Modeling Writing Development as Coordinated Change Across Linguistic and Semantic Dimensions

Michelle Banawan, Andrew Potter,
Tracy Arner, and **Danielle S. McNamara**

Learning Engineering Institute
Arizona State University
Tempe, AZ, USA
{mbanawan, ahpotter, tarner, dsmcnama}@asu.edu

Abstract

Writing development is often assessed through aggregate improvements in surface-level features, yet less attention has been given to how multiple linguistic dimensions evolve jointly over time. We model writing development as a multidimensional system shaped by stable individual variation and instructional progression across staged assignments, using interpretable linguistic features from the Writing Analytics Toolkit (WAT) and transformer-based sentence embeddings.

Variance partitioning reveals substantial between-student stability alongside stage-dependent change. Mixed-effects models identify non-uniform developmental trajectories: academic focus, information density, and conventional language increase, whereas development of ideas and lexical variety decline, indicating tradeoffs across competing dimensions. Cross-lagged analyses further show dynamic dependencies between dimensions, suggesting coordinated change rather than independent progression.

Embedding-based analyses capture stage-dependent shifts in semantic representation, with larger changes in earlier stages and greater stability in later stages. Although assignment structure contributes to observed variation, stable learner-level variation and cross-stage dependencies extend across instructional tasks.

Together, these findings characterize writing development as structured change in a multidimensional representational system, highlighting the need for computational models that capture stable variation, non-monotonic trajectories, and interactions among linguistic components.

1 Introduction

Understanding how writing develops among college writers under instruction is central to both educational research and computational models of language learning. Prior computational and applied linguistics research emphasizes that writing reflects interactions between individual linguistic repertoires and learning environments (Crossley, 2020; Crossley and Kim, 2022). Computational studies further show that writing exhibits stable individual variation across linguistic and representational spaces, although different representational frameworks capture distinct aspects of writing, including stylistic, functional, and semantic organization (Zhu and Jurgens, 2021). However, most computational approaches evaluate language ability cross-sectionally, with limited attention to how linguistic abilities evolve over time in structured instructional contexts (Crossley, 2020). Writing development is inherently multidimensional, involving coordinated changes across interacting linguistic dimensions that may evolve at different rates (Crossley and Kim, 2022; Goulart and Wood, 2021).

In academic settings, writing instruction typically progresses through sequenced assignments of increasing rhetorical complexity, moving from structured comparison to descriptive elaboration and research-based argumentation. These transitions provide a natural framework for examining development across instructional milestones. However, existing work often focuses on overall improvement or isolated features, rather than modeling the organization and interaction of multiple linguistic dimensions within a staged curriculum.

This study models writing development using a repeated-measures dataset of student essays collected across five instructional stages in a first-year university writing course, replicated across three sections. Each essay is represented using nine interpretable functional indices (component scores) derived from automated linguistic analysis, including Sophisticated Wording, Development of Ideas, Word Variety, Information Density, Academic Focus, and Conversational Writing Style. These in-

Proceedings of the 1st Workshop on Computational Developmental Linguistics (CDL), pages 83–91
July 3, 2026 ©2026 Association for Computational Linguistics

dices capture higher-order functional properties of writing rather than surface-level counts.

We address three research questions: (1) How is variance in writing dimensions partitioned across students and course sections? (2) Do these dimensions exhibit systematic change across instructional stages? (3) Are dimensions dynamically coupled, such that prior values of one dimension predict subsequent change in another?

To answer these questions, we integrate variance partitioning, growth modeling, and cross-lagged analyses to characterize writing development as a multidimensional system shaped by stable individual variation and stage-dependent reorganization.

From a computational perspective, we model writing development as representational change over time, examining how texts evolve across interpretable linguistic features and neural representations rather than treating writing ability as a single outcome. We combine WAT-derived linguistic features with transformer-based sentence embeddings to link interpretable linguistic analysis with representation learning. Linguistic features capture variation and interactions across dimensions of writing, whereas embeddings reflect stage-dependent changes in semantic representation.

This work makes three contributions. First, it models writing development as coordinated change across interacting linguistic dimensions within a structured instructional setting. Second, it provides empirical evidence that developmental trajectories are non-uniform and involve tradeoffs between competing dimensions. Third, it links interpretable linguistic features with neural representations, showing how each captures distinct aspects of development.

2 Feature Spaces for Modeling Writing Development

2.1 Functional Linguistic Dimensions (WAT Components)

Functional linguistic dimensions are derived from the Writing Analytics Toolkit (WAT), which organizes large sets of linguistic features into a smaller set of interpretable dimensions using principal component analysis (PCA) (Potter et al., 2026). PCA identifies patterns of covariance among features and reduces them into components representing broader functional properties of writing.

The models are trained on a corpus of academic writing, enabling identification of stable patterns across genres (McNamara et al., 2026). The resulting components capture constructs such as academic language use, cohesion, elaboration, and lexical variety.

In this study, component scores are used as interpretable indices of writing. Because each dimension reflects coordinated variation among multiple features, changes across instructional stages are interpreted as shifts in underlying linguistic subsystems rather than isolated feature-level variation.

2.2 Neural Embedding Representation of Student Writing

Neural embeddings represent text as dense vectors that encode semantic and contextual relationships learned from large corpora (Devlin et al., 2019). We use Sentence-BERT (SBERT), which produces fixed-length sentence embeddings optimized for semantic similarity, enabling efficient comparison using cosine similarity (Reimers and Gurevych, 2019).

Embedding-based representations have been applied to student writing to capture semantic similarity and discourse structure, complementing feature-based approaches that emphasize surface-level variation (Fiacco et al., 2022). However, transformer-based semantic embeddings are known to entangle topical and stylistic information, and are therefore not optimized for isolating stable authorial style independently of content (Wegmann et al., 2022).

In this study, embeddings provide a complementary representation of student writing alongside WAT-derived features. While WAT components capture interpretable variation across linguistic dimensions, embeddings encode the overall semantic structure of texts and are used to analyze trajectories in semantic representation across instructional stages, capturing how assignment structure shapes writing over time.

3 Methods

3.1 Data Construction and Longitudinal Design

The dataset was constructed from institutional Canvas exports containing submission-level metadata for student writing assignments, including assignment identifiers, timestamps, student identifiers, and course section identifiers (McNamara et al., 2022). We restrict the analysis to three sections of a first-year university writing course (English 101) that implemented an identical staged curriculum.

After filtering to include final draft submissions from students who completed the course and consolidating records, the final dataset comprises 308 student essays produced by 62 students across five instructional stages and three course sections, with essays serving as the unit of analysis.

The curriculum consists of five sequenced writing assignments of increasing rhetorical complexity: Compare/Contrast, Illustration, Descriptive, Persuasive Research, and a discussion-based synthesis assignment. These assignments reflect standard genres in introductory composition courses, progressing from structured comparison to descriptive elaboration and research-based argumentation, and provide a structured framework for examining development across instructional stages.

Assignments are mapped to a stage index ($stage\_index \in \{1, 2, 3, 4, 5\}$) based on their curricular order. Because each stage is implemented across all sections, the dataset forms a partially replicated longitudinal design.

Stage 1 corresponds to the Compare/Contrast assignment, Stage 2 to Illustration, Stage 3 to Descriptive writing, Stage 4 to Persuasive Research, and Stage 5 to the synthesis assignment.

Each observation corresponds to a single student submission at a given stage and includes assignment metadata, section identifiers, and the stage index used in subsequent analyses.

3.2 Linguistic Feature Extraction

Linguistic features were extracted from each essay using WAT. PCA-derived component scores were used as functional indices of writing, with each score representing a text's position along an underlying linguistic dimension rather than the frequency of a single feature.

The nine dimensions analyzed are: Sophisticated Wording, Sentence Cohesion, Word Variety, Conversational Writing Style, Academic Focus, Conventional Language, Information Density, Word Concreteness, and Development of Ideas. These dimensions capture higher-order properties of writing, including lexical sophistication, stylistic orientation, informational density, and discourse elaboration.

Because each dimension reflects coordinated variation among multiple features, changes across instructional stages are interpreted as shifts in underlying linguistic subsystems rather than isolated feature-level variation.

3.3 Embedding-based Modeling

To model writing at the representational level, essays were encoded using Sentence-BERT (SBERT) with the `all-mpnet-base-v2` pretrained model (Reimers and Gurevych, 2019), a transformer-based model optimized for semantic similarity. Because essay-length inputs exceeded the model's maximum token length, essays were segmented into smaller text units prior to encoding. Embeddings were generated for each segment and aggregated using mean pooling to obtain a single essay-level semantic representation. The resulting embeddings capture contextual and semantic relationships across the full document in a high-dimensional space (Devlin et al., 2019).

Representational change across instructional stages was quantified by computing cosine similarity between embeddings of consecutive essays written by the same student:

$$sim(v_i, v_j) = \frac{v_i \cdot v_j}{\|v_i\| \, \|v_j\|} \tag{1}$$

Representational change was operationalized as:

$$drift = 1 - sim(v_i, v_j) \tag{2}$$

Higher drift values correspond to lower cosine similarity and therefore indicate greater semantic change between stages. Because subsequent analyses report cosine similarity directly, lower similarity values should be interpreted as greater representational drift.

Because each student contributes one essay per stage, embedding comparisons primarily capture assignment-driven variation rather than within-stage stylistic differences. Accordingly, embeddings are used to model trajectories in semantic representation across instructional stages.

This analysis complements the feature-based approach, in which WAT dimensions capture interpretable variation across linguistic functions, while embeddings capture movement through semantic representation space.

3.4 Functional Linguistic Dimension Models

3.4.1 Variance Structure of Functional Writing Dimensions

To characterize the structure of writing development, variance for each dimension was partitioned into student-level, section-level, and residual components using ANOVA-based decomposition ($Y \sim C(student\_id) + C(asu\_class\_id)$).

Results show substantial between-student stability across several dimensions (Table 2). Sophisticated Wording (0.57), Sentence Cohesion (0.54), Word Variety (0.54), and Conversational Writing Style (0.52) exhibit the largest student-level variance components, with Academic Focus (0.47) and Conventional Language (0.41) also showing substantial stability.

In contrast, Development of Ideas (0.24), Word Concreteness (0.32), and Information Density (0.37) show greater residual variance, indicating higher within-student fluctuation across instructional stages. Section-level variance is negligible across all dimensions (≤ 0.01), suggesting minimal structural differences between course sections.

These results indicate that writing development comprises both stable, trait-like dimensions and more plastic, stage-sensitive components, reflecting partially independent subsystems that differ in responsiveness to instructional context.

3.4.2 Developmental Change Across Instructional Stages

Linear mixed-effects models were estimated for each dimension with stage index as a fixed effect and random intercepts and random slopes for instructional stage at the student level, allowing individual variation in both baseline performance and stage-dependent change. Model convergence diagnostics were inspected for all mixed-effects models, and no major convergence issues were observed. For the growth-model analyses, p-values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure ($\alpha = .05$) across linguistic dimensions.

Results reveal systematic but non-monotonic stage-dependent shifts (Figure 1). Conventional Language shows the strongest positive growth ($b = 0.22$, $SE = 0.04$), followed by Academic Focus ($b = 0.11$, $SE = 0.03$) and Information Density ($b = 0.09$, $SE = 0.04$), indicating increasing formalization, informational density, and greater adherence to grammatical and conventional language use across stages.

In contrast, Development of Ideas decreases significantly ($b = -0.15$, $SE = 0.04$), and Word Variety shows a smaller decline ($b = -0.07$, $SE = 0.03$), suggesting a shift from more exploratory writing toward more constrained, concise, and rhetorically focused expression.

Sophisticated Wording, Sentence Cohesion, and Word Concreteness do not exhibit significant linear change, consistent with their higher trait-level stability.

Functional Dimension	b	SE	Adj. p
Development of Ideas	-0.15	0.04	$< .001$
Word Variety	-0.07	0.03	.025
Conversational Writing	-0.01	0.04	.813
Sentence Cohesion	0.03	0.03	.415
Sophisticated Wording	0.04	0.03	.234
Word Concreteness	0.07	0.05	.115
Information Density	0.09	0.04	.010
Academic Focus	0.11	0.03	.001
Conventional Language	0.22	0.04	$< .001$

Table 1: Linear Growth Effects Across Instructional Stages

Dimension	Student	Section	Residual
Soph. Wording	0.57	0.01	0.43
Sent. Cohesion	0.54	0.00	0.46
Word Variety	0.54	0.00	0.46
Conv. Writing	0.52	0.01	0.48
Academic Focus	0.47	0.00	0.53
Conv. Language	0.41	0.01	0.58
Info Density	0.37	0.01	0.62
Word Concr.	0.32	0.01	0.68
Dev. Ideas	0.24	0.01	0.75

Table 2: Variance partitioning of functional writing dimensions.

Overall, these findings suggest stage-dependent reorganization toward academic register, with gains in grammatical control, concision, and focus rather than uniform improvement across all dimensions.

3.4.3 Cross-Dimensional Coupling

To examine dynamic coupling between functional dimensions, lag–difference models were estimated in which stage-to-stage change in a focal dimension (ΔY_t) was regressed on the prior-stage value of another dimension (X_{t-1}). Models were estimated with cluster-robust standard errors at the student level. Based on theoretical expectations regarding register formalization and the relationship between elaboration and academic focus, consistent with prior work contrasting conversational and informational academic discourse (Biber, 2012), three directional pairings were specified for cross-lagged testing; p-values are reported without adjustment given the small number of pre-specified comparisons:

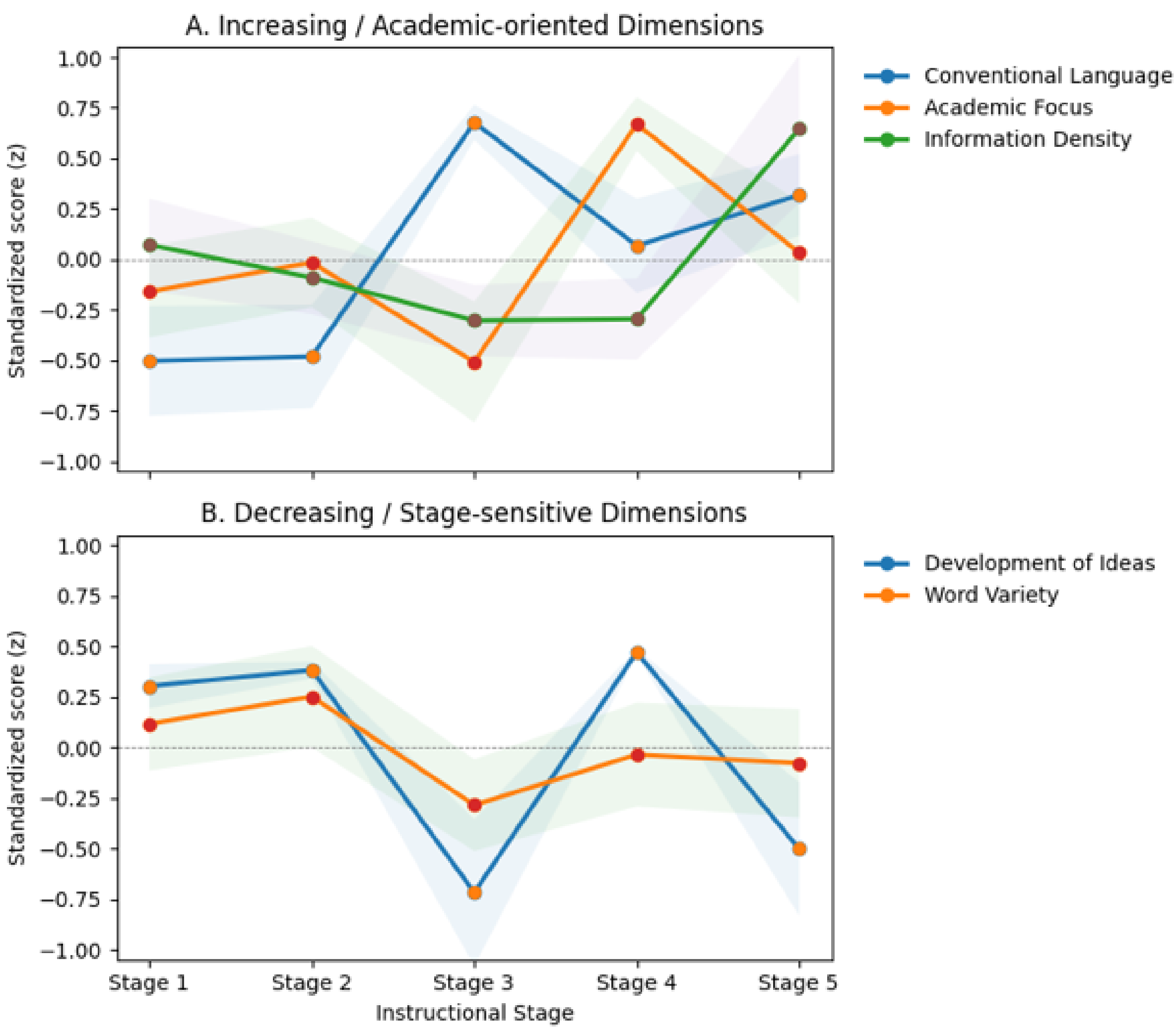

Figure 1: Developmental trajectories of selected WAT dimensions across instructional stages. (A) Increasing dimensions (Conventional Language, Academic Focus, and Information Density) show consistent growth across stages, reflecting movement toward formal and academically oriented writing. (B) Stage-sensitive dimensions (Development of Ideas and Word Variety) exhibit declines and fluctuations, indicating shifts away from exploratory discourse toward more constrained rhetorical structure. Shaded regions indicate variability across students.

$$\Delta Y_t = Y_t - Y_{t-1} = \alpha + \beta X_{t-1} + \varepsilon_t \quad (3)$$

Results indicate significant cross-dimensional interactions across the three pre-specified directional pairings (Table 3). Prior Academic Focus positively predicts change in Conversational Writing ($b = 0.18$, $SE = 0.07$, $p = .012$), while prior Conventional Language predicts increases in Information Density ($b = 0.13$, $SE = 0.06$, $p = .046$), suggesting coordinated and partially reinforcing dynamics across dimensions.

The strongest effect is observed for Development of Ideas, which negatively predicts change in Academic Focus ($b = -0.70$, $SE = 0.11$, $p < .001$), suggesting an inverse relationship between exploratory and more formal writing across stages, although this effect should be interpreted cautiously because the lag–difference specification does not control for prior outcome levels.

Pred.	**Outcome (Δ)**	**b**	**p**
Acad. Focus	Δ Conv. Writing	0.18	.012
Conv. Lang.	Δ Info Density	0.13	.046
Dev. Ideas	Δ Acad. Focus	-0.70	$<$ .001

Table 3: Cross-lagged effects on stage-to-stage change.

Overall, these results show that writing development reflects interacting linguistic subsystems under instructional constraints.

Comparison	Cosine Similarity (SD)
Same student essays	0.22 (0.13)
Different student essays	0.23 (0.16)

Table 4: Representational similarity within and between students.

Stage Transition	Cosine Similarity (SD)
Stage 1 → 2	0.15 (0.10)
Stage 2 → 3	0.10 (0.08)
Stage 3 → 4	0.23 (0.10)
Stage 4 → 5	0.28 (0.13)

Table 5: Representational movement between instructional stages.

3.4.4 Embedding-based Results

We first compared semantic similarity between essays written by the same student and those written by different students. As shown in Table 4, similarity was comparable (same-student: $M = 0.22$, $SD = 0.13$; different-student: $M = 0.23$, $SD = 0.16$), with no significant difference ($p = .30$). This indicates that embeddings primarily capture assignment-driven variation. This pattern is consistent with the semantic orientation of SBERT-based representations, which are optimized to encode contextual and topical similarity rather than stable author-specific stylistic identity.

We then examined similarity across consecutive stages for the same student (Table 5). Early transitions show lower similarity (Stage 1→2: $M = 0.15$; Stage 2→3: $M = 0.10$), indicating substantial representational change. Later transitions show higher similarity (Stage 3→4: $M = 0.23$; Stage 4→5: $M = 0.28$), suggesting increased stability in later stages.

Finally, cross-student similarity varies across stages (Table 6), with the highest similarity at Stage 3 ($M = 0.54$) and lower similarity in earlier and later stages (e.g., Stage 2: $M = 0.20$, Stage 5: $M = 0.23$). This pattern indicates that assignment structure shapes the distribution of texts in semantic space.

Taken together, these results show that embeddings capture stage-dependent, non-linear trajectories driven by instructional context, complementing feature-based analyses that reflect stable variation and subsystem dynamics.

Stage	Mean Pairwise Similarity (SD)
Stage 1	0.37 (0.21)
Stage 2	0.20 (0.11)
Stage 3	0.54 (0.12)
Stage 4	0.43 (0.22)
Stage 5	0.23 (0.13)

Table 6: Stage-level cross-student similarity in embedding space.

4 Discussion

4.1 WAT Dimensions and Implications for Computational Language Learning

The variance structure indicates that writing development preserves stable individual differences alongside stage-dependent change. Several dimensions—particularly lexical sophistication, cohesion, and stylistic features such as academic focus and conversational writing style—exhibit strong trait-level stability, suggesting that models of language development should capture persistent individual variation in addition to aggregate trends.

Developmental trajectories are non-uniform. While academic focus, formal language, and information density increase over time, development of ideas and lexical variety decline. This pattern reflects reweighting among partially competing dimensions rather than uniform improvement, highlighting a limitation of computational models that assume monotonic performance scaling.

These patterns emerge within a structured sequence of assignments with distinct rhetorical demands, suggesting that observed trajectories reflect stage-dependent reorganization under instructional and rhetorical constraints rather than isolated developmental progression alone. However, the presence of stable between-student variation and cross-stage dependencies indicates that the results are not solely attributable to assignment effects. Rather than isolating development from instructional context, writing reorganizes across pedagogically sequenced tasks, with different assignments eliciting shifts in linguistic priorities. Although assignments differ in rhetorical demands, they are designed within introductory composition courses to develop transferable knowledge about writing and rhetorical situations, and prior work suggests that writing quality reflects multiple linguistic dimensions that generalize across tasks (Crossley, 2020).

Cross-lagged analyses further show that dimensions interact over time. The negative relation-

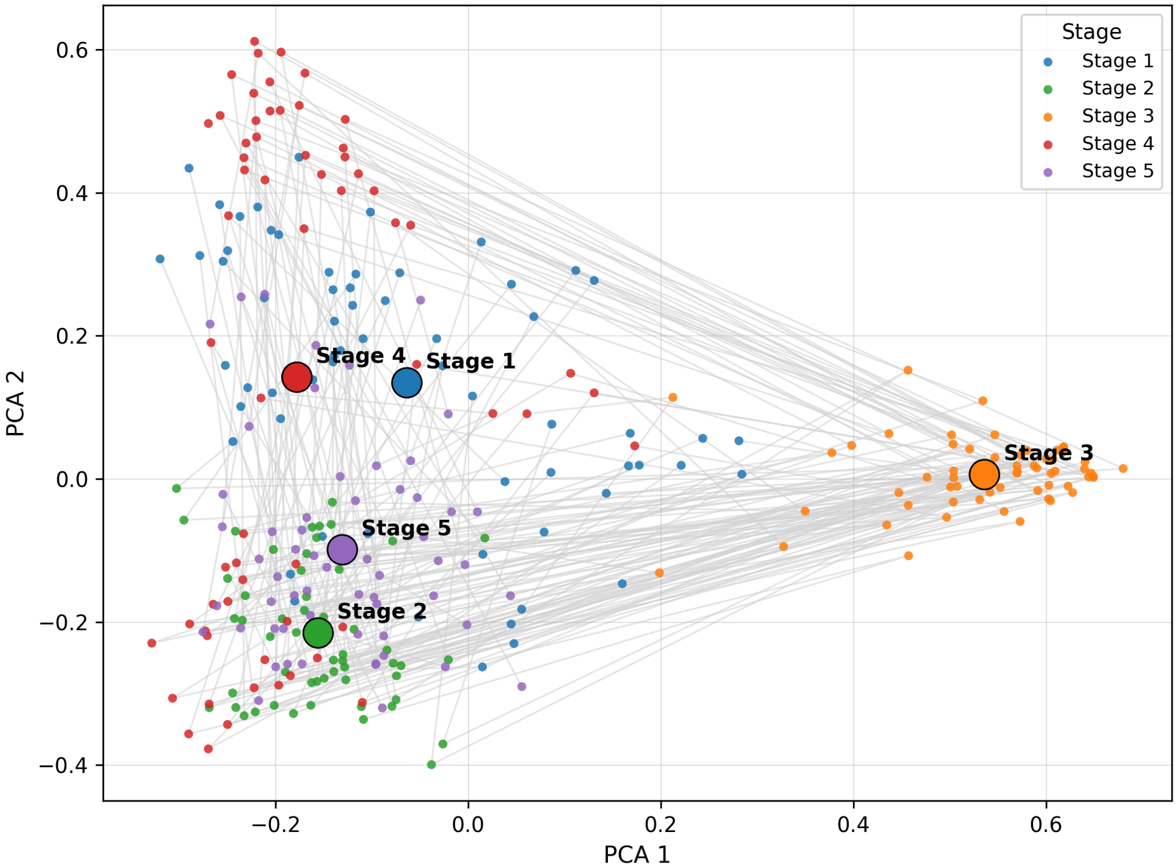

Figure 2: SBERT representation trajectories across instructional stages. Thin gray lines connect successive essays by the same student in PCA-projected embedding space. Colored points represent essays by stage, and larger markers indicate stage centroids. The projection shows substantial overlap and dispersion across stages, with heterogeneous and non-linear trajectories. Certain stages (e.g., Stage 3) occupy relatively distinct regions, consistent with stage-dependent variation observed in similarity analyses.

ship between Development of Ideas and subsequent Academic Focus suggests a possible inverse relationship between exploratory and formal writing, consistent with coordinated adjustments across interacting components rather than independent progression.

Together, these results support a view of writing development as a multidimensional system characterized by stable variation, stage-dependent change, and interactions among competing dimensions (Zhu and Jurgens, 2021). For computational models to approximate human language development, they should account for variability, non-monotonic trajectories, interactions among representational components, and the influence of task structure on observed behavior.

4.2 Embedding-based Analysis

Figure 2 visualizes SBERT-based representation trajectories across instructional stages. The projection shows substantial overlap and dispersion across stages, indicating that semantic representations vary widely across students and assignments. While certain stages, such as Stage 3, occupy relatively distinct regions of embedding space, most stages exhibit considerable overlap, suggesting that assignment-driven semantic structure does not produce clear separation in low-dimensional projections.

Individual trajectories further illustrate that representational change is non-linear and heterogeneous across students, with paths frequently crossing and diverging rather than following a uniform progression. This pattern is consistent with the quantitative results, which show stage-dependent changes in similarity but no strong evidence of stable author-specific clustering.

Embedding-based analyses indicate that semantic representations are primarily shaped by instructional context. The absence of differentiation between same-student and different-student essays suggests that embedding similarity reflects assignment-driven variation rather than stable stylistic differences. At the same time, embeddings capture stage-dependent, non-linear trajectories, with larger representational shifts occurring in earlier stages and greater stability emerging in later stages.

Cross-student similarity patterns further indicate that assignment structure shapes the distribution of texts in representation space, with some stages producing more homogeneous responses than others. Taken together, these findings show that embeddings capture trajectories in semantic space driven by instructional sequencing, complementing feature-based analyses that reflect stable variation and interactions among linguistic dimensions. Writing development can therefore be understood as an interaction between learner-level variation and task-level constraints, with different representational frameworks capturing distinct aspects of this process.

5 Conclusion

This study models writing development as a multidimensional system shaped by stable individual variation and instructional progression across staged assignments. Using interpretable linguistic dimensions, we show that some aspects of writing exhibit strong stability across students, while others change systematically with instructional demands.

Patterns across instructional stages are non-uniform: academic focus, formal language, and information density increase over time, whereas development of ideas and lexical variety decline. Cross-dimensional analyses further reveal interactions among dimensions, including a possible inverse relationship between exploratory and more formal writing.

Embedding-based analyses complement these findings by capturing stage-dependent, non-linear trajectories in semantic representation. Rather than reflecting stable stylistic similarity, embeddings primarily encode assignment-driven variation, with larger shifts in early stages and greater stability later.

Together, these results characterize writing development as coordinated change across interacting dimensions under instructional constraints.

Limitations

A key consideration is that instructional stages correspond to different assignment types with distinct rhetorical demands that are intentionally sequenced within the curriculum. As a result, observed changes reflect both developmental processes and structured task progression. This design provides a controlled instructional framework for examining how writing reorganizes across pedagogically defined stages, but it does not fully isolate development from task-specific effects. Future work could extend this approach by examining development within repeated or more tightly controlled task settings.

Because lag–difference models did not include lagged outcome controls, cross-dimensional effects should be interpreted as suggestive dependencies rather than strong causal or mechanistic relationships.

Additionally, the dataset is drawn from a single institutional context and course, with 62 students contributing repeated observations across instructional stages. Although the longitudinal design provides multiple observations per student, the modest sample size may limit statistical power and the stability of more complex model estimates. Findings may therefore not generalize to other instructional settings, disciplines, or curricula.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305N210041 and R305T240035 to Arizona State University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Douglas Biber. 2012. Register as a predictor of linguistic variation. *Corpus linguistics and linguistic theory*, 8(1):9–37.

Scott A Crossley. 2020. Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, 11(3):415–443.

Scott A Crossley and Minkyung Kim. 2022. Linguistic features of writing quality and development: A longitudinal approach. *The Journal of Writing Analytics*, 6(1):59–93.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pages 4171–4186.

James Fiacco, Shiyan Jiang, David Adamson, and Carolyn Rosé. 2022. Toward automatic discourse parsing of student writing motivated by neural interpretation. In *Proceedings of the 17th workshop on innovative use of NLP for building educational applications (BEA 2022)*, pages 204–215.

Larissa Goulart and Margaret Wood. 2021. Methodological synthesis of research using multi-dimensional analysis. *Journal of Research Design and Statistics in Linguistics and Communication Science*, 6(2):107–137.

Danielle S McNamara, Tracy Arner, Elizabeth Reilley, Paul Alvarado, Chani Clark, Thomas Fikes, Annie Hale, and Betheny Weigele. 2022. The asu learning at scale (asu l@ s) digital learning network platform. *Grantee Submission*.

Danielle S. McNamara, Andrew Potter, Zeinab Serhan, Manmeet Singh, Nishad Patne, Tracy Arner, Rod D. Roscoe, Laura K. Allen, and Scott A. Crossley. 2026. The writing analytics tool: A learning engineering approach to designing ai-supported writing instruction. In *Proceedings of the Learning Engineering Research Network Convening (LERN 2026): From Insights to Implementation, Learning Engineering in Action*. LERN Convention Proceedings.

Andrew Potter, Zeinab Serhan, Nishad Patne, Püren Öncel, Ishrat Ahmed, Tracy Arner, Rezwana Islam, Rod Roscoe, Laura Allen, Scott Crossley, and Danielle McNamara. 2026. Human-ai collaboration for qualitative analysis in participatory design: Refining the writing analytics tool. *Journal of Educational Data Mining*, 18(1):113–155.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*, pages 3982–3992.

Anna Wegmann, Marijn Schraagen, and Dong Nguyen. 2022. Same author or just same topic? towards content-independent style representations. In *Proceedings of the 7th Workshop on Representation Learning for NLP*, pages 249–268.

Jian Zhu and David Jurgens. 2021. Idiosyncratic but not arbitrary: Learning idiolects in online registers reveals distinctive yet consistent individual styles. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 279–297.