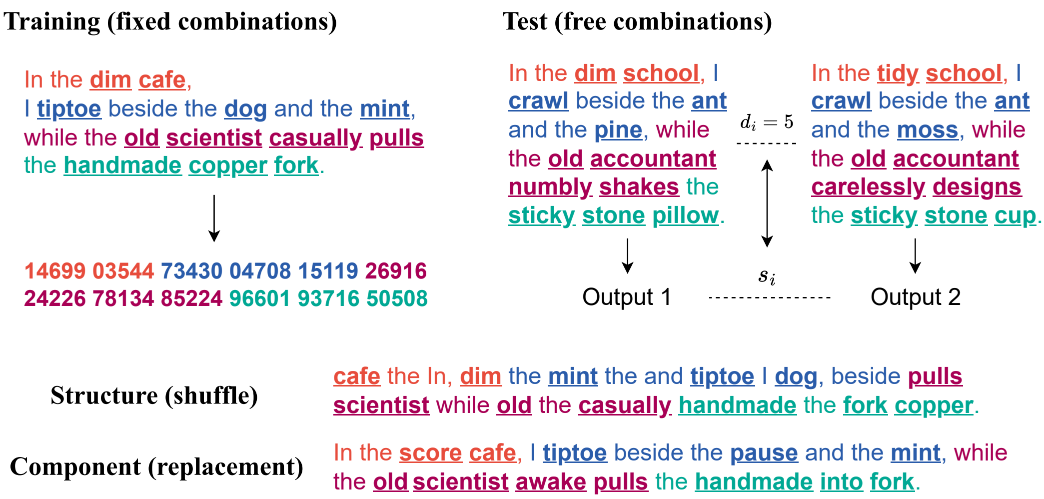

Figure 1: Illustration of compositionality quantification and how inputs change under two types of settings. Each **component** has 50 candidate words. In the training data, the combination of components within a semantic group (indicated by color) is fixed. In the test data, components can be freely combined.

ever, during training set construction, for each semantic group represented by a color in the template, we restrict the selection of components to 50 fixed combinations, and any candidate word appears in exactly one combination. In such a training set, any word always appears alongside fixed other components within the semantic group. Consequently, there is no evidence that a word independently determines an output component (i.e., the training set provides no evidence for compositionality). When fine-tuning a language model using this training set, the extent to which the model perceives a word as independently determining an output component indicates the degree of its spontaneous compositionality development. We randomly construct 20,000 distinct samples in this manner to form the training set for fine-tuning the pre-trained model.

Following Sevestre and Dupoux (2025), we employ the concept of Topographic Similarity to quantify the compositionality of language models. We first generate 2,000 text inputs based on the template, where candidate word selection no longer imposes combination constraints as in training set construction. For the i-th input x_i, we randomly select an input distance $d_i \in [1, 12]$, forming the input x_i^* by replacing the d_i components in x_i with other candidate words. We then measure the distance between the model's outputs for inputs x_i and x_i^* (i.e., the number of differing number codes) as s_i. Finally, we compute the Pearson correlation coefficient between $\{d_i\}$ and $\{s_i\}$ as a quantitative metric for compositionality. When the model exhibits maximum compositionality, each component in the input independently determines a corresponding number code in the output, resulting in a correlation coefficient of 1. A lower correlation coefficient indicates lower model compositionality.

3 Experiments

We conduct experiments on the following pre-trained language models: gemma-7b-it (Gemma-Team, 2024), Llama-3.1-8B-Instruct (Llama-Team, 2024), Qwen2.5-7B-Instruct (Qwen-Team, 2024), and Olmo-3-7B-Instruct (Olmo-Team, 2025). All results are averaged across three independent runs with different random seeds. Implementation details are provided in Appendix B.

3.1 Structure

Regarding structure, we seek to investigate whether the compositionality development of language models during fine-tuning changes when the symbolic essence of training data remains identical but inputs lose their typical natural language structure.

Following Kallini et al. (2024), we employ LocalShuffle to construct instances that lose natural language structure. Under the shuffle setting, for each original instance, we randomly shuffle all words within each semantic group, causing the input to lose its typical natural language structure. The shuffling pattern remains consistent across different instances within the same semantic group,

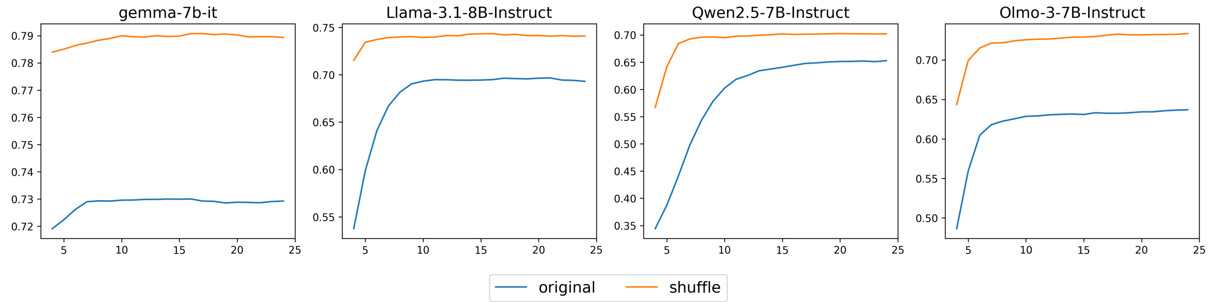

Figure 2: Curve showing the model compositionality over epochs under the original and shuffle settings.

and the order of the output number codes is adjusted to match the input component sequence after shuffling. This ensures that the symbolic essence of the training set remains identical between the original and shuffled settings. We observe whether there are differences in the compositionality development of models during fine-tuning under the original and shuffled settings.

Figure 2 presents the experimental results. We find that all pre-trained models exhibit higher compositionality under the shuffle setting compared to the original setting. This indicates that even when the training set provides mappings of identical symbolic essence, structural alterations still influence compositionality development during fine-tuning. When the structure deviates from the natural language structure common in pre-training, the model develops higher compositionality, exhibiting a greater degree of component independence.

3.2 Component

Regarding component, we seek to investigate whether the compositionality development of language models during fine-tuning changes when the symbolic essence of training data remains identical but inputs present uncommon combinations of components in natural language.

We first employ a random replacement setting to construct instances with uncommon component combinations. Under the random replacement setting, for each original instance, we randomly replace one word component in each semantic group with a word from a shared vocabulary of the four models (containing 13,740 words, none of which are candidate words when constructing the original instances; see Appendix C for details). Different instances share identical component selections within the same semantic group and employ the same replacement word for the same original word. This ensures that the symbolic essence of the training set remains identical between the original and random replacement settings. We observe whether there are differences in the compositionality development of models during fine-tuning under the two settings.

Figure 3 presents the experimental results (blue and orange lines). We observe that random replacement settings influence the compositionality development of models, but the direction varies across different models: compositionality increases on gemma-7b-it and Olmo-3-7B-Instruct, while compositionality decreases on Llama-3.1-8B-Instruct and Qwen2.5-7B-Instruct.

This result prompts us to consider whether other factors influence compositionality development in component replacement. Through observation of instances, we find that under the random replacement setting, there is a high probability of selecting uncommon words from the large vocabulary. Based on frequency and compositionality theory (Kirby, 2001; Sevestre and Dupoux, 2025), we hypothesize that the commonness of the replacement words themselves may also affect compositionality development. To verify this, we supplement our experiments with a common replacement setting. The common replacement setting follows the same construction as the random replacement setting, but the vocabulary used to select replacement words is a subset of the vocabulary in the random replacement setting, containing only 1,312 common words.

The green line in Figure 3 shows the results of the supplementary experiment. We find that all pre-trained models exhibit higher compositionality under the random replacement setting than under the common replacement setting, indicating that low word commonality leads to high compositionality. Comparison between the original setting and the common replacement setting eliminates the influence of word commonality. The results show that all pre-trained models exhibit lower or essentially unchanged compositionality the common re-

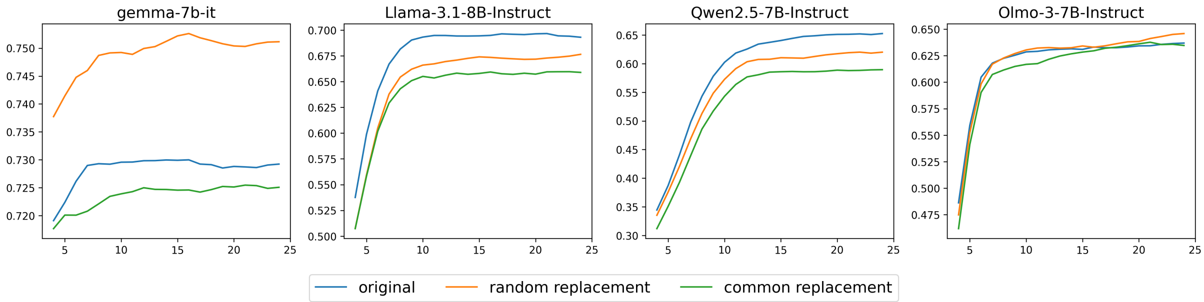

Figure 3: Curve showing the model compositionality over epochs under the original and replacement settings.

placement setting. This implies that even when the training set provides mappings of identical symbolic essence, changes in component still influence compositionality development during fine-tuning. The influencing factors include both the commonality of components and the commonality of component combinations. Uncommon components lead models to develop higher compositionality during fine-tuning, while uncommon component combinations result in lower compositionality.

4 Discussion of the Phenomenon

In the field of emergent communication, prior research has demonstrated that limited data exposure due to low commonality leads neural networks to develop higher compositionality (Sevestre and Dupoux, 2025). However, the differences in compositionality observed across various settings emerge during fine-tuning that is symbolically identical. The training data used for fine-tuning exhibits no differences in the degree of data exposure. Therefore, these compositionality differences cannot be attributed to data exposure differences during fine-tuning.

Past research indicates that pre-training on natural language leads to language models exhibiting poorer learning capabilities for languages that share the same symbolic essence as natural language but are linguistically impossible (Kallini et al., 2024). We therefore conjecture that the influence of structure and component factors on the compositionality development of language models during fine-tuning stems from characteristics inherent to language models that arise from pre-training on natural language text. This characteristic manifests during fine-tuning, preventing compositionality development from being entirely symbolic.

If this conjecture is accepted, the phenomenon can be further explained based on the theory of data exposure. For pre-trained language models, the concept of data exposure in the development of model compositionality may align with the concept of linguistic compositionality—that is, the exposed objects influencing model compositionality are the component and the structure. Under this explanation, less common components lead to higher compositionality because they are less frequently exposed during pre-training. Similarly, unnatural syntactic structures lead to higher compositionality due to less exposure of the structures during pre-training. For less common component combinations, language models may prefer to interpret them as special wholes, rather than understanding them through independent components based on linguistic compositionality. Consequently, less common component combinations lead to lower compositionality.

5 Conclusion

In this work, we explore the factors influencing the compositionality development of pre-trained language models during fine-tuning from two perspectives: structure and component. By constructing instances through local shuffling and component replacement, we discover that fine-tuning of identical symbolic essence can lead to different compositionality developments in the same pre-trained language model. Two pathways emerge regarding structure and component: (1) Structures uncommon in natural language lead to high compositionality; (2) Uncommon components in natural language lead to high compositionality, while uncommon component combinations lead to low compositionality. This indicates that the compositionality development of pre-trained language models during fine-tuning is not entirely symbolic and may be influenced by characteristics linked to linguistic compositionality, inherited from pre-training on natural language. Our work offers new insights into compositionality studies on language models.

Limitations

Limited experimental scenarios. To accurately quantify model compositionality and precisely construct mappings of identical symbolic essence for studying factors influencing the compositionality development of models during fine-tuning, our experiments are conducted on a fixed, artificially constructed transformation task where inputs are template-generated. Despite enriching the linguistic components within these templates as much as possible, such inputs cannot comprehensively cover all observable phenomena in natural language. Furthermore, model compositionality is easier to quantify under artificial tasks, whereas real-world tasks may present more complex scenarios where compositionality is difficult to measure. We will explore ways to broaden our experimental scenarios in future work.

The explanation for this phenomenon remains under discussion. Although compositionality can be quantified and the phenomenon can be clearly observed (i.e., fine-tuning with identical symbolic essence can lead to different compositional developments in the same model), explanations for this phenomenon are still being debated. In Section 4, our interpretive approach to the phenomenon involves excluding the influence of data exposure differences inherent in fine-tuning itself, attributing it to pre-training, and providing an explanation linked to linguistic compositionality. While this explanation is theoretically plausible, it may lack substantive proof. Further exploration may be required to explain this phenomenon. Nevertheless, we believe the discovery of this phenomenon itself constitutes a nontrivial contribution.

Related work. Due to space constraints, we have placed the related work section in Appendix D.

Ethics Statement

We comply with the license to use language models for scientific research purposes only. The datasets we use do not contain any information that names or uniquely identifies individual people or offensive content.

We use the translation tool DeepL for our writing, which may provide translations based on LLMs. Nevertheless, we meticulously review the translated content to avoid potential issues.

References

Shengnan An, Zeqi Lin, Qiang Fu, Bei Chen, Nanning Zheng, Jian-Guang Lou, and Dongmei Zhang. 2023. How do in-context examples affect compositional generalization? In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 11027–11052. Association for Computational Linguistics.

Dzmitry Bahdanau, Shikhar Murty, Michael Noukhovitch, Thien Huu Nguyen, Harm de Vries, and Aaron C. Courville. 2019. Systematic generalization: What is required and can it be learned? In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net.

Noam Chomsky. 1965. *Aspects of the Theory of Syntax*. 11. MIT press.

Eric Elmoznino, Thomas Jiralerspong, Yoshua Bengio, and Guillaume Lajoie. 2025. A complexity-based theory of compositionality. *Preprint*, arXiv:2410.14817.

Gemma-Team. 2024. Gemma: Open models based on gemini research and technology. *CoRR*, abs/2403.08295.

Chengang Hu, Xiao Liu, and Yansong Feng. 2023. DiNeR: A large realistic dataset for evaluating compositional generalization. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 14938–14947, Singapore. Association for Computational Linguistics.

Dieuwke Hupkes, Verna Dankers, Mathijs Mul, and Elia Bruni. 2020. Compositionality decomposed: How do neural networks generalise? *J. Artif. Intell. Res.*, 67:757–795.

Ahmad Jabbar, Cleo Condoravdi, and Christopher Potts. 2025. Distinguishing fair from unfair compositional generalization tasks. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 20796–20807, Suzhou, China. Association for Computational Linguistics.

Julie Kallini, Isabel Papadimitriou, Richard Futrell, Kyle Mahowald, and Christopher Potts. 2024. Mission: Impossible language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2024, Bangkok, Thailand, August 11-16, 2024*, pages 14691–14714. Association for Computational Linguistics.

Daniel Keysers, Nathanael Schärli, Nathan Scales, Hylke Buisman, Daniel Furrer, Sergii Kashubin, Nikola Momchev, Danila Sinopalnikov, Lukasz Stafiniak, Tibor Tihon, Dmitry Tsarkov, Xiao Wang, Marc van Zee, and Olivier Bousquet. 2020. Measuring compositional generalization: A comprehensive method on realistic data. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.

Najoung Kim and Tal Linzen. 2020. COGS: A compositional generalization challenge based on semantic interpretation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 9087–9105. Association for Computational Linguistics.

S. Kirby. 2001. Spontaneous evolution of linguistic structure-an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5(2):102–110.

Brenden M. Lake and Marco Baroni. 2018. Generalization without systematicity: On the compositional skills of sequence-to-sequence recurrent networks. In *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 2879–2888. PMLR.

Jin Hwa Lee, Thomas Jiralerspong, Lei Yu, Yoshua Bengio, and Emily Cheng. 2025. Geometric signatures of compositionality across a language model's lifetime. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2025, Vienna, Austria, July 27 - August 1, 2025*, pages 5292–5320. Association for Computational Linguistics.

Itay Levy, Ben Bogin, and Jonathan Berant. 2023. Diverse demonstrations improve in-context compositional generalization. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 1401–1422. Association for Computational Linguistics.

Llama-Team. 2024. The llama 3 herd of models. *CoRR*, abs/2407.21783.

Kate McCurdy, Paul Soulos, Paul Smolensky, Roland Fernandez, and Jianfeng Gao. 2024. Toward compositional behavior in neural models: A survey of current views. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, EMNLP 2024, Miami, FL, USA, November 12-16, 2024*, pages 9323–9339. Association for Computational Linguistics.

Olmo-Team. 2025. Olmo-3-7b-instruct.

Peter Pagin and Dag Westerståhl. 2010. Compositionality i: Definitions and variants. *Philosophy Compass*, 5(3):250–264.

Qwen-Team. 2024. Qwen2.5: A party of foundation models.

Jean-Baptiste Sevestre and Emmanuel Dupoux. 2025. Frequency & compositionality in emergent communication. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 27250–27262, Suzhou, China. Association for Computational Linguistics.

Peter Shaw, Ming-Wei Chang, Panupong Pasupat, and Kristina Toutanova. 2021. Compositional generalization and natural language variation: Can a semantic parsing approach handle both? In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 922–938. Association for Computational Linguistics.

Zoltán Gendler Szabó. 2004. Compositionality. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Thaddäus Wiedemer, Prasanna Mayilvahanan, Matthias Bethge, and Wieland Brendel. 2023. Compositional generalization from first principles. In *Advances in Neural Information Processing Systems*, volume 36, pages 6941–6960. Curran Associates, Inc.

Ziyao Xu and Houfeng Wang. 2024. SPOR: A comprehensive and practical evaluation method for compositional generalization in data-to-text generation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 604–621, Bangkok, Thailand. Association for Computational Linguistics.

Ziyao Xu and Houfeng Wang. 2025. Investigating the (de)composition capabilities of large language models in natural-to-formal language conversion. *CoRR*, abs/2501.14649.

A Candidate Words

Tables 1 and 2 show the candidate words for each component. For each component, we provide DeepSeek-V3.2 with a description of the component's function and request it to generate 50 commonly used candidate words. When generated candidate words have already served as candidates for other components, we require the model to regenerate new candidate words, thereby ensuring no overlap among candidate words for different components.

B Implementation Details

For training, we use a fixed learning rate of 1e-6 and a batch size of 1024. We train for 30 epochs, evaluating at the end of each epoch. Each point on the curve represents the average result over the most recent 5 epochs. Three runs of experiments take approximately 7 to 9 hours when using 8 A800 GPUs.

C Replacement Vocabulary

The replacement vocabulary is derived from the intersection of the vocabularies from the four evaluated pre-training models. We identify words within this vocabulary intersection that can function as independent new tokens following a blank space. These words are then used to construct a vocabulary for the random replacement setting, excluding any component candidate words used to build the original instances.

For the common replacement setting, we first require DeepSeek-V3.2 to provide a list containing 3,000 distinct common words. We then take the intersection of this list with the vocabulary used for the random replacement setting to obtain the vocabulary for the common replacement setting.

D Related Work

Quantification of model compositionality on tasks. For the model compositionality on tasks, the most common quantitative approach is to characterize it by the degree of manifestation of the ability to achieve compositional generalization, also known as the compositional generalization test (Bahdanau et al., 2019). Specifically, this test constructs a special training-test split for specific downstream tasks, ensuring that the components in the test set are rarely or never present in the training set. The model's performance on the test set is then evaluated after fine-tuning on the training set (Lake and Baroni, 2018; Kim and Linzen, 2020; Keysers et al., 2020). This test has also been further extended to in-context learning scenarios (An et al., 2023; Levy et al., 2023; Xu and Wang, 2025). Although the compositional generalization tests are now widely accepted as a measure of model compositionality, it requires models to adhere to a prescribed compositional structure that may not be sufficiently supported by limited fine-tuned data (Hupkes et al., 2020; Jabbar et al., 2025). The sufficiency of this support is difficult to define, leading to ambiguity in characterization (Wiedemer et al., 2023). Some studies have attempted to overcome this ambiguity through novel quantitative approaches, such as Complexity-based Theory (Elmoznino et al., 2025) or Topographic Similarity (Sevestre and Dupoux, 2025). Our work adopts the latter approach.

Linguistic compositionality. Humans possess the ability to understand and generate linguistic expressions based on linguistic compositionality (Chomsky, 1965). Through independent perception of linguistic components, we achieve comprehension and generation of arbitrary new combinations formed from known components. In the field of natural language processing, whether language models understand and generate based on linguistic compositionality has long been a contentious issue (McCurdy et al., 2024). For language models, directly quantifying their degree of leveraging linguistic compositionality is quite challenging. Although some work has attempted to explore compositional generalization of models under natural language variation (Shaw et al., 2021; Hu et al., 2023), the relationship between model compositionality on tasks and linguistic compositionality remains quite unclear. Our work explicitly establishes a connection between these two concepts for the first time.

Component	Candidate Words
place.ADJ	quiet, noisy, clean, dirty, bright, dim, modern, traditional, spacious, cramped, comfortable, plain, luxurious, shabby, safe, dangerous, peaceful, bustling, tidy, messy, warm, cool, humid, dry, ventilated, stuffy, fresh, polluted, barren, advanced, backward, lifeless, deserted, private, historic, newly-built, dilapidated, cozy, ornate, crowded, public, terrifying, dedicated, versatile, abandoned, vacant, heavily-guarded, fireproof, oppressive, relaxing
place.N	library, supermarket, school, hospital, park, restaurant, cinema, mall, bank, postoffice, airport, station, cafe, bar, gym, pool, museum, gallery, theater, stadium, zoo, arcade, bookstore, market, office, apartment, house, classroom, hotel, dormitory, garage, basement, attic, kitchen, bedroom, bathroom, church, temple, courthouse, prison, clinic, laboratory, factory, farm, garden, playground, parking, lobby, elevator, stairwell
action.VI	stand, sit, lie, sleep, run, walk, jump, laugh, cry, crawl, fly, rest, wait, think, dream, sing, dance, play, work, study, read, write, breathe, yawn, sneeze, cough, shiver, tremble, smile, shout, whisper, spin, roll, wriggle, complain, pray, meditate, slip, glide, stomp, bend, stretch, linger, hide, squat, kneel, tiptoe, blink, bow, wander
animal.N	dog, cat, cow, pig, chicken, duck, goose, horse, sheep, goat, rabbit, mouse, bear, lion, tiger, leopard, wolf, fox, elephant, giraffe, zebra, hippopotamus, rhinoceros, crocodile, shark, whale, dolphin, eagle, pigeon, sparrow, crow, peacock, parrot, snake, lizard, turtle, frog, butterfly, bee, ant, spider, mosquito, cockroach, cricket, kangaroo, panda, koala, penguin, monkey, gorilla
plant.N	apple, banana, grape, strawberry, watermelon, pineapple, mango, peach, lemon, pomelo, papaya, tomato, potato, carrot, cucumber, cabbage, spinach, onion, garlic, broccoli, eggplant, rose, sunflower, lily, orchid, jasmine, peony, chrysanthemum, tulip, daisy, lavender, cactus, oak, maple, willow, pine, palm, birch, cherry, ginkgo, moss, fern, grass, mint, aloe, ivy, lotus, wheat, rice, corn
person.ADJ	kind, intelligent, brave, friendly, diligent, humorous, honest, confident, optimistic, patient, lazy, selfish, rude, arrogant, timid, pessimistic, hypocritical, irritable, stupid, jealous, tall, short, fat, thin, young, old, outgoing, introverted, serious, lively, gentle, strong, fragile, enthusiastic, indifferent, careful, careless, cheerful, melancholy, generous, stingy, cautious, impulsive, reliable, unreliable, creative, conservative, open-minded, steady, easygoing

Table 1: The candidate words for the first six components.

Component	Candidate Words
job.N	doctor, teacher, engineer, lawyer, nurse, police-officer, firefighter, chef, driver, programmer, accountant, salesperson, manager, waiter, architect, artist, musician, writer, journalist, scientist, pilot, soldier, civil-servant, barber, farmer, worker, businessperson, consultant, designer, translator, actor, director, photographer, editor, researcher, dentist, veterinarian, pharmacist, psychologist, mechanic, electrician, plumber, courier, librarian, professor, judge, student, athlete, tailor, welder
action.ADV	happily, sadly, angrily, excitedly, patiently, hurriedly, easily, arduously, carefully, carelessly, bravely, fearfully, confidently, hesitantly, resolutely, gently, roughly, enthusiastically, coldly, sincerely, hypocritically, diligently, lazily, painfully, calmly, nervously, naturally, quickly, slowly, gracefully, clumsily, seriously, dazedly, stealthily, casually, actively, passively, wearily, bewilderedly, contentedly, composedly, anxiously, attentively, distractedly, smugly, desperately, numbly, unhurriedly, restlessly, blankly
action.VT	kicks, throws, licks, eats, scratches, bites, carries, hugs, rubs, hits, holds, squeezes, grabs, punches, pats, touches, pushes, pulls, lifts, moves, rotates, shakes, presses, cuts, destroys, repairs, cleans, washes, heats, observes, examines, analyzes, describes, finds, hides, shows, protects, uses, makes, designs, draws, measures, wraps, isolates, hangs, monitors, scans, kisses, curses, twists
object.ADJ	red, blue, green, yellow, black, white, purple, orange, brown, gray, pink, cyan, transparent, huge, tiny, round, square, triangular, oval, rectangular, fragrant, smelly, odorless, smooth, rough, soft, hard, sticky, new, opaque, worn-out, cheap, expensive, durable, premium, inferior, practical, flashy, classic, fashionable, outdated, user-friendly, eco-friendly, portable, custom-made, handmade, sturdy, lightweight, minimalist, exquisite
texture.ADJ	plastic, wooden, metal, glass, ceramic, leather, rubber, paper, stone, iron, steel, aluminum, copper, gold, silver, cotton, silk, wool, bamboo, titanium, brass, bronze, marble, granite, concrete, sponge, wax, nylon, polyester, silicone, linen, canvas, velvet, carbon-fiber, acrylic, resin, zinc, enamel, memory-foam, rattan, cork, chrome-plated, corduroy, beeswax, nephrite, graphite, mica, pearl, zirconia, silicon-carbide
object.N	necklace, bottle, table, chair, book, pen, mobile-phone, computer, cup, bowl, spoon, fork, knife, lamp, clock, watch, shoes, clothing, pants, hat, bag, wallet, key, glasses, headphones, television, refrigerator, air-conditioner, fan, mirror, toothbrush, towel, pillow, quilt, sofa, bookshelf, vase, plate, pot, pan, kettle, oven, broom, keyboard, hammer, hanger, suitcase, umbrella, basket, scissors

Table 2: The candidate words for the last six components.