

Original Paper

Stance-taking in Narrative Completion: Attitude Resources in Chinese EFL Students' Continuation Writing

Hezhen Sun¹

¹ College of Foreign Languages, Xinjiang University, Urumqi, 830046, Xinjiang, China

Received: July 12, 2026

Accepted: August 29, 2026

Online Published: September 14, 2026

doi:10.22158/eltls.v8n5p9

URL: <http://dx.doi.org/10.22158/eltls.v8n5p9>

Abstract

Stance-taking poses a significant challenge for writers, especially for those writing in English as a foreign language. To explore the realization and patterning of stance resources in EFL writing, this study investigates English continuation writing, a narrative-like school task, through the lens of the attitude system within the Appraisal framework. Adopting a corpus-assisted comparative design, the study analyzes 99 continuation texts produced under two topics, Homesickness and Alligator Race, focusing on affect, judgment, appreciation, and their inscribed and invoked realizations. Results indicate that topic is the strongest factor affecting attitude distribution: Homesickness elicits more affect, whereas Alligator Race elicits more judgment. Besides, disciplinary background mainly affects appreciation and the use of invoked attitude, and its effect on appreciation depends on the writer's disciplinary background, that is, science students use more invoked resources, especially affording. Finally, essay rating is significantly related to affect and invoked attitude: low-rated essays use fewer affect resources and fewer invoked attitude resources. The study extends the application of the Appraisal framework to narrative-like school writing and has implications for genre-sensitive writing instruction in senior high school English classrooms.

Keywords

stance-taking, attitude system, EFL students, continuation writing

1. Introduction

Merriam-Webster (n.d.) defines stance as “a publicly stated opinion”, and researchers have referred to this concept using various terms such as posture, attitude, appraisal, evaluation, and metadiscourse (Hyland & Jiang, 2016). Biber and Finegan (1988, 1989) were among the first to systematically define stance, stating, “By stance we mean the lexical and grammatical expression of attitudes, feelings, judgments, or commitment concerning the propositional content of a message”. Building on this

foundation, with the proliferation of stance research, studies from diverse perspectives have emerged, each offering its own conceptualization of stance. From a functional linguistic perspective, Hyland (2005) defines stance as the author's textual voice or personality recognized by a particular community, seen as an attitudinal dimension. Stance-taking, in turn, refers to how individuals express a stance in actual language use, either in speech or in writing (Du Bois, 2007; Englebretson, 2007; Jaffe, 2009, as cited in Chung, 2024). After citing a range of studies, Zhang and Zhang (2023) emphasized that how to strategically use stance resources is important for both novice and expert to incorporate multiple perspectives and convey authorial opinions. And Valerdi Zárate (2025) also examined whether attitudinal language can improve the effectiveness of EFL writing, who compared attitude resource of essays written by Mexican undergraduate students to the grades assigned by university professors. His findings concluded that attitudinal resources and the perceived effectiveness of student writing have a significant relationship. We can see the strategic use of evaluative resources plays an important role in language teaching and learning. Absorbed a systemic-functional perspective, we view language as a shared resource for constructing meaning. Individual texts do not use all available linguistic options. Instead, writers select specific resources from the system to express their particular meanings. However, how to appropriately select particular resources remains challenging for novice writers, especially those writing in English as a foreign language (Zhang & Zhang, 2023). To tackle it, this study draws on the Appraisal framework to examine stance-taking in students' writing, as it provides a systematic system that integrates diverse evaluative options according to Hyland (2005, as cited in Deng, Chen, & Gao, 2024), which are often examined separately in prior research (e.g., Hyland, 1998). Previous studies have shown that explicit instruction in evaluation can help L2 learners improve their deployment of stance resources (Chang & Schleppegrell, 2016; Humphrey and MacNaught, 2016; Jou, 2019; Zhang & Zhang, 2021). Recent research has also examined the patterning and distribution of stance resources in learner writing (Deng et al., 2024; Chung & Crosthwaite, 2024; Zhang, 2023; Lam & Crosthwaite, 2018). Nonetheless, much of the current literature focuses on argumentative or academic writing (eg., Xuan, Chen & Zhang, 2025) and relatively less attention has been paid to other genres, especially narratives. Cross-genre research has further shown that the use of stance resources varies across genres. For instance, Biber and Finegan's (1989) cluster analysis identified six distinct stance styles, with personal letters dominated by frequent affect markers and emphatics while academic prose and press reportage overwhelmingly favored a near total absence of stance features. Therefore, findings based on argumentative or academic discourse cannot be readily generalized to other types of writing. In recent years, English continuation writing has been introduced as a new writing task in the college entrance exams of several provinces in western China and many high school English teachers in western China still face challenges in teaching it. This task is a narrative-like genre requiring a clear thematic position, and thus have its own stance pattern. Actually, this task is popular in China for its learning facilitation effects, that is, linguistic alignment (Wang, 2016, 2017). Previous research has reported that this task can help learners acquire vocabulary (Gao & Xu, 2025; Li & Wang, 2025; Zhang, 2017), syntactic

structure (Yan & Li, 2016), discourse organization (Peng, 2017), and overall writing performance (Jiang & Chen, 2015; Wang & Wang, 2025). However, such research has focused primarily on the formal properties of texts, often without systematic attention to the meaning-making functions of linguistic forms in writing. The author's teaching experience also revealed some problems in current classroom practices. During internship, classroom observations and conversations with students suggested that instruction often focused on collecting useful expressions, understanding input passages, and finding possible links between story events. Although students had received training in continuation writing for nearly one year, some still had difficulties in clearly describing out a good continuation essay. When asked how they judged their own writing, several students responded that their evaluation relied on "feeling". Yet, they were unable to identify the specific linguistic features behind this "feeling". Teachers' views further suggested that a clear theme is regarded as a key feature of successful continuation writing. For this reason, the present study adopts the Appraisal framework to explore whether high-rated narrative continuations display distinctive patterns of stance use. Such an analysis may offer a more detailed understanding of how effective continuation writing is achieved. Given that attitude resources represent a major component of the Appraisal framework, this study concentrates on the attitude system.

1.1 Attitude System within the Appraisal Framework

Attitude system within the Appraisal framework deals with evaluative language, including evaluations of people, behaviors, objects, and events. It consists of three main domains, which represent different types of evaluation, including affect, judgment, and appreciation. Affect focuses on emotional responses and reflects how individuals express positive or negative feelings. As Martin and White (2005, p. 41) explain, affect is related to whether people experience emotions such as "happy or sad, confident or anxious, interested or bored." Judgment "deals with attitudes towards behaviour, which we admire or criticize, praise or condemn" (p. 43). Appreciation involves "evaluations of semiotic and natural phenomena, according to the ways in which they are valued or not in a given field" (p. 49). Table 1 presents some examples drawn from Martin and White (2005, pp. 48-56) to illustrate the attitude system.

Table 1. Resources of Attitude

Domains of Appraisal	Sub-domains of Appraisal	Examples
Attitude	Affect	The captain felt anxious.
	Judgment	He played skillfully.
	Appreciation	It was a skillful innings.

Attitude meanings, however, are not always realized through words that directly encode evaluation. They may also be conveyed in a more indirect way. In the Appraisal framework, attitude can therefore

be realized either as inscribed attitude or as invoked attitude. According to Martin and White (2005), these mechanisms can be understood as a cline ranging from provoke through flag to afford, depending on the degree of freedom readers have in aligning with the reading position naturalised by the text. Table 2 presents examples taken from Martin and White (2005, p. 67). Accordingly, the present study takes affect, judgment, and appreciation as its main coding categories, while also considering both inscribed and invoked attitude in the analysis of evaluative meaning.

Table 2. Explicit and Implicit Attitude

Explicit or inscribed attitude	It was our ignorance and our prejudice	
	Flagging	we smashed their way of life
Implicit or invoked attitude	Affording	we brought the diseases
	Provoking	we fenced them in like sheep

1.2 Corpus-Assisted Discourse Studies

Corpus-assisted discourse studies (CADS) is generally regarded as a methodological approach that combines corpus linguistics with discourse analysis. According to Gillings, Mautner and Baker (2023, p. 7), “corpus-assisted discourse studies explores discourse (i.e., language as social practice) through examining corpora (i.e., large computerized sets of textual data).” It emphasizes that the analytical process should integrate quantitative analysis of large-scale corpora with qualitative interpretation of specific texts, thereby achieving a balanced and complementary use of both approaches (Baker et al., 2008; Partington, 2004). A growing body of literature has used CADS in studies based on the Appraisal framework. For example, Xie (2016) employed a corpus-assisted discourse studies method and Appraisal framework (Martin & White, 2005) to analyze a self-compiled corpus of 25 Chinese MA thesis literature reviews. Through manual annotation of attitude, graduation, and engagement, plus quantitative analysis, he investigated the students’ evaluative resources and stance-taking. Similar corpus-based perspectives have also been adopted in L2 writing research beyond the Appraisal framework. For example, researchers have compared linguistic choices among writers from different disciplinary backgrounds (McGrath & Liardét, 2022) and among learners with different levels of proficiency (Crosthwaite & Jiang, 2017). Taken together, these studies suggest that CADS is suitable for comparing attitude resources across groups of texts while also paying attention to contextual meanings created through linguistic choices. In the current study, CADS is combined with the attitude system of the Appraisal framework to explore stance-taking in English continuation writing produced by Chinese Grade 12 students from liberal arts and science backgrounds. Three questions guide the analysis:

- 1) How do Chinese Grade 12 liberal arts and science students use attitude resources in English continuation writing task?

- 2) To what extent does topic influence students' use of attitude resources in English continuation writing tasks?
- 3) How is the use of attitude resources related to essay ratings in English continuation writing tasks?

2. Method

2.1 Data Collection

Participants involved were Grade 12 students enrolled in a senior high school in western China. Most were around 18 years old, and all received English instruction from the same teacher. Collecting writing data from this educational setting required intense contact with the school and classroom teachers. The dataset consisted of two continuation writing tasks, including Alligator Race and Homesickness. The reason why these topics were chosen because they were considered familiar to senior high school students and were connected to their everyday experiences. Regarding the aspect of controlling variables, both writing tasks were completed in an examination-like setting, and all students were asked to write approximately 150 words for each continuation task within a limited time. They were not allowed to use dictionaries or other materials during writing. All compositions were handwritten. After their completion, the classroom teacher collected and assessed their essays, which were subsequently converted into electronic texts for corpus analysis by the author. Table 3 reports the number of essays and word counts included in each sub-corpus across the two writing topics.

Table 3. Details of the Collected Corpora

	Liberal arts		Science	
	No. of essays	Words	No. of essays	Words
Alligator Race	27	5939	25	5454
Homesickness	25	5676	22	4840
Total	52	11615	47	10294
Total corpus size = (n = 99)				

The study was conducted in a natural teaching context, without a control group or a comparison class. Before the students' continuation writing texts were used for analysis, each participant received an information sheet and signed an informed consent form. During data analysis and publication, all participants were coded and anonymized to protect their privacy.

2.2 Essay Rating and Text Annotation

The essays produced by the liberal arts and science students were rated by two raters, both of whom were senior English teachers from key high schools in western China, using the official rating scale for the continuation writing task (see Appendix B). The scale consists of seven levels, with Level 7 as the highest and Level 1 as the lowest. No essay was rated at Level 1 or Level 2. Accordingly, essays at

Levels 7 and 6 were grouped as high-rated essays, whereas essays at Levels 4 and 3 were grouped as low-rated essays. Table 4 presents the distribution of high-rated and low-rated essays across the two topics.

Table 4. High-rated and Low-rated Essays by Topic

Topic	Rating Level	No. of essays
Homesickness	High-rated	36
Homesickness	Low-rated	11
Alligator Race	High-rated	44
Alligator Race	Low-rated	10

Based on the attitude system within the Appraisal framework (Martin & White, 2005), 20% of the texts were pre-annotated by the author and a second trained rater. Interrater reliability was assessed using the intraclass correlation coefficient (ICC). An ICC of over 0.84 indicates a high level of interrater consistency, and the small number of disagreements in affect and invoked resources were resolved through discussion. On this basis, the remaining 80% of the corpus was then annotated.

2.3 Statistical Analysis

To ensure equitable cross-corpus comparisons, the raw frequency of each annotated attitude feature was first normalized to a frequency per 1,000 words for each essay. The analytical procedure was adapted from Chung (2024), with modifications to suit the current research context. Conditional inference trees (CITs) were constructed in R using the partykit package to explore variation in the use of attitude resources across three dimensions: disciplinary background (liberal arts vs. science), topic, and essay rating (high-rated vs. low-rated). CITs are a non-parametric multivariate approach that do not require distributional assumptions and yields easily interpretable results. They follow a recursive data-partitioning procedure aimed at optimizing predictive accuracy (Baayen et al., 2017). For multivariate model selection, the glmulti package was employed (Calcagno & de Mazancourt, 2010; Calcagno, 2020). For each normalized frequency variable, a linear regression model was initially fitted with disciplinary background, topic, essay rating, and all pairwise interactions as candidate predictors. Candidate models were compared based on the corrected Akaike information criterion (AICc), and evidence weights were calculated to quantify the relative support for each model. The best-fitting model for each outcome was identified by the lowest AICc value. Because the data exhibited heteroscedasticity, HC3 heteroscedasticity-consistent standard errors were used for significance testing, as recommended by Long and Ervin (2000). All regression results reported in the tables are based on these robust standard errors, with 95% confidence intervals computed accordingly.

3. Result

Table 5 presents the descriptive statistics of attitude resources across disciplinary background (liberal arts vs. science), topic (Homesickness vs. Alligator Race), and essay rating (high-rated vs. low-rated). The following sections report the effects of disciplinary background, topic, and essay rating on the three categories of the attitude system, namely affect, judgment, and appreciation.

Table 5. Descriptive Statistics of Attitude Resources by Group (M $\pm$ SD, per 1,000 words)

Group	n	Affect	judgment	Appreciation	Inscribed	Invoked	Affording	Flagging	Provoking
Science	47	33.95 $\pm$ 12.17	34.08 $\pm$ 15.71	21.38 $\pm$ 12.31	35.51 $\pm$ 5.28	52.62 $\pm$ 13.00	17.99 $\pm$ 8.99	14.34 $\pm$ 6.38	20.10 $\pm$ 7.98
Liberal arts	52	33.04 $\pm$ 9.04	32.36 $\pm$ 19.98	17.39 $\pm$ 7.49	42.18 $\pm$ 20.21	40.50 $\pm$ 14.12	11.08 $\pm$ 8.08	13.32 $\pm$ 9.04	16.10 $\pm$ 9.79
Alligator Race	52	30.33 $\pm$ 7.79	46.34 $\pm$ 12.80	15.90 $\pm$ 7.08	41.30 $\pm$ 20.88	50.21 $\pm$ 14.98	16.00 $\pm$ 10.84	15.54 $\pm$ 7.75	18.58 $\pm$ 9.07
Homesickness	47	36.95 $\pm$ 12.18	18.61 $\pm$ 9.84	23.02 $\pm$ 11.81	36.49 $\pm$ 14.65	41.88 $\pm$ 13.51	12.54 $\pm$ 6.49	11.88 $\pm$ 7.62	17.36 $\pm$ 9.30
Low-rated	21	29.91 $\pm$ 8.63	32.99 $\pm$ 20.51	18.17 $\pm$ 8.90	39.72 $\pm$ 14.39	41.11 $\pm$ 13.08	12.98 $\pm$ 7.03	12.45 $\pm$ 8.40	15.67 $\pm$ 8.06
High-rated	78	34.43 $\pm$ 10.91	33.2 $\pm$ 17.42	19.58 $\pm$ 10.57	38.82 $\pm$ 19.24	47.64 $\pm$ 15.05	14.73 $\pm$ 9.66	14.17 $\pm$ 7.73	18.63 $\pm$ 9.37

3.1 Affect

A CIT analysis showed that topic was a significant predictor of affect. As shown in Fig. 1, the samples were split into two terminal nodes, with the Homesickness texts showing a higher frequency of affect than the Alligator Race texts.

Model selection using glmulti further identified a model including topic and essay rating as the best-fitting model (AICc = 741.34, evidence weight = 0.36). The regression results (Table 6) based on heteroscedasticity-consistent standard errors showed that topic had a significant effect on affect frequency ($\beta = 6.83$, $p = .002$), while essay rating also reached significance ($\beta = -4.95$, $p = .023$), although its effect was weaker than that of topic. The adjusted R^2 was 0.117. Discipline background was not retained in the top model and had negligible relative importance.

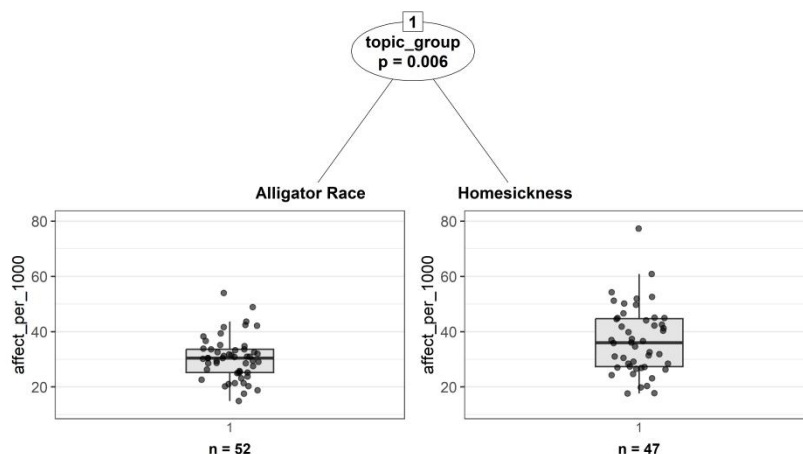


Figure 1. CIT Analysis of the Impact of Topic on Use of Affect

Table 6. Results of Regression Analysis on Factors Affecting the Use of Affect

Predictors	Estimates(β)	95% CI	p
(Intercept)	31.28	[28.96, 33.61]	<0.001
topic[Homesickness]	6.83	[2.67, 10.98]	0.002
score[Low-rated]	-4.95	[-9.22, -0.69]	0.023

A qualitative comparison of two low-rated essays (A.2 and A.5) reveals that the Homesickness text deploys notably more affect resources than the Alligator text. In A.2, positive affect resources such as “excitement,” “enthusiasm,” “love,” “joy,” and “contentment” dominate the narrative and are sustained throughout. Only a few negative tokens (e.g., tears and grief at departure) appear, and these serve to reinforce the value of family bonds rather than undermine the overall positive tone. By contrast, A.5 exhibits a mixed affective profile. Negative emotions, including “racing” heart, “nervousness”, “guilt”, prevail in the first half, while positive affects, including “relief”, “joy”, “exhilaration”, “belonging”, emerge only after the obstacle is overcome. In addition to this cross-topic difference, a further comparison within the same topic, two Homesickness essays (A.1, high-rated, and A.2, low-rated), confirms that the low-rated text deploys fewer affect resources. Although the raw counts are not drastically different, A.1 demonstrates greater lexical diversity and intensity, featuring nuanced affective terms such as “awe”, “content”, “anticipation”, “blissed out”, and “bonding”, whereas A.2 relies on more basic and repeated items like “excitement”, “enthusiasm”, “tears”, and “love”.

3.2 Judgment

A CIT analysis showed that topic was a significant predictor of judgment. As shown in Fig. 2, the samples were divided into two terminal nodes, with the Alligator Race texts showing a higher frequency of judgment than the Homesickness texts.

The best-fitting model included discipline background, topic, essay rating, and the interactions of discipline with topic and rating ($AICc = 767.93$, evidence weight = 0.18). The regression analysis

(Table 7) based on heteroscedasticity-consistent standard errors showed that topic was the strongest predictor ($\beta = -31.52$, $p < .001$). The discipline $\times$ rating interaction was also significant ($\beta = 12.30$, $p = .023$), indicating that the effect of rating on judgment differed across disciplines. The discipline $\times$ topic interaction was marginally significant ($\beta = 8.55$, $p = .065$). The adjusted R^2 was 0.615.

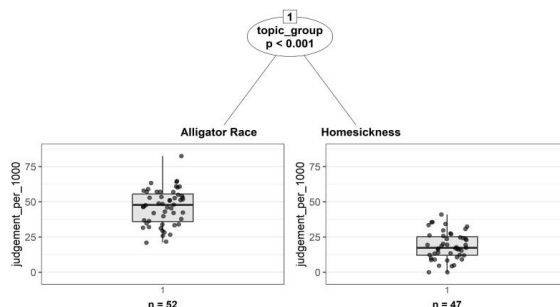


Figure 2. CIT Analysis of the Impact of Topic on Use of Judgment

Table 7. Results of Regression Analysis on Factors Affecting the Use of Judgment

Predictors	Estimates(β)	95% CI	p
(Intercept)	48.32	[42.48, 54.16]	<0.001
Discipline[Science]	-5.17	[-12.74, 2.29]	0.178
Topic[Homesickness]	-31.52	[-38.23, -24.81]	< 0.001
Score[Low-rated]	-3.50	[-10.28, 3.28]	0.308
Discipline[Science] $\times$ topic[Homesickness]	8.55	[-0.54, 17.65]	0.065
Discipline[Science] $\times$ score[Low-rated]	12.30	[1.71, 22.89]	0.023

Qualitatively, the two low-rated essays (A.2 and A.5) reveal a clear disparity in the deployment of judgment resources. The Alligator text (A.5) contains substantially more judgment than the Homesickness text (A.2). In A.2, judgment is scarce, with only about two explicit instances, both of which are rather abstract and general. For example, the father states that visiting grandparents is necessary regardless of grades, and the narrator concludes by reflecting on family qualities such as love and responsibility. These judgments are primarily related to propriety. By contrast, A.5 exhibits a dense cluster of judgment expressions, which are mostly Social Esteem. Most prominently, capacity judgments dominate, as seen in the narrator's self-assessment such as "I did all my might", self-encouragement such as "You can make it!", and external evaluations from teammates such as "You really did a good job!". The Judgment subcategory of Tenacity is also evident in the emphasis on gradual progress and the final moral that "trying your best" truly matters. Furthermore, judgment resources in A.5 are frequently intensified by graduation resources, such as "all my might," "quickly

yet strongly,” and “really,” which amplify the strength of the capacity evaluations, whereas A.2 lacks such intensification for its judgment instances.

3.3 Appreciation

A CIT analysis showed that appreciation was influenced by both topic and disciplinary background. As shown in Fig. 3, the first split was by topic ($p < .001$), with the Homesickness texts showing a higher frequency of appreciation than the Alligator Race texts. The samples were then further split by disciplinary background within each topic. In the Homesickness group, science students used appreciation more frequently than liberal arts students, different from the Alligator group. Overall, topic appeared to be the primary factor affecting the distribution of appreciation, while disciplinary background showed a secondary effect.

Model selection results identified a model including discipline background, topic, and their interaction ($AICc = 689.17$, evidence weight = 0.59). A significant discipline $\times$ topic interaction ($\beta = 22.35$, $p < .001$) indicated that the effect of topic on appreciation depended entirely on the writer’s disciplinary background (Table 8). The adjusted R^2 was 0.446.

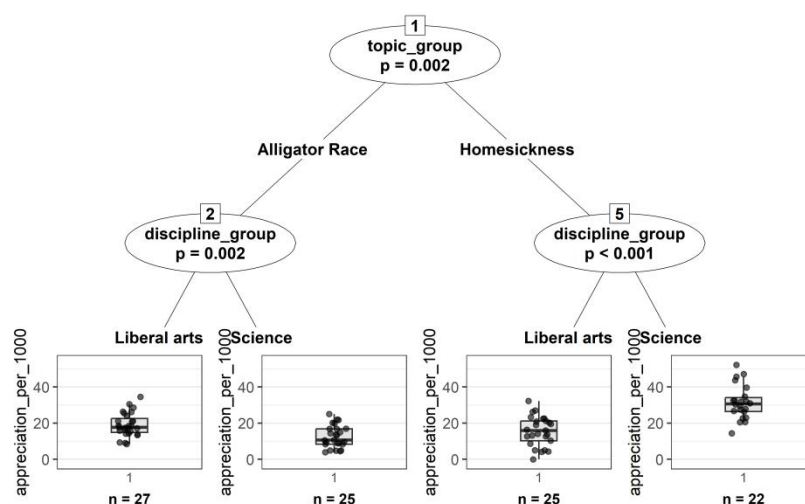


Figure 3. CIT Analysis of the Impact of Discipline and Topic on Use of Appreciation

Table 8. Results of Regression Analysis on Factors Affecting the Use of Appreciation

Predictors	Estimates(β)	95% CI	p
(Intercept)	19.03	[16.51, 21.56]	<0.001
Discipline[Science]	-6.51	[-10.06, -2.96]	<0.001
Topic[Homesickness]	-3.43	[-7.60, 0.75]	0.107
Discipline[Science] $\times$ topic[Homesickness]	22.35	[15.99, 28.70]	<0.001

A.2 and A.3, written by a liberal arts student and a science student respectively, reveal that the science student’s text deploys noticeably more appreciation resources than the liberal arts student’s text. In A.2,

appreciation tokens are relatively scarce, approximately two to three instances, and are primarily limited to reaction, as seen in the phrase dramatic purple-pink sunset, and valuation, such as most valuable qualities and precious love, with no evidence of composition. By contrast, A.3 contains a denser and more diverse set of appreciation expressions, about five instances, covering all three subcategories: reaction through expressions like “delicate flowers”, “lovely birds”, and “lively creatures”; composition through the description of a “big multi-story house”; and valuation through the appreciation of “best” time. However, this pattern reverses in the Alligator topic. Here, the liberal arts student’s text (A.5) employs more appreciation resources than its science counterpart (A.4). A.5 contains a richer array of reaction expressions, such as the metaphor “not the result but trying your best was what truly matters”, and valuation expressions, such as “not the result but trying your best was what truly matters”, whereas A.4 relies on simpler evaluations such as “first rank” and “confidence is crucial”.

3.4 Inscribed and Invoked Attitude

Attitude meanings in the texts were mainly realized in two ways: inscribed attitude and invoked attitude. Invoked attitude could be further realized through affording, flagging, and provoking. A CIT analysis did not identify any single significant splitting variable for inscribed attitude. By contrast, invoked attitude showed a clearer distributional pattern, and the following analysis therefore focuses on this mode of realization.

A CIT analysis showed that the overall frequency of invoked attitude was jointly influenced by disciplinary background and topic. As shown in Fig. 4, the first split was by disciplinary background ($p < .001$), with science students showing a higher frequency of invoked attitude than liberal arts students. The science group was then further split by topic, with Homesickness texts showing a lower frequency of invoked attitude than Alligator Race texts.

Model selection results ($AICc = 783.52$, evidence weight = 0.42) identified discipline background, topic, essay rating, and the discipline $\times$ topic interaction as the best-fitting model. Regression analysis (Table 9) based on heteroscedasticity-consistent standard errors showed that a significant interaction ($\beta = -17.56$, $p < .001$) indicated the effect of topic on invoked attitude depended entirely on the writer’s disciplinary background. Specifically, topic had no effect for non-science students ($\beta = 0.39$, $p = .923$), whereas for science students, the Homesickness topic elicited significantly fewer invoked attitude than the Alligator Race topic. In addition, low-rated essays contained fewer invoked attitude than high-rated ones ($\beta = -6.08$, $p = .025$). The adjusted R^2 was 0.244.

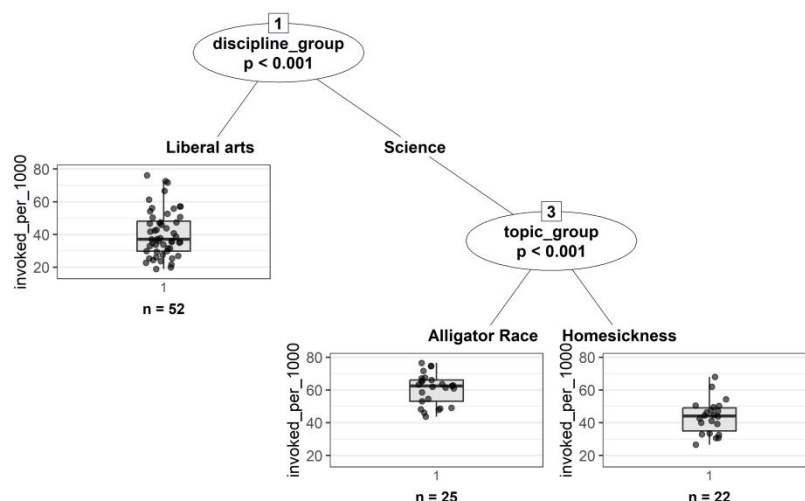


Figure 4. CIT Analysis of the Impact of Discipline and Topic on Use of Invoked Attitude

Table 9. Results of Regression Analysis on Factors Affecting the Use of Invoked Attitude

Predictors	Estimates(β)	95% CI	p
(Intercept)	41.72	[36.62, 46.81]	<0.001
Discipline[Science]	20.11	[13.90, 26.31]	<0.001
Topic[Homesickness]	0.39	[-7.62, 8.41]	0.923
Score[Low-rated]	-6.08	[-11.37, -0.80]	0.025
Discipline[Science] $\times$ topic[Homesickness]	-17.56	[-27.51, -7.62]	< 0.001

Qualitatively, in the Science group texts (A.3 and A.4), the student used invoked resources in writing on Homesickness, including “No sooner had I...”, “rushed”, and “snuggled” in writing on homesickness. And more invoked resources appeared in the Alligator Race texts, including “can’t reach the bottom absolutely”, “a whirlpool of anxiety”, “a shivering tone”, “put his hands on my back”, “tried our best to swim”, “the swim GOAT” and “won the first rank”. However, in Liberal Arts group, students used fewer invoked attitude, such as “twinkled with excitement”, “sprang to my feet” in A.2 and “My heart racing rapidly”, “leaving other lifeguards behind” and “patted my shoulder” in A.4. Across the texts, we observed a recurrent rhetorical move: writers often used an inscribed attitudinal opener to establish the emotional key of a paragraph.

3.4.1 Affording

To further examine the internal composition of invoked attitude, this study analyzed its three sub-strategies separately. A CIT analysis showed that affording was the only invoked strategy significantly split by the variables ($p < .001$), forming a relatively clear decision path (Fig. 5). More specifically, the first split was by disciplinary background. Within the liberal arts group, topic emerged

as a further significant splitting variable, with Homesickness showing a higher frequency of affording than Alligator Race. Within the science group, topic was also significant, and within the Alligator Race texts, essay rating further split the data, with high-rated essays showing a higher frequency of affording than low-rated essays.

Model selection results showed that the best-fitting model for affording included discipline background, topic, and their interaction ($AICc = 657.35$, evidence weight = 0.37). As is shown in Table 10, a significant interaction ($\beta = -21.29$, $p < .001$) revealed that the effect of topic depended entirely on the writer's disciplinary background. Specifically, for non-science students, the Homesickness topic elicited significantly more affording evaluations than the Alligator Race topic ($\beta = 6.71$, $p = .002$), whereas for science students, the pattern was reversed, with Homesickness eliciting fewer affording evaluations. The adjusted R^2 was 0.501.

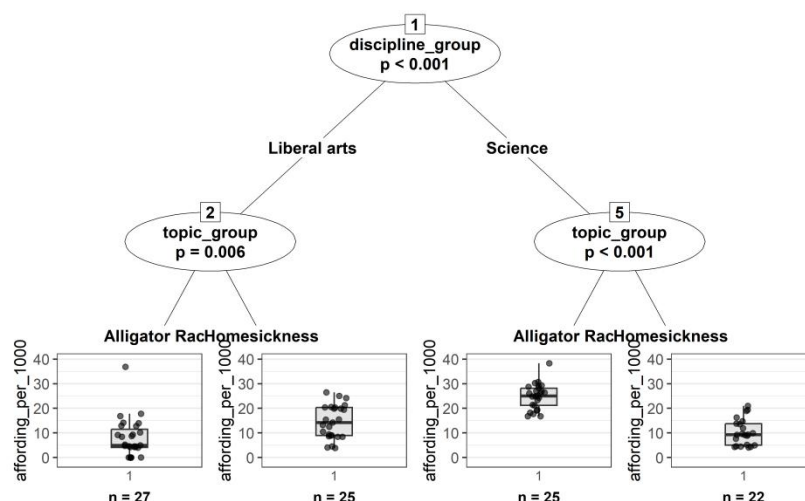


Figure 5. CIT Analysis of the Impact of Discipline, Topic and Rating on Use of Affording

Table 10. Results of Regression Analysis on Factors Affecting the Use of Affording

Predictors	Estimates(β)	95% CI	p
(Intercept)	7.85	[4.77, 10.93]	<0.001
Discipline[Science]	16.96	[13.24, 20.68]	<0.001
Topic[Homesickness]	6.71	[2.57, 10.86]	0.002
Discipline[Science] $\times$ topic[Homesickness]	-21.29	[-26.47, -16.11]	<0.001

No significant predictors were identified for flagging or provoking. In A.4, expressions such as “can’t reach the bottom absolutely”, “try our best to swim”, “put his hands on my back”, and “won the first rank” afford meanings of fear, effort, support and achievement through the description of actions. And in A.3, expressions such as “shared stories and had dinner together” and “promise to be regular visitors

to grandparents” implicitly convey affect and judgment. As for liberal arts group, in A.5, they use “my heart racing rapidly”, and “leaving other lifeguards for behind” to implicitly express nervousness and excitement, as well as their judgment toward others. And in A.1, this student use “answered by a nod and a hug” to afford his attitude.

4. Discussion

4.1 How do Chinese Grade 12 Liberal Arts and Science Students Use Attitude Resources in English Continuation Writing Task?

Overall, both Grade 12 science students and liberal arts students relied heavily on affect and judgment in realizing attitude meanings in the continuation writing task. The descriptive statistics showed that the mean frequencies of affect and judgment were broadly similar across the two groups: 33.95 and 34.08 per 1,000 words for science students, and 33.04 and 32.36 for liberal arts students. This suggests that, regardless of disciplinary background, students in this narrative-like task tended to draw on affect and judgment to organize characters’ emotions, behavior, and event meanings. However, previous research has shown that affect and judgment are less dominant in argumentative texts, where appreciation often was used more in expressing stance (Allami et al., 2024; Chung & Crosthwaite, 2024; Hashemi & Mahdavi-rad, 2023; Hood, 2010; Valerdi Zárate, 2025). For example, Hood (2010) examined the introduction part of four published research articles with approximately 650 to 1,200 words, which found that appreciation accounted for 80% to 100% of explicit Attitude resources, whereas affect and judgment appeared only occasionally. When comparing with narrative studies, we gain lots of similar findings. As for the study of Fitriati et al. (2018), our result is partly consistent with it, as he reported that affect was the most frequently used attitude resource in Indonesian EFL students’ narrative writing and judgment resources are less than affect. In the current dataset, judgment occurred at a level similar to affect, while Fitriati et al. observed a much smaller proportion of judgment compared with affect and appreciation. The difference is likely related to the different topics examined in the two studies. To be specific, Fitriati et al.’s participants produced narratives based on folklore, different from real-life story. The true similar pattern is the studies of Pasaribu (2020) and Magfiroh, Herdiawan, and Rofi’i (2021). Both studies found that affect appeared most frequently in narrative writing, followed by judgment, while appreciation occurred least often. Theoretically, the core rhetorical task of narrative is to foster reader empathy and ethical discernment, with affect and judgment functioning as the primary (Macken-Horarik, 2003).

A further difference between the two student groups appeared in the realization of attitude resources. Science students employed more invoked attitude than liberal arts students, particularly affording. The difference may reflect processes of language socialization (Ochs & Schieffelin, 1984; Schieffelin & Ochs, 1986). From this perspective, language development occurs within social and cultural practices. When students learn languages, they will be influenced by the target people, culture and the like. Therefore, the distinction between inscribed and invoked attitude should not be viewed as an inherent

feature of learners. Instead, it may reflect the influence of proficiency, classroom experience, and broader educational environments. That is to say that the higher use of invoked attitude among science students may relate to differences in disciplinary experiences and learning orientations. Such experiences may encourage a more indirect style of evaluation compared with that found among liberal arts students. From a theoretical perspective of attitude resources, psychological narratives often rely on an interaction between evoked and inscribed appraisal. Evoked evaluation gradually guides readers toward particular interpretations, while inscribed evaluation highlights important moments such as turning points or moments of realization (Macken-Horarik, 2003). Together, these two forms contribute to the persuasive force of narrative discourse.

4.2 To What Extent Does Topic Influence Students' Use of Attitude Resources in English Continuation Writing Tasks?

Topic influences the selection of attitudinal resources in writing (Chung, 2022; Chung et al., 2022; Liu & McCabe, 2018; Zhang & Zhang, 2023). The present study found that Alligator Race produced more frequent use of judgment than Homesickness, whereas affect appeared less. It aligns with Chung and Crosthwaite (2024)' research, who found that different topics in argumentative writing could lead to differences in judgment use. Although the two studies examined different genres, both indicate that judgment may be particularly sensitive to topic variation. Among the three attitude subsystems, judgment seems more closely connected with the situations and events shown in a writing task. As for the explicitness of attitude resources, results showed that science students used more invoked attitude in Alligator Race than in Homesickness. In fact, Homesickness involves a high degree of emotional engagement and more readily encourages students to rely on inscribed resources to encode feelings and values. By contrast, Alligator Race emphasizes process development, bodily sensation, and teamwork, and therefore provides more favorable conditions for realizing attitude through invoked resources embedded in action, state, and outcome. Drawing on the notion of the "cline of instantiation" originally proposed by Halliday and Matthiessen (1999, as cited in Martin & White, 2005), Martin and White (2005) further specified a five-level framework for evaluative meaning. This framework moves from the most abstract system to the most concrete reading, including system, register, text type, instance, and reading, which correspond respectively to appraisal, key, stance, evaluation, and reaction. From the perspective of instantiation, discourses of different topic types can be categorized into different text types, and it is at the level of text type that generalizable patterns of use of evaluative options become manifest (i.e., the stance level).

4.3 How is the use of attitude resources related to essay ratings in English continuation writing tasks?

Affect was found to have a marginally significant relationship with essay rating ($p = .046$), with lower-rated texts containing fewer affect resources. This result differs from the findings of earlier studies. For example, as for affect resources, Chung and Crosthwaite (2024) and Derewianka (2007) reported that lower-rated argumentative essays tended to include more affect resources; as for judgment resources, Derewianka (2007) found that judgment appeared more frequently in lower-rated essays,

whereas Chung and Crosthwaite (2024) reported judgment appeared more frequently in higher-rated writing. They explained this difference in terms of variations in essay topics. The present findings appear closer to the explanation. In the continuation writing task examined here, judgment seems to be influenced more by task content than by essay quality across genres. Next, when we comes to the role of appreciation, it also differs from previous research. For instance, Fauziah, Warsono, and Widhiyanto (2019) observed that higher-rated argumentative essays contained more appreciation resources. This linguistics phenomenon may result from the different communicative purposes of the two genres. Argumentative and academic writing emphasizes clarity, objectivity, and logical persuasion. In such contexts, successful writers may regulate emotional expression and subjective opinions more carefully. By cantrast, continuation writing depends largely on character representation, emotional development, and narrative progression to achieve a meaningful ending. Theoretically, Hogan (2011) emphasizes that emotion is central to narrative understanding and that story structures are shaped by emotional changes. Thus, the role of affective resources in continuation writing may differ from their role in argumentative writing. Apart from the distribution of attitude resources, the present study also discovered that scoring and the realization of attitude resources has some significant relationship. Specifically, higher-rated essays showed a greater tendency to employ invoked attitude, while lower-rated essays used this resource less frequently. This pattern corresponds with Chung and Crosthwaite (2024) and Valerdi Zárate (2025), who both found that writing performance was related to the use of invoked evaluation. Different from inscribed attitude, invoked attitude does not directly encode evaluation. Readers need to recover its evaluative meanings from contextual cues. Therefore, writers must manage and organize linguistic choices more carefully when using invoked resources, so as to allow readers to participate in interpretation. Thus, such use of invoked resources might indicates a more developed interlanguage system.

5. Conclusion

In conclusion, the realization and patterning of attitude resources in the continuation writing task change with topic, disciplinary background and essay grade. More specifically, three main findings are as follows. First, topic was the key factor influencing the distribution of attitude resources and the way attitude was realized. Compared with Homesickness, Alligator Race elicited more judgment, whereas affect and appreciation were relatively less frequent. In addition, the Alligator Race group used more invoked attitude than the Homesickness group. Second, disciplinary background significantly affected appreciation and the inscribed or invoked realization of attitude resources. Specifically, the effect of topic on appreciation depended on the writer's disciplinary background. In terms of realization, compared with liberal arts students, science students used more invoked attitude, especially affording. Third, essay rating was associated with affect and invoked attitude, that is, low-rated essays used fewer affect resources and invoked attitude. The current findings provide one possible explanation for the unclear sense of "feeling". Students with this kind of feeling may have an implicit awareness of how

attitudinal resources contribute to effective narrative writing.

These findings have implications for writing instruction. First and foremost, the observed disciplinary variations in this study suggest that instructors teaching mixed-disciplinary classes might supplement instruction with discipline-specific examples and contrastive analyses. Besides, it shows that topic strongly affects how students mobilize different attitude resources, which means that instruction should help students adjust their evaluative language according to topic and narrative purpose, especially affording. Finally, it extends the application of the Appraisal framework from the more frequently studied field of argumentative and academic writing to a less studied narrative-like school writing task and further shows the distribution of attitude resources varies across genres, which stresses the importance of a genre-sensitive approach to writing instruction. In line with this view, Mochizuki's (2025) findings reveal students' needs in both genre knowledge and genre-based language use.

However, there are some limitations. For one thing, the manual identification of appraisal resources involved subjective decisions. As Tian and Wang (2019) note, lower-level attitude categories are relatively stable because they are closely related to the nature of the evaluated object, whereas more detailed subcategories become increasingly dependent on contextual interpretation. While the three broad categories of affect, judgment, and appreciation could be identified with reasonable consistency, finer distinctions required greater attention to contextual meaning. For another, some methodological limitations remain in the statistical analysis. Although the use of HC3 robust standard errors reduced the possible influence of heteroscedasticity on significance testing, model selection was conducted within a linear regression framework, which assumes normality. But, given the moderate sample size and categorical predictors, this is unlikely to substantially affect model ranking, but future research could consider Tweedie regression.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the authors used DeepSeek Academic (DeepSeek AI) in order to assist with generating R code for statistical analysis and data visualization. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Data availability

The data that have been used are confidential.

References

- Becker, L. J., & Seligman, C. (1981). Welcome to the energy crisis. *Journal of Social Issues*, 37(2), 1-7.
- Hamner, H. A. (in press). *Research in cognitive development*. New York: Springer.
- Hilts, P. J. (1999, February 16). In forecasting their emotions, most people flunk out. *New York Times*. Retrieved November 21, 2000, from <http://www.nytimes.com>
- Laplace, P. S. (1951). *A philosophical essay on probabilities* (F. W. Truscott & F. L. Emory, Trans.). New York: Dover. (Original work published 1814)

- Publication Manual of the American Psychological Association* (5th ed.). (2001). Washington, DC: American Psychological Association.
- Rocklin, T. R. (1989). *Individual differences in item selection in computerized self-adapted testing*. Paper presented at the Annual Meeting of the American Research Association, San Francisco, CA.
- Ryerson, J. F. (1983). *Effective management training: Two models*. Unpublished master's thesis. College of Technology, Potsdam, New York.
- Schwartz, J. (1993, September 30). Obesity affects economic, social status. *The Washington Post*, pp. A1, A4.
- Allami, H. , Karlsson, M. , & Shahroosvand, H. R.. (2024). Functions of idioms in english as lingua franca: an appraisal system account. *International Journal of Applied Linguistics*, 34(4). DOI: 10.1111/ijal.12588
- Biber, D., & Finegan, E. (1988). Adverbial stance types in English. *Discourse Processes*, 11(1), 1-34. <https://doi.org/10.1080/01638538809544689>
- Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*, 9(1), 93-124. <https://doi.org/10.1515/text.1.1989.9.1.93>
- Baayen, R. H., Vasishth, S., Kliegl, R., & Bates, D. M. (2017). The cave of shadows: Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language*, 94, 206-234. <https://doi.org/10.1016/j.jml.2016.11.006>
- Baker, P., Gabrielatos, C., Khosravini, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273-306. <https://doi.org/10.1177/0957926508088962>
- Calcagno, V., & de Mazancourt, C. (2010). glmulti: An R package for easy automated model selection with (generalized) linear models. *Journal of Statistical Software*, 34(12), 1-29. <https://doi.org/10.18637/jss.v034.i12>
- Chang, P., and Schleppegrell, M. (2016). Explicit learning of authorial stance-taking by L2 doctoral students. *Journal Of Writing Research*, 8, 49-80. DOI: 10.17239/jowr-2016.08.01.02
- Crosthwaite, P., & Jiang, F.K. (2017). Does EAP affect written L2 academic stance? A longitudinal learner corpus study. *System*, 69, 92-107. <https://doi.org/10.1016/j.system.2017.06.010>
- Calcagno, V. (2020). *glmulti: Model selection and multimodel inference made easy* (Version 1.0.8) [Computer software]. Comprehensive R Archive Network (CRAN). <https://CRAN.R-project.org/package=glmulti>
- Chung, T. T. (2022). Writing with attitude: A learner corpus study of APPRAISAL resources of Vietnamese and Khmer's L2 English writing [Doctoral dissertation, The University of Queensland]. UQ eSpace. <https://doi.org/10.14264/1a5776b>
- Chung, T. T., Bui, L. T., & Crosthwaite, P. (2022). Evaluative stance in Vietnamese and English writing by the same authors: A corpus-informed appraisal study. *Research in Corpus Linguistics*, 10(1),

- 1-30. <https://doi.org/10.32714/ricl.10.01.01>
- Chung, T. T., & Crosthwaite, P. (2024). Stance-taking through APPRAISAL in L1 and L2 english argumentative essays: Insights from vietnamese L2 english. *International Review of Applied Linguistics in Language Teaching*. Advance online publication. <https://doi.org/10.1515/iral-2024-0079>
- Deng, L., Cheng, Y., & Gao, X. (2024). Engagement patterns in research article introductions: A cross-disciplinary study. *System*, 120, 103204. <https://doi.org/10.1016/j.system.2023.103204>
- Derewianka, B. (2007). Using Appraisal framework to track interpersonal development in adolescent academic writing. In A. McCabe, M. O'Donnell, & R. Whittaker (Eds.), *Advances in language and education*(pp. 142-165). Continuum. DOI: 10.5040/9781474212045.ch-007
- Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stance taking in discourse: Subjectivity, evaluation, interaction*(pp. 139-182). John Benjamins. <https://doi.org/10.1075/pbns.164.07du>
- Englebretson, Robert. (2007). Stance taking in discourse: An introduction. In Robert Englebretson (Ed.), *Stance taking in discourse: Subjectivity, evaluation, interaction, Pragmatics & beyond new series*, Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.164.02eng>
- Gao, W., & Xu, B. (2025). Effects of multi-turn xu-based comparative continuation writing on the development of phraseological complexity: A bigram-based longitudinal study. *System*, 132, 103720. <https://doi.org/10.1016/j.system.2025.103720>
- Gillings, M., Mautner, G., & Baker, P. (2023). Corpus-assisted discourse studies. In S. Hunston (Ed.), *Cambridge elements in corpus linguistics*. Cambridge University Press. <https://doi.org/10.1017/9781009168144>
- Hood, S. (2010). *Appraising research: Taking a stance in academic writing*. Sydney: University of Technology Sydney dissertation. DOI: 10.1057/9780230274662
- Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *Text*, 3, 349-382. DOI: 10.1515/text.1.1998.18.3.349
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192. <https://doi.org/10.1177/1461445605050365>
- Hogan, P. C. (2011). *Affective narratology: The emotional structure of stories*. University of Nebraska Press. <https://doi.org/10.2307/j.ctt1df4gnk>
- Humphrey, S., & MacNaught, L. (2016). Functional language instruction and the writing growth of English language learners in the middle years. *TESOL Quarterly*, 50(4), 792-816. <https://doi.org/10.1002/tesq.247>
- Hyland, K., & Jiang, F. K. (2016). Change of attitude? A diachronic study of stance. *Written Communication*, 33(3), 251-274. <https://doi.org/10.1177/0741088316650399>
- Hashemi, A., & Mahdavarad, F. . (2023). A cross-cultural, cross-disciplinary, and cross-gender study on appraisal resources in phd dissertation abstracts: martin & white's (2005) Appraisal framework in

- focus. *Heliyon*, 9(11), 13.
- DOI: 10.1016/j.heliyon.2023.e22074
- Liu, X., & McCabe, A. (2018). Attitudinal evaluation in Chinese university students' English writing: A contrastive perspective. Springer Singapore. <https://doi.org/10.1007/9789811064159>
- Jaffe, A. (2009). *Stance: Sociolinguistic perspectives*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780195331646.003.0008
- Jiang, L., & Chen, J. (2015). Effects of continuation writing on the development of accuracy, complexity and fluency in L2 English writing [Original title: 读后续写对英语写作语言准确性、复杂性和流利性发展的影响]. *Modern Foreign Languages* [Original title: 现代外语], 38(3), 366-375, 438. Retrieved from <https://www.cnki.net/>
- Jou, Y. (2019). Scaffolding L2 writers' metacognitive awareness of voice in article reviews: A case study of SFL-based pedagogy. *Journal of English for Academic Purposes*, 41, 100770. <https://doi.org/10.1016/j.jeap.2019.100770>
- Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, 54(3), 217-224. <https://doi.org/10.1080/00031305.2000.10474549>
- Lam, S. L., & Crosthwaite, P. (2018). APPRAISAL resources in L1 and L2 argumentative essays: A contrastive learner corpus-informed study of evaluative stance. *Journal of Corpora and Discourse Studies*, 1(1), 8-37. <https://doi.org/10.18573/jcads.1>
- Liu, X., & McCabe, A. (2018). Attitudinal evaluation in Chinese university students' English writing: A contrastive perspective. Springer Singapore. <https://doi.org/10.1007/978-981-10-6415-9>
- Li, Y., & Wang, J. (2025). EFL learners' development of multiword sequences and constructions in longitudinal integrated narrative writings and their relationships to essay scores. *International Journal of Applied Linguistics*. Advance online publication. <https://doi.org/10.1111/ijal.12813>
- Macken-Horarik, M. (2003). Appraisal and the special instructiveness of narrative. *Text*, 23(2), 285-312. <https://doi.org/10.1515/text.2003.012>
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan. DOI: 10.1590/S0102-44502007000200008
- Magfiroh, I., Herdiawan, R. D., & Rofi'i, A. (2021). An appraisal analysis of narrative text from the 11th grade English textbook. *Applied Linguistics, Linguistics, and Literature*, 1(1). DOI: 10.26877/allure.v1i1.9270
- Mochizuki, N. (2025). EFL undergraduate students' thinking and awareness in genre-based writing: Conceptualizing rhetorical moves and experimenting with linguistic knowledge. *Journal of Second Language Writing*, 70, 101243. <https://doi.org/10.1016/j.jslw.2025.101243>
- Merriam-Webster. (n.d.). *Stance*. In Merriam-Webster.com dictionary [simple definition]. Retrieved August 15, 2026, from <https://www.merriam-webster.com/simple/stance>
- Fauziah, M., Warsono, W., & Widhiyanto, W. (2019). The comparison of appraisal resources in

- argumentative essays written by students with different proficiency levels. *English Education Journal*, 9(4), 484-491. <https://doi.org/10.15294/eej.v9i4.31980>
- Ochs, E., & Schieffelin, B. B. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. A. Shweder & R.A. LeVine (Eds.), *Culture theory: Essays on mind, self and emotion* (pp. 276-320). New York: Cambridge University Press.
- Partington, A. (2004). Corpora and discourse, a most congruous beast. In A. Partington, J. Morley, & L. Haarman (Eds.), *Corpora and discourse* (pp. 11-20). Peter Lang. <https://hdl.handle.net/11585/9617>
- Peng, H. (2017). An empirical study on writing coherence of English learners. *Journal of PLA University of Foreign Languages*, 40(4), 87-92. Retrieved from <https://www.cnki.net/>
- Sun, L. (2021). A study on the use of metadiscourse from the perspective of pragmatic identity theory: A case study of research articles in international journals of applied linguistics. *Journal of PLA University of Foreign Languages*, 44(1), 79-86. Retrieved from <https://www.cnki.net/>
- Schieffelin, B. B., & Ochs, E. (Eds.). (1986). *Language socialization across cultures*. Cambridge University Press.
- Valerdi Zárate, J. C. (2025). Evaluative language in undergraduate academic writing: Expressions of attitude as sources of text effectiveness in English as a foreign language. *International Review of Applied Linguistics in Language Teaching*, 63(2), 1139-1168. <https://doi.org/10.1515/iral-2023-0103>
- Wang, C. (2016). Learning by continuation. *Modern Foreign Languages*, 39(6), 784-793, 873. Retrieved from <https://www.cnki.net/>
- Wang, C. (2017). From “writing to learn” to “continuing to learn”. *Foreign Language Teaching and Research*, 49(4), 547-556, 639-640. Retrieved from <https://www.cnki.net/>
- Wang, C., & Wang, Q. (2025). A comparative study of learning effects between oral and written continuation tasks. *Foreign Language Teaching and Research*, 57(1), 55-67. <https://doi.org/10.19923/j.cnki.fltr.2025.01.004>
- Wu, Y. (2025). Authorial identity construction through implicit stance-taking in the introduction of research articles. *English for Specific Purposes*, 78, 88-108. <https://doi.org/10.1016/j.esp.2024.12.003>
- Xie, J. (2016). Direct or indirect? Critical or uncritical? Evaluation in chinese english-major MA thesis literature reviews. *Journal of English for Academic Purposes*, 23, 1-15. <https://doi.org/10.1016/j.jeap.2016.05.001>
- Xuan, W. W., Chen, S., & Zhang, Y. (2025). Applying Appraisal framework in EAP writing research: A methodological synthesis. *International Journal of Applied Linguistics*. Advance online publication. <https://doi.org/10.1111/ijal.70047>
- Yan, X., & Li, G. (2016). An investigation of L2 syntactic structure production in continuation writing tasks. *English Teacher*, 16(12), 44-49. Retrieved from <https://www.cnki.net/>

- Zhang, X. (2017). Effects of receptive-productive integration tasks and prior knowledge of component words on L2 collocation development. *System*, 66, 156-167. <https://doi.org/10.1016/j.system.2017.03.019>
- Zhang, L., & Zhang, L. J. (2021). Fostering stance-taking as a sustainable goal in developing EFL students' academic writing skills: Exploring the effects of explicit instruction on academic writing skills and stance deployment. *Sustainability*, 13(8), 4270. <https://doi.org/10.3390/su13084270>
- Zhang, G. (2023). Authorial stance in citations: Variation by writer expertise and research article part-genres. *English for Specific Purposes*, 70, 131-147. <https://doi.org/10.1016/j.esp.2022.12.002>
- Zhou, Z. (2023). Analysis of Donald Trump's inaugural speech based on attitude within Appraisal framework. *Journal of Education and Educational Research*, 4(1), 151-158. <https://doi.org/10.54097/jeer.v4i1.10275>
- Zhang, L., & Zhang, L. J. (2023). Improving EFL students' stance-taking in academic writing with SFL-based instruction: A qualitative inquiry. *Language Teaching Research*. <https://doi.org/10.1177/13621688231164758>
- Zhang, L., Zhang, L. J., & Xu, T. S. (2023). Effects of teacher explicit instruction in stance-taking on students' perceptions of stance and on their academic writing beliefs. *Frontiers in Psychology*, 14, 1140050. <https://doi.org/10.3389/fpsyg.2023.1140050>

Appendix A

These sample student essays are available upon request from the corresponding author.

- A.1. *Essay of a Student from the High-Rated Liberal Arts Homesickness Group*
- A.2. *Essay of a Student from the Low-Rated Liberal Arts Homesickness Group*
- A.3. *Essay of a Student from the Low-Rated Science Homesickness Group*
- A.4. *Essay of a Student from the Low-Rated Science Alligator Group*
- A.5. *Essay of a Student from the Low-Rated Liberal Arts Alligator Group*
- A.6. *Essay of a Student from the High-Rated Science Alligator Group*

Appendix B. Rating Scale for the Continuation Writing Task

Level	Criteria
Level 7 (22-25 points)	The essay develops original, rich, and reasonable content with strong logic. The continuation is complete and highly consistent with the original context. A wide range of appropriate vocabulary and grammatical structures is used. The expression is fluent, with very few language errors, none of which affects understanding. Cohesive devices are used naturally and effectively across and within paragraphs. The overall structure is clear, with good internal connection and coherence.
Level 6 (18-21 points)	The essay develops relatively rich and reasonable content with fairly good logic. The continuation is relatively complete and largely consistent with the original context. A relatively varied and appropriate range of vocabulary and grammatical structures is used. The expression is fairly fluent, with only a few errors that do not affect understanding. Cohesive devices are used relatively effectively. The overall structure is fairly clear and the meaning is relatively coherent.
Level 5 (15-17 points)	The essay develops basically reasonable content with some logic. The continuation is basically complete and relevant to the original context. Vocabulary and grammatical structures are generally appropriate, but the expression lacks variety. There are some errors, but they do not seriously affect understanding. Cohesive devices are used. The overall structure is relatively clear and the meaning is relatively coherent.
Level 4 (11-14 points)	The essay presents a basically complete story, but some plot details are not sufficiently reasonable or logical. The continuation is basically related to the original context. Simple vocabulary and grammatical structures are used. There are some language errors and inappropriate expressions, and some parts may affect understanding. Some cohesion is present. The overall structure is basically clear and the meaning is basically coherent.
Level 3 (6-10 points)	The essay shows some major problems in content and logic. The continuation is incomplete and somewhat disconnected from the original text. Vocabulary is limited and grammatical structures are simple and repetitive. There are many, often basic, errors, which affect understanding.

Level	Criteria
	Cohesive devices are not used effectively. The overall structure is unclear and the meaning lacks coherence
	The essay shows many major problems in content and logic, or partly copies the original text. The continuation is incomplete and largely disconnected from the original context.
Level 2 (1-5 points)	Vocabulary is very limited and grammatical structures are simple. There are many errors, which seriously affect understanding. Cohesive devices are hardly used. The overall structure is unclear and the meaning is incoherent.
Level 1 (0 points)	No response: the writing is too short or illegible to be evaluated; or the content is entirely copied from the original text or completely irrelevant to the task.
