

# LEXICAL CREATIVITY OF LARGE LANGUAGE MODELS: EVIDENCE FROM SUFFIXATION

**Alina REȘCEANU**

**University of Craiova (Romania)**

**Anca DINU**

**University of Bucharest (Romania)**

**Andra-Maria FLORESCU**

**University of Bucharest (Romania)**

<https://doi.org/10.52846/5tnybg68>

## **Abstract**

In this study, we examine lexical creativity in Large Language Models (LLMs) through a focused analysis of suffixation, a highly productive and well-described word-formation process in English. Using a selection of suffixation outputs generated by 24 LLMs (Dinu et al. 2025a), the study operationalises lexical creativity through a transparent, theory-driven scoring framework based on morphological well-formedness, productivity and semantic interpretability. Rather than treating creativity as a psychological construct or attempting to measure it directly, our analysis follows a descriptive and exploratory approach, relying on linguistically motivated composite indices. The results show that LLMs demonstrate strong morphological competence in suffixation, while creative behaviour varies across the two suffixes (*-ish* vs. *-able*) and across the tested LLMs. Creativity emerges most clearly at the boundaries of productivity, where increased novelty often goes hand in hand with overgeneralisation. The findings suggest that lexical creativity in LLMs involves exploring the limits of existing word-formation patterns, rather than creating words without constraints, and highlights how artificial language systems manage the balance between structure and innovation.

**Keywords:** *lexical creativity, suffixation, word formation, large language models (LLMs), morphological productivity.*

## **Résumé**

Dans cette étude, nous examinons la créativité lexicale des grands modèles de langage (Large Language Models, LLMs) à travers une analyse ciblée de la suffixation, un procédé de formation des mots en anglais à la fois très productif et largement décrit. À partir d'un ensemble de productions suffixées générées par 24 LLMs (Dinu et al. 2025a), l'étude opérationnalise la créativité lexicale au moyen

d'un cadre d'évaluation transparent et théoriquement fondé, reposant sur la bonne formation morphologique, la productivité et l'interprétabilité sémantique. Plutôt que de considérer la créativité comme un construit psychologique ou de chercher à la mesurer directement, notre analyse adopte une approche à la fois descriptive et exploratoire, fondée sur des indices composites linguistiquement motivés. Les résultats montrent que les LLMs font preuve d'une forte compétence morphologique en suffixation, tandis que le comportement créatif varie selon le type de suffixe (-ish vs -able) et selon les modèles. La créativité apparaît plus nettement aux frontières de la productivité, où l'augmentation de la nouveauté s'accompagne souvent de phénomènes de surgénéralisation. Les résultats suggèrent que la créativité lexicale des LLMs consiste plutôt à explorer les limites des schémas de formation des mots existants qu'à produire des formes dépourvues de contraintes, et ils mettent en lumière la manière dont les systèmes linguistiques artificiels gèrent l'équilibre entre structure et innovation.

**Mots-clés:** *créativité lexicale, suffixation, formation des mots, grands modèles de langage (LLMs), productivité morphologique.*

## **1. Introduction**

Large Language Models (LLMs) have demonstrated a growing capacity to generate new lexical items, which increased the researchers' interest in the question of the linguistic creativity in artificial language generation. In our recent work, we have explored this capacity empirically, comparing LLM outputs to human performance across a range of creative language tasks, including word formation and figurative language use (Dinu and Florescu 2024a, Dinu and Florescu 2024b, Dinu and Florescu 2025, Dinu et al. 2025a). At the same time, more general discussions in computational creativity emphasize the need for transparent, theoretically grounded analyses (Boden 2004, 2009; Colton and Wiggins 2012; Chakrabarty et al. 2023a; Franceschelli and Musolesi 2024). Within the field of linguistics, creativity is increasingly viewed not as unrestricted novelty, but as patterned innovation constrained by linguistic structure (Bergs 2019; Sampson 2016; Jones 2015). Despite this growing body of work, relatively little attention has been paid to the morphological mechanisms underlying LLM-generated lexical novelty.

Morphology offers a particularly interesting perspective for examining lexical creativity, as word formation combines rule-governed structure with opportunities for innovation. Among word-formation processes, suffixation occupies a central role in English morphology due to its high productivity and well-established formal and semantic constraints (Bauer 1983; Kortvélyessy et al. 2021, 2022). Recent computational studies have shown that LLMs display non-trivial morphological

generalisation abilities, including sensitivity to cross-linguistic morphological patterns and nonce-word formation (Anh et al. 2024; Ismayilzada et al. 2025). At the same time, evidence suggests that such generalisation is uneven and prone to overextension (Anh et al. 2024), making suffixation an ideal domain for investigating how productivity and constraint interact in LLM-generated lexical items.

In this context, the aim of our study is to explore lexical creativity in LLM-generated suffixation by operationalising creativity through a set of theoretically motivated linguistic criteria. Following common practice in linguistic and computational creativity research, creativity is not treated as a psychological construct, nor is it directly measured, but rather operationalised as observable linguistic behaviour (Boden 2004; Gaut 2003; Jones 2015). That is one of the motivations on which we build the rationale of our approach. The second motivation results from the observations we made in some of our earlier work (Dinu et al. 2024; 2025) on linguistic creativity in LLMs, where analyses based on computational analysis tools had to be complemented by manual inspection of model outputs in order to capture finer-grained creative patterns.

While such qualitative examination led to valuable empirical insights, it also highlighted the need for a more systematic and linguistically grounded approach to analysing machine-generated creativity.

Therefore, the framework proposed here responds to this need by offering a transparent model for the structured analysis of lexical creativity, with a particular focus on word-formation processes. In other words, our model uses a small set of already established linguistic criteria (i.e. whether the forms created are morphologically well formed, how productively the suffixes are used and whether the resulting words are semantically interpretable) and combines them into simple scores that help us compare suffixation outputs across the tested models and tasks. Because there is no standard numerical measure of lexical creativity in word formation, this approach allows us to analyse creative patterns systematically without treating creativity as a single number.

From the methodological standpoint, our paper adopts a descriptive and exploratory approach, reflecting both the challenges of studying lexical creativity in LLMs and the linguistic focus of the analysis. Rather than evaluating models in machine learning terms or applying complex statistical or psychological tests of creativity (Guilford 1950, 1967; Kaufman & Sternberg 2010), we rely on simple descriptive statistics based on scored items which are linguistically motivated. This makes our analysis transparent and interpretable, and also allows us to focus on how suffixation is used creatively within clear morphological constraints.

Moreover, by concentrating on a single word-formation process, we want to shed some light on the patterns of lexical creativity in LLMs while remaining cautious about making broader generalisations. Therefore, we have to point out that our intention is not to “measure creativity” as a cognitive trait, but to operationalise somehow LLMs’ creativity in a more replicable way.

## **2. Research background**

Recent work has played a key role in shaping current approaches to the study of linguistic creativity in LLMs, particularly through systematic comparison with human language use. For example, Dinu & Florescu (2024a&b) introduced controlled benchmarks aimed at testing verbal creativity across a variety of tasks, including word formation, figurative language, and semantic flexibility and showed that LLMs are capable of producing outputs that are both novel and linguistically plausible, while also revealing differences in how creative behaviour manifests across linguistic domains. Importantly, creativity was treated not as a single, unified ability, but as a set of observable linguistic behaviours that can be elicited and analysed under carefully controlled conditions.

Subsequent work focused on the design and structure of creativity assessment tasks, with a strong emphasis on methodological clarity and comparability between human and model participants. Dinu et al. (2025a) argue that meaningful assessment of linguistic creativity requires tasks that are sufficiently open to allow creative responses, yet constrained enough to support linguistic analysis. In this context, word formation is highlighted as a particularly suitable domain, as it enables the study of creative potential without conflating creativity with general fluency or stylistic variation. This focus on task design underscores the importance of examining creativity at the level of specific linguistic processes rather than relying on broad or undifferentiated measures.

More recent comparative studies further refine this perspective by combining qualitative linguistic analysis with descriptive quantitative summaries. Dinu et al. (2025a) propose a framework for assessing linguistic creativity that prioritises transparency and theoretical grounding, explicitly avoiding claims about psychological creativity or cognitive equivalence between humans and language models. Similarly, Dinu and Florescu (2025) show that although LLMs often achieve high levels of surface novelty, their creative outputs remain closely shaped by structural and distributional constraints. Taken together, this body of work provides a clear motivation for our study’s focus on suffixation as a well-defined site of lexical creativity, extending existing benchmarks with a more fine-grained, morphology-oriented analysis that remains consistent with the methodological principles established in earlier research.

### **3. Methodology**

#### ***3.1. Research design***

As mentioned in the introductory section, we adopt here a descriptive and exploratory research design to investigate lexical creativity in LLMs, by focusing specifically on suffixation. Our choice reflects both the novelty of the research area and the need for standard quantitative metrics for lexical creativity in word formation. Therefore, based on such a methodological approach, we emphasize linguistic interpretability, transparency and theoretical grounding rather than predictive modelling or statistical inference. The analysis is not conceived as a machine learning evaluation or a psychological test of creativity, but as a linguistic examination of how suffixal constructions are used to produce new lexical items.

#### ***3.2. Data and task design***

The data used in this study consist of LLM-generated lexical items<sup>1</sup> produced in response to controlled suffixation tasks, especially the responses to the test described and applied in Dinu et al. (2025 a & b). The test was applied to both humans and LLMs, but, for the purpose of this analysis, we selected only the responses of the LLMs included in the survey. The task was designed to elicit the productive use of English suffixes by requiring the model to generate new words formed through suffixation with *-ish* and *-able*. Each generated lexical item constitutes a single unit of analysis. In the test, the requirements explicitly encouraged originality while implicitly constraining the word-formation process to suffixation, thereby balancing creative freedom with formal restriction.

#### ***3.3. Operationalisation of lexical creativity: description of the model and the scoring procedure***

Lexical creativity is operationalised through a set of linguistically motivated evaluation criteria. Following common practice in linguistic and computational creativity research, creativity is understood as constrained innovation, emerging from the interaction between productivity and rule-governed structure.

Our model examines and evaluates each lexical item along two broad dimensions:

---

<sup>1</sup> The dataset comprises outputs generated by 24 large language models, for a total of 2,304 responses in all tasks. Of these, 20 models were run using default parameter settings, while 4 models were tested using maximised generation parameters (e.g. temperature and top-p). These parameters are commonly associated in the literature with increased output variability and are sometimes informally referred to as “creativity parameters” (see details in Dinu et al. 2025a).

1. Morphological competence, reflecting the LLM's ability to apply suffixes in accordance with formal and semantic constraints.

2. Creativity, reflecting the degree of novelty and expressivity achieved within those constraints.

These two dimensions were further specified through six criteria, each scored on a 0-2 scale (0 = not met, 1 = partially met, 2 = fully met), which are basically described as follows:

1. *Morphological well-formedness*: refers to the formal correctness of the derived form, assessing whether the word conforms to basic morphological requirements, such as recognisable base and suffix segmentation, correct affix realisation and the absence of formal distortion or malformed affixation.

2. *Constructional compatibility*: refers to the degree to which a derived form conforms to the formal and semantic requirements of the suffixation pattern involved. This criterion captures whether the suffix is combined with an appropriate base, whether the resulting form aligns with the expected grammatical category, and whether the base-suffix pairing is semantically coherent within the relevant constructional schema.

3. *Productivity*: refers to the model's sensitivity to the productive scope and constraints of a suffix. It evaluates whether productive patterns are extended in a controlled manner, or whether creative extension results in overgeneralisation, semantic mismatch or formal distortion (as was the case in some responses).

4. *Novelty*: refers to the degree to which a derived form departs from attested or highly conventional lexical items. This criterion assesses whether the output introduces a genuinely new combination or extension, rather than reproducing an established or easily predictable form.

5. *Originality*: refers to the extent to which a derived form reflects an unexpected use of a suffixation pattern (*-able* and *-ish* in our case). While novelty captures newness, originality captures the degree of creative insight involved in selecting and combining morphological elements.

6. *Expressiveness*: looks at the extent to which a derived form contributes an expressive or stylistic dimension that enhances its communicative impact beyond its compositional semantics.

The first three criteria were aggregated into a Morphological Competence Index (MCI), while the latter three were aggregated into a Creativity Index (CI). The reason why we use these two composite indices is to offer some analytical clarity to our analysis without implying absolute measurement of creativity.

For each suffixation task, mean MCI and CI scores can in principle be computed by averaging the indices across all generated items within that task. In this study, however, scores are calculated for a sample of responses, with the aim of

illustrating how the proposed evaluation model can be applied in practice. This illustrative use of the scoring framework does not limit its scope: while the empirical analysis is necessarily restricted, the evaluation model itself is designed to be applicable to suffixation tasks in general and can be extended to larger datasets or to additional word-formation processes. We intend to do that in our future work and also to use it to compare responses from humans and LLMs.

#### 4. Examples of scored responses

In this section, we selected some LLM responses, for which we exemplify the scoring procedure in terms of the two composite indices introduced above: the Morphological Competence Index (MCI) and the Creativity Index (CI), briefly justifying the score. We also grouped them according to the relative balance between MCI and CI, rather than according to absolute scores alone. The purpose of this classification is illustrative, since it shows how different combinations of structural control and creative extension manifest in suffixation outputs.

##### 4.1. High MCI, moderate CI (controlled productivity)

Example 1: *gleamish* (-ish)

Criterion	Score	Justification
Morphological		
well-formedness	2	clear base and suffix
Constructional		
compatibility	2	canonical adjectival <i>-ish</i>
Productivity	2	productive extension
MCI	6	
Novelty	1	slight lexical novelty
Originality	1	predictable extension
Expressiveness	1	mild evaluative tone
CI	3	

Example 2: *chirpish* (-ish)

Criterion	Score	Justification
Morphological		
well-formedness	2	correct suffixation
Constructional		
compatibility	2	typical <i>-ish</i> semantics
Productivity	2	acceptable anthropomorphic use
MCI	6	
Novelty	1	limited surprise

Originality	1	conventional extension
Expressiveness	1	weak expressivity
CI	3	

Example 3: *quirkish* (-ish)

Criterion	Score	Justification
Morphological		
well-formedness	2	similar to attested forms
Constructional		
compatibility	2	fully canonical
Productivity	2	productive extension
MCI	6	
Novelty	0	near-attested
Originality	0	highly predictable
Expressiveness	2	strong evaluative meaning
CI	2	

Examples such as *gleamish*, *chirpish*, and *quirkish* are grouped together because they exhibit maximal morphological competence (MCI = 6), while their creativity scores remain moderate or low (CI = 2–3). These forms are fully well formed, constructionally compatible and reflect appropriate use of productive suffixation patterns. However, they involve limited lexical surprise, as the base-suffix combinations are either attested or highly predictable. This profile corresponds to controlled productivity, where innovation remains safely within well-established morphological schemas.

#### 4.2. High MCI, higher CI (constrained creativity)

Example 4: *sniffish* (-ish)

Criterion	Score	Justification
Morphological		
well-formedness	2	correct suffixation
Constructional		
compatibility	2	acceptable metaphorical use
Productivity	2	productive extension
MCI	6	
Novelty	1	slightly unexpected
Originality	1	moderate semantic shift
Expressiveness	1	subtle imagery
CI	3	

Example 5: *grumbleable* (-able)

Criterion	Score	Justification
Morphological		
well-formedness	2	base + suffix identifiable
Constructional		
compatibility	1	verb marginal for -able
Productivity	1	creative overextension
MCI	4	
Novelty	2	non-attested
Originality	1	pattern-based creativity
Expressiveness	1	slightly humorous
CI	4	

Example 6: *dreamable* (-able)

Criterion	Score	Justification
Morphological		
well-formedness	2	correct suffixation
Constructional		
compatibility	1	abstract base
Productivity	2	still interpretable
MCI	5	
Novelty	1	limited novelty (attested)
Originality	1	safe extension
Expressiveness	1	neutral
CI	3	

We included in this category forms that maintain strong morphological control but exhibit greater creative extension. For instance, *sniffish* preserves full morphological well-formedness but introduces a mild semantic shift, while *grumbleable* and *dreamable* extend the productive scope of *-able* beyond prototypical cases. These examples are characterised by relatively high MCI values (4–6) combined with higher CI scores, reflecting creativity that stretches constraints without fully violating them. They illustrate how creativity can emerge within a still recognisable constructional framework.

**4.3. Lower MCI, higher CI**Example 7: *wobbleble* (-able)

Criterion	Score	Justification
-----------	-------	---------------

Morphological		
well-formedness	0	suffix distortion/truncation
Constructional		
compatibility	0	violation of form
Productivity	1	creative intent visible
MCI	1	
Novelty	2	highly unexpected
Originality	1	non-canonical (error pattern)
Expressiveness	0	poor interpretability
CI	3	

Example 8: *flimsyble* (-able)

Criterion	Score	Justification
Morphological		
well-formedness	0	suffix distortion/truncation
Constructional		
compatibility	0	constraint violation
Productivity	1	attempted derivation
MCI	1	
Novelty	2	strong deviation
Originality	1	non-canonical (error pattern)
Expressiveness	0	low clarity
CI	3	

The reason we grouped forms such as *wobbleble* and *flimsyble* separately is because they exhibit a clear imbalance between creativity and morphological control. Their low MCI scores reflect formal distortion and violation of suffixation schemas, while their higher CI scores result from strong novelty and deviation from conventional patterns. These examples illustrate cases where creative extension exceeds the limits of productive morphology, leading to overgeneralisation rather than controlled innovation.

#### **4.4. High creativity within constraint**

Example 9: *pawprintable* (-able)

Criterion	Score	Justification
Morphological		
well-formedness	2	multi-morphemic base intact
Constructional		
compatibility	1	noun-based derivation marginal

Productivity	1	creative but strained
MCI	4	
Novelty	2	unexpected base
Originality	2	non-obvious choice
Expressiveness	1	evocative imagery
CI	5	

Example 10: *snuggable* (-able)

Criterion	Score	Justification
Morphological		
well-formedness	2	canonical
Constructional		
compatibility	2	prototypical -able use
Productivity	2	highly productive
MCI	6	
Novelty	0	attested
Originality	0	conventional
Expressiveness	2	strong affective meaning
CI	2	

Examples like *pawprintable* are somehow on an intermediate position. They are not fully canonical (MCI = 4), but they remain interpretable and structurally recognisable, and achieve high creativity scores due to unexpected but meaningful base–suffix combinations. This category highlights instances where creativity is maximised without complete loss of constructional integrity.

**4.5. Borderline, yet interpretable**

Example 11: *jitterish* (-ish)

Criterion	Score	Justification
Morphological		
well-formedness	2	correct suffixation
Constructional		
compatibility	2	clear adjectival meaning
Productivity	2	productive
MCI	6	
Novelty	1	mildly new
Originality	1	limited surprise
Expressiveness	1	moderate affective meaning
CI	3	

Example 12: *cloudable* (-able)

Criterion	Score	Justification
Morphological		
well-formedness	2	correct suffixation
Constructional		
compatibility	1	noun-based derivation
Productivity	1	creative extension
MCI	4	
Novelty	1	some novelty
Originality	1	expected extension
Expressiveness	1	neutral
CI	3	

Finally, the examples above (*jitterish* and *cloudable*) are grouped as borderline cases because they are morphologically well formed and interpretable, but only mildly creative. Their scores reflect neither clear overgeneralisation, nor high creativity.

#### 4.6. Discussion of results

The results illustrated above of the suffixation-based analysis provide evidence that lexical creativity in LLMs emerges as a form of constrained innovation, rather than unrestricted novelty. Across the analysed responses, suffixation outputs exhibit rather high Morphological Competence Index (MCI) scores, while Creativity Index (CI) scores display greater variability. This dissociation aligns closely with theoretical accounts of word formation in human language, in which productivity is understood as rule-governed yet gradient (Bauer 1983).

From the perspective of classical morphological theory, Bauer (1983) emphasises that productive affixation is not merely a matter of formal attachment, but also of semantic compatibility and conventional usage. The present findings reflect this distinction. Outputs such as *gleamish*, *chirpish* and *jitterish* demonstrate high MCI scores due to their conformity to established *-ish* constructional schemas, but achieve only moderate CI values because they remain predictable. These forms illustrate what Bauer characterises as regular productivity, where innovation operates within well-established morphological patterns.

At the same time, several examples reveal the limits of productive extension. Forms such as *grumbleable* and *pawprintable* illustrate creative attempts to extend *-able* beyond its prototypical verb-based domain. While these items remain interpretable, their lower MCI scores reflect marginal constructional compatibility and overgeneralisation. This behaviour resonates with the framework proposed by

Kortvélyessy, Štekauer, and Kačmár (2021, 2022), who argue that creativity in word formation arises from tension between creative potential (the availability of productive patterns) and creative performance (their contextually appropriate realisation). In this light, LLM outputs exhibiting moderate MCI but high CI values can be interpreted as instances of creative performance that stretch, but do not entirely abandon, morphological constraints.

More interesting cases are *wobbleble* and *flimsyble*, which score high on novelty but low on morphological competence, reflecting formal degradation rather than productive creativity. From a theoretical standpoint, such outputs resemble what Kortvélyessy et al. (2021) describe as constraint violation, where innovation exceeds the limits of the morphological system. We have to point out that these cases are relatively infrequent, suggesting that LLMs generally maintain constructional integrity even when engaging in creative extension.

The observed patterns also align with recent computational studies probing morphological generalisation in LLMs. Anh et al. (2024), using a Wug-test paradigm across languages, demonstrate that LLMs possess non-trivial morphological generalisation abilities, while also exhibiting systematic weaknesses in constraint-sensitive application. Our findings extend this insight to the domain of lexical creativity, showing that while LLMs can generalise suffixation patterns effectively, creative extension often increases the risk of overgeneralisation.

All in all, our results support the idea that LLM’s lexical creativity is fundamentally construction-based and rule-governed, and does not manifest as free-form invention. High creativity scores are most often achieved when LLMs push against semantic or distributional constraints, rather than when they abandon them altogether.

Finally, the contrast between the analysed suffixes further emphasize the role of morphological structure in shaping creative outcomes. The consistently higher MCI scores observed for *-ish* compared to *-able* reflect differences in semantic flexibility and productivity documented in the morphological literature (Bauer 1983). Conversely, the higher CI values associated with *-able* indicate that creativity is more likely to emerge in domains where constraints are tighter and therefore more salient. This asymmetry provides empirical support for the claim that lexical creativity is most visible at the edges of productivity, where rule-governed behaviour encounters resistance.

### 5. Limitations

There are several limitations of our analysis. First, the analysis focuses exclusively on suffixation as a word-formation process. While this choice enables theoretical evaluation of morphological productivity and constraint, it necessarily limits the generalisability of the findings to other word-formation

mechanisms such as compounding or blending. We intend in our future work to examine whether similar patterns of constrained creativity appear across other morphological processes.

Second, lexical creativity was operationalised through a set of linguistically motivated evaluation criteria combined into composite indices. Although these criteria draw on well-established notions in morphological theory, such as well-formedness, productivity and semantic interpretability, no standardised quantitative metric for lexical creativity currently exists. Consequently, the proposed indices should be understood as analytical tools for comparison rather than as absolute measures of creativity.

Finally, the scoring procedure relies on our linguistic judgment, which inevitably introduces a degree of subjectivity, particularly in the evaluation of novelty and expressiveness. We tried to somehow reduce this subjectivity by relying on clearly defined scoring guidelines rather than individual scores taken in isolation. Future work could build on this approach by involving larger groups of evaluators/annotators or by combining expert evaluation with corpus-based evidence of how newly formed words are taken up in actual language use, especially when considering human responses to such tasks.

## 6. Conclusions

This study examined lexical creativity in LLMs through a focused analysis of suffixation, one of the most productive and theoretically well-understood word-formation processes in English. By operationalising creativity using linguistically motivated criteria and applying this framework to a selection of suffixation outputs from 24 LLMs, the analysis shows a clear pattern: while LLMs are generally very good at producing well-formed suffixal words, their creative behaviour is more uneven. In other words, morphological competence appears to be stable across models, whereas creativity emerges more selectively and under specific conditions.

A closer examination of the data shows that creativity tends to surface where morphological constraints become more visible. Suffixes such as *-ish*, which are semantically flexible and widely productive, mostly lead to safe and predictable extensions. By contrast, suffixes like *-able* encourage more experimental uses, resulting in higher novelty but also in a greater number of marginal or ill-formed outputs.

More generally, these findings contribute to ongoing discussions about linguistic creativity of LLMs by showing that lexical creativity is best understood as exploration within constraints rather than as free invention. When LLMs generate novel words, they do not abandon morphological structure; instead, they explore the limits of existing word-formation patterns. Although our study focuses only on suffixation, it demonstrates how careful, linguistically informed analysis can clarify

both the possibilities and the boundaries of creativity in artificial language systems, and it provides a clear starting point for future work on other word-formation processes and languages.

## **BIBLIOGRAPHY**

1. Anh, Duc; Raviv, Limor; Galke, Lukas, “Morphology matters: Probing the cross-linguistic morphological generalization abilities of large language models through a wug test”, in Kuribayashi, Taro; Rambelli, Giulia; Takmaz, Ece; Wicke, Philipp; Oseki, Yohei (coords), *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, Association for Computational Linguistics, Bangkok, 2024, pp. 177–188. DOI: 10.18653/v1/2024.cmcl-1.15
2. Bauer, Laurie, *English Word-Formation*, Cambridge Textbooks in Linguistics, Cambridge University Press, Cambridge, 1983.
3. Bergs, Alexander, “What, if anything, is linguistic creativity?”, *Gestalt Theory*, vol. 41, 2019, pp. 173–183, doi:10.2478/gth-2019-0017.
4. Boden, Margaret A., *The Creative Mind: Myths and Mechanisms*, Routledge, London, 2004.
5. Boden, Margaret A., “Computer models of creativity”, *AI Magazine*, vol. 30, no. 3, 2009, pp. 23–34, doi:10.1609/aimag.v30i3.2254.
6. Chakrabarty, Tuhin; Laban, Philippe; Agarwal, Divyansh; Muresan, Smaranda; Wu, Chien-Sheng, “Art or artifice? Large language models and the false promise of creativity”, in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, ACM, Hamburg, 2023.
7. Colton, Simon; Wiggins, Geraint, “Computational creativity: The final frontier?”, in *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI 2012)*, IOS Press, Amsterdam, 2012, pp. 21–26, doi:10.3233/978-1-61499-098-7-21.
8. Dinu, Anca; Florescu, Andra-Maria; Reșceanu, Alina, “A comparative approach to assessing linguistic creativity of Large Language Models and Humans”, *Procedia Computer Science*, vol. 270, 2025, pp. 1292–1301 (Dinu et al. 2025a).
9. Dinu, Anca; Florescu, Andra-Maria; Reșceanu, Alina, “Test Design and Structure for Assessing Humans and LLMs’ Linguistic Creativity”, *Analele Universității din Craiova, Seria Științe Filologice, Lingvistică*, Anul XLVII, nr. 1-2, 2025, *Editura Universitaria*, Craiova, 2025, pp. 62-74, <https://doi.org/10.52846/aucssflingv.v47i1-2.178> (Dinu et al. 2025b).
10. Dinu, Anca; Florescu, Andra-Maria, “Testing language creativity of large language models and humans”, in Hamäläinen, Mika; Öhman, Emily; Bizzoni, Yuri; Miyagawa, So; Alnajjar, Khalid (coords), *Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities*, Association

for *Computational Linguistics*, Albuquerque, 2025, pp. 426–436. <https://aclanthology.org/2025.nlp4dh-1.37.pdf>, DOI: 10.18653/v1/2025.nlp4dh-1.37

11. Dinu, Anca; Florescu, Andra-Maria, “An integrated benchmark for verbal creativity testing of LLMs and humans”, *Procedia Computer Science*, vol. 246, 2024, pp. 2902–2911.

<https://www.sciencedirect.com/science/article/pii/S1877050924024141>,  
<https://doi.org/10.1016/j.procs.2024.09.380>. (Dinu and Florescu 2024a)

12. Dinu, Anca & Florescu, Andra-Maria, “Comparing Large Language Models’ Verbal Creativity to Human Verbal Creativity”, in F. Dell’Orletta, A. Lenci, S. Montemagni & R. Sprugnoli (eds.), *Proceedings of the Tenth Italian Conference on Computational Linguistics (CLiC-it 2024)*, Pisa, CEUR-WS.org, 2024. (Dinu and Florescu 2024b)

13. Franceschelli, Giuseppe; Musolesi, Mirco, „On the creativity of large language models”, *AI & Society*, 2024, DOI:10.1007/s00146-024-02127-3.

14. Gaut, Berys, „Creativity and imagination”, in Kaufman, James; Sternberg, Robert (coords), *The Cambridge Handbook of Creativity*, Cambridge University Press, Cambridge, 2010, pp. 148–173.

15. Jones, Rodney H. (ed.), *The Routledge Handbook of Language and Creativity*, Routledge, London, 2015.