

# Creative Transformation in Literary Texts: Modelling Change Across Representational Levels

Ioana-Roxana Boriceanu<sup>1</sup>, Liviu P. Dinu<sup>1,2</sup>

<sup>1</sup>Faculty of Mathematics and Computer Science

<sup>2</sup>Human Language Technologies Research Center  
University of Bucharest, Romania

ioana-roxana.boriceanu@fmi.unibuc.ro, ldinu@fmi.unibuc.ro

## Abstract

Creativity is often framed as the production of novelty, yet many cultural works emerge through transformation of earlier artifacts and not through isolated invention. Drawing on theories of imitation by Gabriel Tarde and James Mark Baldwin, this paper models creativity as selective transformation across multiple levels of textual representation. We introduce a multi-level framework that compares literary texts across lexical, semantic, conceptual, structural, and narrative dimensions using directional alignment and control calibrated similarity measures. Applying the model to historically documented literary relationships, we show that different pairs preserve source structure at different representational levels while diverging in others. These transformation profiles provide a quantitative method for characterizing how imitation persists and where creative divergence occurs within literary works.

## Introduction

Creativity has long been a central concern in psychology, philosophy, and the arts, and has gained renewed urgency in the context of artificial intelligence, where questions of human distinctiveness and creative capacity are increasingly foregrounded (Boden 2016; Marcus and Davis 2019). A widely accepted formulation is the standard definition proposed in (Runco and Jaeger 2012), which holds that creativity requires both originality and effectiveness. Novelty alone is insufficient, as arbitrary variation may be original without being meaningful, while usefulness without originality reduces to routine competence.

Margaret Boden further clarified the relation between novelty and value in *The Creative Mind* (Boden 2004), distinguishing combinational, exploratory, and transformational creativity. In her account, creative production operates within structured spaces of possibility, and novelty emerges through transformations of those spaces instead of unconstrained deviation. Building on this perspective, computational creativity research has proposed formal frameworks for evaluating artificial artifacts (Maher 2010; Ritchie 2007; Franceschelli and Musolesi 2024). Many approaches measure novelty as distance from other artifacts within a class and value as performance relative to domain norms (Maher 2010). While these frameworks provide domain-independent evaluation criteria, they treat creativity

as a property of an artifact relative to a distribution and not as a structured relation between specific works.

In many artistic domains, especially literature, new works do not simply occupy distant regions of a space of possibilities. They emerge through transformation of identifiable predecessors, retaining aspects of form, theme, or organization while reconfiguring others. Modelling creativity in such contexts therefore requires moving beyond class-level deviation toward an analysis of how representational structure is preserved and reshaped across works.

This relational view resonates with the social theory of Gabriel Tarde and the developmental psychology of James Mark Baldwin. In *The Laws of Imitation* (Tarde 1903), Tarde argued that social phenomena propagate through imitation, with variations introduced in transmission becoming the basis of innovation. Baldwin similarly described imitation as a recursive process through which repetition consolidates habits while enabling modification (Baldwin 1894). In both accounts, invention is not opposed to imitation but emerges from it. Novelty arises through selective transformation of inherited structure instead of originating from nothing.

Julia Kristeva, drawing on Bakhtinian dialogism, described every text as "a mosaic of quotations," framing intertextuality as the intersection of prior textual surfaces rather than the expression of a self-contained origin (Kristeva 1980). Dialogism here refers to the idea that every utterance is shaped by and answers to prior and anticipated utterances (Bakhtin 2010). Roland Barthes likewise emphasized the plurality of cultural codes that converge within a work, shifting attention from authorial origin to the network of textual traces that constitute it (Barthes 1977). Gérard Genette introduced analytic precision through his typology of transtextual relations, particularly his concept of hypertextuality, which foregrounds structured transformation between a later text and its precursor (Genette 1997). Across these accounts, literary production appears not as isolated invention but as patterned reconfiguration of prior forms.

We operationalise this insight through a layered framework for textual comparison. Instead of collapsing two texts into a single similarity score, we decompose comparison across five representational strata: lexical surface patterns, semantic similarity in embedding space, conceptual category distributions, structural organization, and narrative dynamics across the text. This design draws on

established traditions in stylistic analysis, stylometry, and influence detection (Burrows 2002; Reagan et al. 2016; Argamon et al. 2009; Burns et al. 2021; Xing 2025), which have long examined these features as distinct dimensions of textual relation. This framework integrates these dimensions into a unified relational model oriented away from authorship attribution and class-level similarity and toward modelling selective transformation between specific works.

We further hypothesise that creative transformation is characterised by non-trivial retention combined with substantive divergence. If retention is negligible, the relation between works is indistinguishable from random variation; if retention is near total, the result approaches reproduction. Creativity therefore occupies a peaked region of the retention spectrum, contingent upon the presence of meaningful divergence in one or more other representational channels. It arises not at the extremes of imitation or rupture, but where structured inheritance coexists with structured transformation.

This paper makes three contributions:

1. We introduce a multi-channel framework for modelling creative transformation between literary texts across lexical, semantic, conceptual, structural, and narrative levels.
2. We propose a directional alignment and comparison pipeline that measures retention and divergence between texts while accounting for structural variation in narrative progression.
3. We introduce a channelled aggregation mechanism that captures selective imitation by identifying the representational dimension in which structural inheritance is strongest while allowing other dimensions to diverge.

The framework is intended as a diagnostic model of literary transformation, not as a definitive classifier of literary influence. It combines directional alignment, multi-channel representation, control-calibrated comparison, and channelled aggregation to produce interpretable transformation profiles for historically motivated reference pairs.

## Related Work

This work draws on two research areas: computational models of creativity, and computational approaches to stylometry and literary influence.

### Computational Models of Creativity

Computational creativity research has undergone several conceptual shifts. Early work formalized creativity as structured search within constrained conceptual spaces, modelling combinational, exploratory, and transformational movement over rule-governed domains (Boden 2004; Wiggins 2006). Creativity was therefore treated primarily as a generative process operating within a space of possibilities. Later work shifted attention toward evaluation, proposing empirical criteria and standardized procedures for assessing novelty and value in creative artifacts (Ritchie 2007; Maher 2010; Jordanous 2012). Creativity increasingly became a property to be measured as well as a process to be described.

With the rise of large-scale machine learning, a distributional perspective emerged. Generative systems learn statistical regularities and produce outputs that deviate from learned patterns (Franceschelli and Musolesi 2022). Novelty has been framed as prediction violation (Grace and Maher 2019), as a balance between novelty and usefulness (Mukherjee and Chang 2023), or as a quantity difficult to stabilize across automatic metrics in large language models (Lu et al. 2025). In parallel, feature-based approaches operationalize creativity as deviation within a corpus, computing scalar scores from similarity networks, semantic distance, or temporal drift in feature space (Shrivastava et al. 2017; Orwig et al. 2025; Jimenez-Mavillard and Suarez 2022).

### Stylometry, Authorship and Influence

Computational stylometry is grounded in the premise that authors exhibit statistically stable stylistic patterns. Early quantitative studies showed that high frequency function words can reliably distinguish authors (Mosteller and Wallace 1963). Later work expanded stylistic modelling to lexical, syntactic, and structural features such as n-grams, part-of-speech distributions, and vocabulary richness (Argamon et al. 2009). Distance-based measures, especially Burrows' Delta, further formalized stylistic similarity as deviation in frequency space (Burrows 2002).

With the adoption of machine learning, authorship attribution became increasingly framed as supervised classification. Stylometric features were combined with statistical and machine learning classifiers to scale across larger author sets and domains (Stamatatos 2009). Studies also examined the relative role of stylistic and topical signals (Sari, Stevenson, and Vlachos 2018), while later approaches combined stylometric features with classical models (Alshamasi and Menai 2022; Boriceanu and Băltoiu 2025) or used pretrained language models to detect stylometric differences (Barlas and Stamatatos 2020; Fabien et al. 2020).

Literary influence poses a more complex problem than authorship attribution. Instead of identifying the stylistic fingerprint of a single author, influence concerns relations between works produced by different authors and often separated by time, genre, or language. Early computational approaches focused on lexical overlap and text reuse, detecting quotation or phrase-level borrowing through n-gram similarity and related methods (Allen 2011; Forstall et al. 2015). Later work moved beyond surface overlap through semantic representations, embeddings, topic models, and latent semantic analysis, allowing texts to be compared through thematic or latent similarity even when wording differs (Burns et al. 2021; Manjavacasa, Karsdorp, and Kestemont 2020; Xing 2025; Blei, Ng, and Jordan 2003; Jockers and Mimno 2013). These approaches motivate the semantic channel of our framework, but our goal is not passage retrieval alone. We use semantic similarity as one component of a broader transformation profile.

Other work compares texts through stylistic, conceptual, and narrative structure. Function-word transition networks model style as relations between frequent grammatical markers (Segarra, Eisen, and Ribeiro 2015). WordNet supersenses and concreteness norms provide interpretable

Author	Work	Year	Literary Period	Tokens
Henry James	The Portrait of a Lady	1881	Late Victorian	233k
Henry James	Daisy Miller	1878	Late Victorian	22k
Edith Wharton	The House of Mirth	1905	American Realism	134k
Edith Wharton	The Age of Innocence	1920	Modernism / Realism	105k
George Eliot	Middlemarch	1871	Victorian Realism	324k
George Eliot	The Mill on the Floss	1860	Victorian Realism	214k
Jane Austen	Pride and Prejudice	1813	Regency	123k
Jane Austen	Emma	1815	Regency	165k
Thomas Hardy	Tess of the d’Urbervilles	1891	Late Victorian	155k
Oscar Wilde	The Picture of Dorian Gray	1890	Aestheticism	80k
Charles Dickens	David Copperfield	1850	Victorian	364k
Charles Dickens	A Tale of Two Cities	1859	Victorian	138k
Bram Stoker	Dracula	1897	Gothic / Late Victorian	163k
Gustave Flaubert	Madame Bovary (Eleanor Marx trans.)	1856	French Realism	118k
Gustave Flaubert	Salammbô (Eleanor Marx trans.)	1862	French Realism	107k
James Joyce	Dubliners	1914	Modernism	69k
James Joyce	Ulysses	1920	High Modernism	270k
Homer	The Odyssey (Alexander Pope trans.)	1725	Neoclassical Translation	115k
Homer	The Odyssey (Samuel Butler trans.)	1900	Victorian Translation	119k
Homer	The Iliad (Pope translation)	1715	Classical Epic	199k
Virginia Woolf	The Voyage Out	1915	Modernism	140k
Virginia Woolf	Mrs Dalloway	1925	High Modernism	65k
Mary Shelley	Frankenstein	1818	Romantic / Gothic	75k

Table 1: Corpus used in the experiments

Source Work	Target Work
The Portrait of a Lady	The House of Mirth
Pride and Prejudice	Middlemarch
Middlemarch	Tess of the d’Urbervilles
Odyssey (Pope)	Ulysses
Odyssey (Pope)	Mrs Dalloway
Odyssey (Butler)	Ulysses
Odyssey (Butler)	Mrs Dalloway
Odyssey (Pope)	Odyssey (Butler)
Madame Bovary	Dubliners
Madame Bovary	The Age of Innocence

Table 2: Reference text pairs based on historically documented literary relationships

conceptual categories (Ciaramita and Altun 2006; Brysbaert, Warriner, and Kuperman 2014), while digital humanities work has compared narrative dynamics through sentiment trajectories, pacing, and plot progression (Reagan et al. 2016). These strands motivate the structural, conceptual, and narrative channels of the model.

Finally, influence may involve compression, expansion, omission, or rearrangement, making direct passage-to-passage comparison insufficient. Dynamic time warping compares ordered sequences that differ in length or pacing while preserving trajectory structure (Berndt and Clifford 1994). Prior studies have usually examined lexical reuse, semantic similarity, stylistic structure, conceptual categories, or narrative trajectories separately. Our framework combines these signals in a single directional, control-calibrated comparison of source and target texts, using the resulting five-channel profile to characterize creative transformation as selective retention and divergence.

## Dataset

The corpus consists of 23 literary works spanning the eighteenth to early twentieth centuries, including major texts from the English, American, and European literary canon. The dataset covers multiple literary periods, including Romanticism, Victorian realism, and high modernism, and contains approximately 3.5 million tokens in total. Table 1 lists the works included in the dataset together with their authors, publication dates, translators where applicable, literary periods, and approximate token counts after preprocessing.

The corpus was assembled to include both historically documented literary relationships and stylistically comparable but unrelated works. Instead of treating influence as a binary classification problem, the corpus serves as a comparative benchmark for analysing patterns of transformation between texts. Pairs of historically related works provide reference cases in which transformation is expected to occur, while unrelated comparisons provide a background distribution of stylistic similarity across the corpus.

To construct these reference comparisons, we define a set of text pairs based on well-established relationships discussed in literary scholarship. These include connections such as the influence of Henry James on Edith Wharton, Austen’s influence on Eliot, Eliot’s influence on Hardy, the influence of Flaubert on later realist and modernist writers, and the well-known engagement of Joyce and Woolf with Homeric epic. The dataset also includes comparisons between different translations of the *Odyssey*, allowing the framework to capture relationships that preserve narrative structure while substantially altering lexical expression. The selected reference pairs are shown in Table 2.

In addition to these reference comparisons, control pairs are constructed by pairing texts from the corpus that are not

historically linked through documented influence. These comparisons include both cross-tradition pairs and pairs drawn from similar literary periods. Together, these controls provide a background distribution of stylistic and structural similarity against which the transformation profiles of the reference pairs can be compared.

The texts were obtained from the Project Gutenberg digital library. Prior to analysis, non-content material was manually removed, including metadata such as source information, author descriptions, and editorial notes, retaining only the main body of the literary work. All texts are analysed in English. Non-English works are included through established English translations. The text was normalized through lowercasing, and tokenization was performed using a rule-based tokenizer.

For downstream comparison, texts were segmented into sequential units using one of two strategies: chapter boundaries or overlapping sliding windows. Overlapping sliding windows were constructed to reach a target size of approximately 2000 tokens, with successive windows advanced by a fixed step of 500 tokens. This segmentation allows structural patterns to be compared across works of different lengths while preserving local narrative organization.

### Method: The Channelled Imitation Model

To operationalise the theoretical framework proposed by Gabriel Tarde and James Mark Baldwin, our pipeline models literary creativity as structured transformation of prior texts. In this view, new works emerge through selective preservation and modification of earlier textual structures. Instead of treating creativity as deviation from a statistical distribution of artifacts, the model measures how representational structures are retained or transformed between specific works.

The proposed framework therefore compares a source text and a derived text across multiple representational layers. Each channel captures a different level of textual organisation. Higher similarity within a channel indicates structural inheritance, while divergence reflects creative transformation. By examining patterns of retention and divergence across channels simultaneously, the model characterises how literary works inherit and reshape earlier forms.

### Preprocessing and Directional Alignment

Before influence can be analysed, the texts must first be placed into a comparable structure. Because derived works frequently omit, reorder, or expand upon material from their sources, direct one-to-one alignment between passages is insufficient. A later work may compress an episode, elaborate on a minor event, or introduce entirely new sections while still maintaining a recognisable structural relationship to the original.

To account for these transformations, each text is segmented into sequential windows, as described in the previous section, and encoded using Sentence-BERT embeddings (specifically the `all-MiniLM-L6-v2` model) (Reimers and Gurevych 2019). Alignment between the two texts is then computed using Dynamic Time Warping (DTW) (Berndt and Clifford 1994), which allows flexible many-

Channel	Representative features
Lexical	Token-frequency and content-word distributions, punctuation statistics, type-token ratio, inter-window lexical overlap
Structural	Function-word adjacency networks (density, degree distributions), part-of-speech transition patterns
Semantic	Sentence-BERT embeddings, LDA topic mixtures, LSA representations
Conceptual	Supersense categories, named-entity-type distributions, lexical concreteness
Narrative	Sentence-length pacing, sentiment trajectories, syntactic-complexity variation

Table 3: Feature layers used to capture each channel of textual transformation.

to-one or one-to-many correspondences between windows while preserving overall narrative order.

The alignment procedure is directional. Instead of treating the texts symmetrically, the algorithm anchors the derived text ( $B$ ) to the source text ( $A$ ), reflecting the theoretical assumption that influence flows from the earlier work toward the later one. To reduce matches between passages that occur at very different relative positions in the two texts, the DTW cost function incorporates a positional bias:

$$D(i, j) = d_{\cosine}(A_i, B_j) + \lambda \left| \frac{i}{N} - \frac{j}{M} \right| \quad (1)$$

where  $d_{\cosine}$  is the cosine distance between embedding vectors for windows  $A_i$  and  $B_j$ ,  $N$  and  $M$  denote the number of windows in the two texts, and  $\lambda$  controls the strength of the positional bias. The second term measures the difference between the normalized positions of the two windows within their respective texts. Alignments between windows that occur at very different relative narrative positions therefore incur a higher cost. This biases the alignment toward passages that occupy comparable positions in the two works while still allowing local expansions or contractions of narrative material. In all reported experiments we set  $\lambda = 0.15$ . Because both the cosine-distance and positional terms are bounded in  $[0, 1]$ , this caps the positional penalty at roughly 15% of the maximum semantic cost, ensuring it acts as a tie-breaker rather than overriding semantic correspondence. To keep alignments locally monotonic, the DTW search is restricted to a band proportional to the length of the longer sequence, and the traceback is biased toward diagonal steps so that each source window receives a corresponding match.

### Multi-Level Representational Channels

Once the texts are aligned, their relationship is analysed across five representational channels corresponding to different levels of textual organisation. Table 3 presents the feature layers extracted for each channel.

**The lexical channel** captures surface-level stylistic similarity. Each window is represented as a normalized token

frequency distribution. Divergence between the source distribution ( $Q$ ) and derived distribution ( $P$ ) is measured using a dampened Kullback–Leibler divergence:

$$D_{KL}^{\text{damp}}(P \parallel Q) = \ln(1 + \max(0, S)) \quad (2)$$

$$S = \sum_x P(x) \ln \frac{P(x)}{Q(x)} \quad (3)$$

This formulation measures lexical novelty while preventing extreme divergence values when the derived text introduces entirely new vocabulary.

**The structural channel** models stylistic organisation independent of topical content. Function word adjacency networks (Segarra, Eisen, and Ribeiro 2015) are constructed from sequences of stop words within each window. Nodes correspond to function words, while edges represent transitions between successive function words in the text.

From these networks we compute structural metrics including network density and node degree distributions, capturing the global topology of function word usage. Structural similarity between aligned windows is evaluated using cosine similarity and Jensen–Shannon divergence over these distributions. Because function words are less topic dependent than content vocabulary, these measures provide a useful proxy for stylistic and structural inheritance.

**The semantic channel** captures thematic similarity through dense latent representations of meaning. Each window is encoded using contextual embeddings derived from Sentence-BERT, producing dense vector representations optimised for semantic similarity. Pairwise similarity between aligned windows is measured using cosine similarity between embedding vectors.

To complement embedding similarity, we also compute thematic similarity using topic modelling and latent semantic analysis (LSA). Topic distributions are obtained with Latent Dirichlet Allocation (LDA) using 50 topics, while LSA representations are derived from a truncated singular value decomposition of the term–document matrix with 100 dimensions. We use LDA because the corpus is relatively small and because topic mixtures provide an interpretable way to compare themes across aligned windows. The number of topics was fixed before evaluating the reference pairs, and the semantic channel combines LDA with other signals, so the results are not based on a single topic-model configuration. Together, these representations capture thematic similarity even when surface wording differs substantially.

**The conceptual channel** captures persistence of semantic categories that may remain stable even when lexical expression and sentence structure change. While the semantic channel models similarity in dense latent representations of meaning, the conceptual channel compares distributions of interpretable semantic categories across the text.

Three types of signals are used: lexical supersense categories (Ciaramita and Altun 2006), named entity type distributions, and lexical concreteness scores (Brysbart, Wariner, and Kuperman 2014). These distributions provide a general representation of the conceptual landscape of a text,

enabling the model to detect continuity of semantic domains even when vocabulary diverges substantially.

**The narrative channel** models large-scale narrative organisation by analysing how stylistic and emotional signals evolve across the text. This layer captures patterns that unfold across sequences of windows, including sentence-length based pacing, sentiment variation, and changes in structural complexity, instead of focusing only on individual sentences.

For each text, these signals are represented as time series defined over the ordered sequence of windows. For example, sentence-length statistics provide a proxy for narrative pacing, while sentiment scores capture shifts in emotional tone (Reagan et al. 2016). Similarity between the resulting narrative trajectories is then computed using correlation and cosine similarity between the corresponding time series. This representation allows the model to measure whether two texts exhibit similar narrative dynamics, such as comparable pacing patterns or sentiment arcs, even when the specific events or wording differ.

### Discriminative Feature Weighting

Literary works produced within the same historical tradition often share stylistic conventions that do not reflect direct influence. To distinguish genuine textual inheritance from background similarity, the model applies a discriminative feature weighting procedure. For each feature  $f$ , a weight is computed as

$$w_f = \max\left(0.01, \frac{\sigma_{\text{tradition}}(f)}{\sigma_{\text{global}}(f) + \epsilon}\right) \quad (4)$$

where  $\sigma_{\text{tradition}}$  measures feature variance among texts belonging to the same literary tradition, assigned by literary and national grouping such as English Victorian, French realist, or classical epic, and  $\sigma_{\text{global}}$  measures variance across the entire corpus. The constant  $\epsilon$  is a small positive value added for numerical stability to prevent division by zero when global variance is extremely small. Features that exhibit higher variance within a tradition relative to their global variance receive higher weight.

These weights are used to compute standardized feature scores relative to tradition-matched control pairs. The resulting weighted score provides a calibrated estimate of transformation-based influence.

### Channelled Aggregation of Influence

The per-channel scores  $z_c$  used in this section are computed from the discriminatively weighted features described above, standardized against tradition-matched control pairs. Because we hypothesise that creative transformation occurs selectively across representational levels, influence is not expected to appear uniformly across all channels. A derived work may preserve narrative structure while transforming lexical expression, or maintain conceptual themes while radically altering stylistic form.

Each representational channel therefore produces an influence signal measuring the degree to which that level of structure is retained between the aligned texts. Let  $z_c$  denote

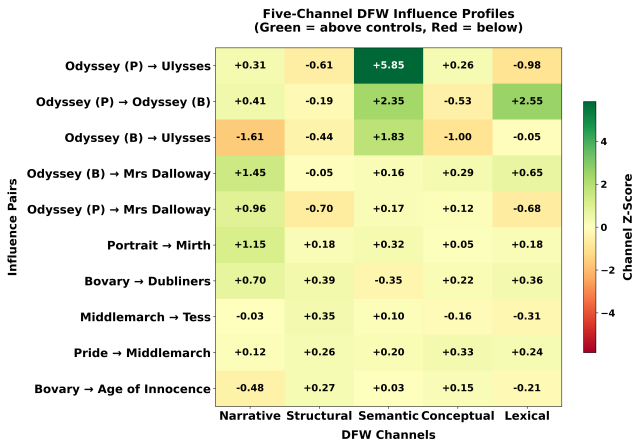


Figure 1: Five-Channel Transformation Profiles

the standardized similarity score for channel  $c$ . The resulting five-dimensional vector

$$\mathbf{z} = (z_{narr}, z_{struct}, z_{sem}, z_{concept}, z_{lex}) \quad (5)$$

represents the influence profile between the two works.

The model identifies the strongest retained structure as the primary locus of inheritance:

$$z_{peak} = \max_c(z_c) \quad (6)$$

Creative transformation often involves strong retention in one channel combined with divergence in others. To capture this pattern, we introduce a diagnostic scoring function that highlights channelled inheritance across representational levels:

$$I = \max(0, z_{peak})(1 + \alpha n_{neg}) \quad (7)$$

where  $n_{neg}$  is the number of channels whose standardized score falls below a divergence threshold of  $z_c < -0.5$ , corresponding to half a standard deviation below the control mean, and  $\alpha$  is a scaling constant set to 0.20. The score is primarily determined by the strength of the dominant channel, with the second factor providing a bounded adjustment for the sharpness of the channelling pattern. Because at most four of the five channels can diverge while one is retained, the maximum bonus is  $1 + 0.20 \times 4 = 1.80$ , ensuring that the divergence term modulates the score without dominating it. This design reflects the theoretical expectation that creative transformation is characterised by a distinctive configuration in which strong inheritance in a specific representational dimension coexists with departure from the source in others, instead of by uniform similarity or uniform distance. The scoring function should therefore be interpreted as a diagnostic of selective transformation rather than as a measure of overall similarity between texts.

## Experiments and Results

This section evaluates whether the framework produces distinctive transformation patterns for historically related texts

Group	$n$	Mean	Median	Std. Dev.
Reference pairs	10	1.94	1.25	2.39
Control pairs	73	0.29	0.19	0.30

Table 4: Aggregate channelled transformation scores for reference and control pairs

compared to unrelated corpus pairs. Instead of treating influence as binary detection, we examine how retention and divergence are distributed across the five representational dimensions introduced above.

The model is applied to the 10 reference relationships in Table 2 and 73 unrelated control pairs constructed from source  $\rightarrow$  target orderings. For instance, *Pride and Prejudice*  $\rightarrow$  *Mrs Dalloway* is retained as a control because Austen predates Woolf, while the reverse direction is excluded as historically impossible. For each pair, the framework computes channel-specific similarity scores and the aggregate transformation score defined in Equation 7.

To assess robustness to segmentation, we run the full pipeline with both chapter-based segmentation and overlapping sliding windows. Reference-control separation is directionally consistent in both settings, but stronger with sliding windows ( $r = 0.74$ ,  $p < .001$ ) than with chapter segmentation ( $r = 0.35$ ,  $p = .07$ ). The remaining analysis therefore reports the sliding-window setting.

The evaluation asks whether historically related works produce stronger transformation signals than controls, and how retention differs across representational layers. Since the model allows selective inheritance, different pairs are not expected to preserve the same kinds of structure. We therefore consider both aggregate score magnitude and the distribution of channel-specific signals.

## Results

Table 4 summarizes the transformation scores. Reference pairs obtain a higher mean than controls (1.94 versus 0.29) and show substantially greater variance. A Mann-Whitney  $U$  test confirms a significant separation ( $U = 635$ ,  $p < 0.001$ , rank-biserial  $r = 0.74$ ), indicating that reference pairs generally receive higher scores than controls. The larger variance among reference pairs suggests that influence does not manifest uniformly, but through different retained layers across different relationships.

Figure 1 presents the five-channel profiles for the reference pairs. Each row shows standardized channel scores relative to the control distribution, where positive values indicate retention and negative values indicate divergence. The profiles show that historically related pairs do not simply display uniformly higher similarity. Instead, they tend to activate one or two representational dimensions while diverging in others. This supports the central claim that creative transformation is best understood as selective inheritance across representational levels.

## Comparison with Single-Channel Baselines

To assess whether channelled aggregation adds information beyond individual dimensions, we compare the full

Signal	$r$	AUC	$p$
Lexical overlap (text reuse) (Allen 2011; Forstall et al. 2015)	-0.04	0.48	.86
POS-trigram structure	0.09	0.55	.64
TF-IDF content similarity	0.10	0.55	.60
LDA topic mixture	0.30	0.65	.13
Sentence-BERT embedding (Burns et al. 2021; Xing 2025)	0.41	0.71	.04
<b>Channelled model (ours)</b>	<b>0.74</b>	<b>0.87</b>	<b>&lt; .001</b>

Table 5: Reference versus control separation for single-channel baselines and the full channelled model under the sliding-window setting. The evaluation includes 10 reference pairs and 73 control pairs.  $r$  is the rank-biserial effect size, AUC is the probability that a reference pair receives a higher score than a control pair, and  $p$  is computed using a two-sided Mann-Whitney  $U$  test.

model with several single-channel baselines. These approximate common approaches to textual relation, including lexical overlap and content similarity used in text-reuse and intertextuality detection (Allen 2011; Forstall et al. 2015), as well as dense embedding similarity used in recent embedding-based intertextual analysis (Burns et al. 2021; Xing 2025). For each baseline, per-window scores are aggregated to the pair level and evaluated by their ability to separate reference pairs from controls.

Table 5 shows that no individual channel separates the groups as strongly as the full model. Lexical overlap, POS-trigram structure, and TF-IDF content similarity remain close to chance ( $AUC \leq 0.55$ ). LDA topic mixtures provide a stronger but non-significant signal, while Sentence-BERT embeddings are the strongest single-channel baseline ( $r = 0.41$ ,  $AUC = 0.71$ ). The channelled model achieves a larger separation ( $r = 0.74$ ,  $AUC = 0.87$ ), suggesting that the relevant signal lies in the configuration of retention and divergence across channels, not in high similarity along one dimension alone.

## Reference Pair Case Studies

Several transformation patterns emerge among the reference pairs. The two English translations of the *Odyssey* show strong semantic and lexical retention, reflecting preservation of narrative content under different stylistic realizations. Pope’s version renders the epic in heroic couplet verse, while Butler’s presents it in unrhymed prose, introducing differences in rhythm, sentence segmentation, and pacing while leaving the underlying semantic structure largely intact.

The Homeric-modernist comparisons show a different configuration. In Joyce’s *Ulysses*, which does not reproduce the *Odyssey* at the level of wording or surface style, the strongest signal appears in the semantic channel. This suggests mediated thematic continuity combined with substantial lexical and structural divergence. Since the aggregate score rewards strong retention in one channel together with divergence in others, this pair scores higher than the more uniformly similar translation pair.

The Woolf comparisons are more narrative-oriented. The Butler *Odyssey*  $\rightarrow$  *Mrs Dalloway* pair produces a stronger signal than the corresponding Pope comparison, with a peak in the narrative channel. Because Butler’s translation uses

prose instead of heroic verse, its form is closer to the narrative conventions Woolf transforms. The stronger Butler signal may therefore reflect mediation through translation, corresponding to hypertextual transformation in Genette’s terms.

Relationships within later prose traditions produce further configurations. The Henry James  $\rightarrow$  Edith Wharton comparison shows elevated narrative and semantic signals, consistent with Wharton’s engagement with Jamesian focalization and psychological interiority. The Eliot  $\rightarrow$  Hardy comparison peaks in the structural channel, likely reflecting continuity in Victorian realist narration and syntactically complex prose without direct content borrowing.

The Flaubert  $\rightarrow$  Joyce comparison is more heterogeneous, with narrative activation and moderate structural and conceptual signals. This fits Flaubert’s role in the development of modern narrative technique, especially free indirect discourse and subjective consciousness, which Joyce extends within a modernist context. Other reference pairs, such as *Madame Bovary*  $\rightarrow$  *The Age of Innocence*, show weaker but interpretable structural and conceptual continuity, suggesting that influence may also appear as distributed partial retention rather than a single dominant channel.

## Reference Pairs and Controls

Table 6 presents representative profiles for reference and control comparisons. Reference pairs more often exhibit a pronounced dominant channel, while controls tend to show weaker or more diffuse similarity across channels.

This contrast is important because many controls are not maximally distant comparisons. They often come from similar periods, related prose traditions, or overlapping stylistic environments, and therefore share broad conventions of narration, syntax, and theme. Even against this stronger baseline, reference pairs more often display a clear dominant channel or coherent retained structure.

The full ranking across all 83 pairs shows the same tendency: the six highest transformation scores are all reference pairs, and seven of the ten reference pairs fall within the top 20. The three lower-ranking reference pairs have flatter profiles, suggesting that some realist influence may be distributed weakly across several channels. Overall, the comparison supports the view that creative transformation

Pair	Type	Score	Narr	Struct	Sem	Concept	Lex	Peak
Odyssey (Pope) → Ulysses	Influence	8.19	0.31	-0.61	5.85	0.26	-0.98	Semantic
Odyssey (Pope) → Odyssey (Butler)	Influence	3.06	0.41	-0.19	2.35	-0.53	2.55	Lexical
Odyssey (Butler) → Ulysses	Influence	2.56	-1.61	-0.44	1.83	-1.00	-0.05	Semantic
The Portrait of a Lady → The House of Mirth	Influence	1.15	1.15	0.18	0.32	0.05	0.18	Narrative
Middlemarch → Tess of the d’Urbervilles	Influence	0.35	-0.03	0.35	0.10	-0.16	-0.31	Structural
Pride and Prejudice → Middlemarch	Influence	0.33	0.12	0.26	0.20	0.33	0.24	Conceptual
Middlemarch → The Portrait of a Lady	Control	0.60	0.60	0.02	0.09	0.08	0.25	Narrative
Pride and Prejudice → The Voyage Out	Control	0.81	0.81	0.39	0.17	0.50	0.69	Narrative
Madame Bovary → Tess of the d’Urbervilles	Control	0.93	0.71	0.93	0.03	0.18	-0.14	Structural
Madame Bovary → The House of Mirth	Control	0.74	0.09	0.61	-0.10	-0.06	-0.53	Structural

Table 6: Representative transformation profiles for influence and control pairs under the chronologically filtered control setting

operates through selective retention and divergence, not uniform similarity.

### Limitations

The present study is limited by the size and composition of its corpus, which focuses on a small set of canonical works from the eighteenth to early twentieth centuries. It should therefore be understood as an exploratory evaluation setting for developing the model, not as a representative sample of literary creativity in general. Other forms of influence, including diffuse genre conventions, multilingual transmission, unconscious stylistic inheritance, or mediation through broader cultural movements, may require larger datasets and additional representational channels.

The reference pairs are based on relationships discussed in literary scholarship rather than formally verifiable ground truth, and should be interpreted as historically plausible cases of transformation, not definitive instances of influence. Likewise, the model’s representational channels are simplified proxies for complex literary phenomena: narrative structure, conceptual domains, and stylistic organization cannot be fully captured by the features used here. The resulting profiles should therefore be read as interpretable evidence of structural relation and as prompts for further literary analysis, not as replacements for close reading. Finally, because the framework assumes a plausible precursor, it is less applicable to works whose originality lies mainly in departing from a broader field; such cases may be better characterized by distributional novelty measures, making genealogical transformation and corpus-level novelty complementary perspectives.

### Conclusion

This paper introduced a multi-channel framework for modelling creative transformation between literary texts. Instead of treating creativity as deviation within a corpus, the proposed approach analyses how representational structures are selectively preserved or transformed across lexical, semantic, conceptual, structural, and narrative levels.

Applying the framework to a corpus of canonical literary works shows that historically related texts tend to exhibit distinctive transformation profiles. Different relationships preserve different layers of structure, ranging from narrative organization in closely related prose traditions to semantic or conceptual continuity in more distant reinterpretations such

as modernist engagements with classical epic. These findings support the view that literary creativity often operates through selective inheritance, with similarity concentrated in specific representational layers instead of appearing uniformly across texts.

More broadly, the results suggest that modelling influence as a multi-layered relation between specific works offers a complementary perspective to existing distribution-based approaches to computational creativity. By decomposing textual comparison across representational channels, the framework provides a way to analyse how creative works simultaneously inherit and transform earlier structures.

The framework may also be useful for evaluating creative text generation. For example, systems prompted to imitate an author or transform an existing narrative may reproduce surface style without preserving deeper semantic, conceptual, or narrative structure. A multi-channel transformation profile could help distinguish shallow stylistic imitation from more substantive transformation of source material.

Future work may extend this approach to larger corpora, additional literary traditions, multilingual settings, and other cultural domains where creativity emerges through structured transformation of prior artifacts.

### Acknowledgements

The authors thank the Romanian writer Ștefan Agopian, whose reflections on literature and the creative process inspired the central idea of this work. They are also grateful to Professor Andra Băltoiu, whose suggestions on stylistic influence and creative transformation helped shape the initial direction of this work.

This work was supported by Ministry of Research, Innovation and Digitization, CNCS-UEFISCDI, project SIROLA, number PN-IV-P1-PCE-20231701, within PNCDI IV.

### References

- Allen, G. 2011. *Intertextuality*. routledge.
- Alshamasi, S., and Menai, M. 2022. Ensemble-based clustering for writing style change detection in multi-authored textual documents. In *CLEF (Working Notes)*, 2357–2374.
- Argamon, S.; Koppel, M.; Pennebaker, J. W.; and Schler, J. 2009. Automatically profiling the author of an anonymous text. *Communications of the ACM* 52(2):119–123.

- Bakhtin, M. M. 2010. *The Dialogic Imagination: Four Essays*, volume 1. University of Texas Press.
- Baldwin, J. M. 1894. Imitation: A chapter in the natural history of consciousness. *Mind* 3(9):26–55.
- Barlas, G., and Stamatatos, E. 2020. Cross-domain authorship attribution using pre-trained language models. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, 255–266. Springer.
- Barthes, R. 1977. *Image-music-text*, volume 6135. Macmillan.
- Berndt, D. J., and Clifford, J. 1994. Using dynamic time warping to find patterns in time series. In *Proceedings of the 3rd international conference on knowledge discovery and data mining*, 359–370.
- Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3(Jan):993–1022.
- Boden, M. A. 2004. *The creative mind: Myths and mechanisms*. Routledge.
- Boden, M. A. 2016. *AI: Its nature and future*. Oxford University Press.
- Boriceanu, I., and Băltoiu, A. 2025. Style change detection using graph and structural-linguistic features for multi-author writing analysis. *Working Notes of CLEF*.
- Brysbaert, M.; Warriner, A. B.; and Kuperman, V. 2014. Concreteness ratings for 40 thousand generally known english word lemmas. *Behavior research methods* 46(3):904–911.
- Burns, P. J.; Brofos, J. A.; Li, K.; Chaudhuri, P.; and Dexter, J. P. 2021. Profiling of intertextuality in latin literature using word embeddings. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4900–4907.
- Burrows, J. 2002. ‘delta’: a measure of stylistic difference and a guide to likely authorship. *Literary and linguistic computing* 17(3):267–287.
- Ciaramita, M., and Altun, Y. 2006. Broad-coverage sense disambiguation and information extraction with a super-sense sequence tagger. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, 594–602.
- Fabien, M.; Villatoro-Tello, E.; Motlicek, P.; and Parida, S. 2020. Bertaa: Bert fine-tuning for authorship attribution. In *Proceedings of the 17th International Conference on Natural Language Processing (ICON)*, 127–137.
- Forstall, C.; Coffee, N.; Buck, T.; Roache, K.; and Jacobson, S. 2015. Modeling the scholars: Detecting intertextuality through enhanced word-level n-gram matching. *Digital Scholarship in the Humanities* 30(4):503–515.
- Franceschelli, G., and Musolesi, M. 2022. Deepcreativity: measuring creativity with deep learning techniques. *Intelligenza Artificiale* 16(2):151–163.
- Franceschelli, G., and Musolesi, M. 2024. Creativity and machine learning: A survey. *ACM Computing Surveys* 56(11):1–41.
- Genette, G. 1997. *Palimpsests: Literature in the second degree*, volume 8. U of Nebraska Press.
- Grace, K., and Maher, M. L. 2019. Expectation-based models of novelty for evaluating computational creativity. In *Computational creativity: The philosophy and engineering of autonomously creative systems*. Springer. 195–209.
- Jimenez-Mavillard, A., and Suarez, J. L. 2022. A computational approach for creativity assessment of culinary products: The case of elbulli. *AI & SOCIETY* 37(1):331–353.
- Jockers, M. L., and Mimno, D. 2013. Significant themes in 19th-century literature. *Poetics* 41(6):750–769.
- Jordanous, A. 2012. A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. *Cognitive Computation* 4(3):246–279.
- Kristeva, J. 1980. Word, dialogue, and novel. desire in language: A semiotic approach to literature and art. *T. Gora et al.*
- Lu, L.-C.; Liu, M.; Lu, P.-C.; Tian, Y.; Sun, S.-H.; and Peng, N. 2025. Rethinking creativity evaluation: A critical analysis of existing creativity evaluations. *arXiv preprint arXiv:2508.05470*.
- Maher, M. L. 2010. Evaluating creativity in humans, computers, and collectively intelligent systems. In *Proceedings of the 1st DESIRE Network Conference on Creativity and Innovation in Design*, 22–28.
- Manjavacasa, E.; Karsdorp, F.; and Kestemont, M. 2020. A statistical foray into contextual aspects of intertextuality. *Proceedings http://ceur-ws.org ISSN 1613:0073*.
- Marcus, G., and Davis, E. 2019. *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon Books.
- Mosteller, F., and Wallace, D. L. 1963. Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed federalist papers. *Journal of the American Statistical Association* 58(302):275–309.
- Mukherjee, A., and Chang, H. 2023. The creative frontier of generative ai: Managing the novelty-usefulness tradeoff. *arXiv preprint arXiv:2306.03601*.
- Orwig, W.; Luchini, S. A.; Beaty, R. E.; and Schacter, D. L. 2025. A ‘sweet spot’ for creative ideation: Non-linear associations between semantic distance and creativity. *The Journal of Creative Behavior* 59(3):e70041.
- Reagan, A. J.; Mitchell, L.; Kiley, D.; Danforth, C. M.; and Dodds, P. S. 2016. The emotional arcs of stories are dominated by six basic shapes. *EPJ data science* 5(1):31.
- Reimers, N., and Gurevych, I. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.
- Ritchie, G. 2007. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines* 17(1):67–99.

- Runco, M. A., and Jaeger, G. J. 2012. The standard definition of creativity. *Creativity research journal* 24(1):92–96.
- Sari, Y.; Stevenson, M.; and Vlachos, A. 2018. Topic or style? exploring the most useful features for authorship attribution. In *Proceedings of the 27th international conference on computational linguistics*, 343–353.
- Segarra, S.; Eisen, M.; and Ribeiro, A. 2015. Authorship attribution through function word adjacency networks. *IEEE Transactions on Signal Processing* 63(20):5464–5478.
- Shrivastava, D.; CG, S. A.; Laha, A.; and Sankaranarayanan, K. 2017. A machine learning approach for evaluating creative artifacts. *arXiv preprint arXiv:1707.05499*.
- Stamatatos, E. 2009. A survey of modern authorship attribution methods. *Journal of the American Society for information Science and Technology* 60(3):538–556.
- Tarde, G. 1903. *The laws of imitation*. H. Holt.
- Wiggins, G. A. 2006. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-based systems* 19(7):449–458.
- Xing, Y. 2025. Modelling intertextuality with n-gram embeddings. *arXiv preprint arXiv:2509.06637*.