



# Evaluation in Chinese University EFL Students' English Argumentative Writing: An APPRAISAL Study

Xinghua Liu

([liuxinghua@sjtu.edu.cn](mailto:liuxinghua@sjtu.edu.cn))

Shanghai Jiao Tong University, China

---

## Abstract

By drawing the on APPRAISAL Theory (Martin & White, 2005), an analytical framework within Systemic Functional Linguistics (SFL) for interpersonal meanings, this article reports on a case study that investigated the use of evaluative language between the high- and low-rated English argumentative essays by two Chinese university EFL students. The study found that the high-rated essay successfully employed appraisal values to foreground authorial voice and position readers, thus manipulating arguability of utterances and building strong persuasion. The study suggests that more attention should be given to the teaching of EFL/ESL writing from an interpersonal perspective and proposes the necessity of explaining EFL/ESL students' use of evaluative language from both linguistic and socio-cultural perspectives.

---

## 1 Introduction

It is widely acknowledged that writing is seen as a site of interaction between writers and readers. Through written texts, writers construct solidarity and alignment with potential or target readers (Thompson, 2001). This kind of evaluation and interaction has been a well established research area in the academic context, and different scholars, for different purposes, employ different terms to account for phenomena such as *attitude* (Halliday, 1994), *stance* (Biber & Finegan, 1989; Hyland, 1999), *evaluation* (Hunston & Thompson, 2000), *metadiscourse* (Crismore, 1989) and *appraisal* (Martin, 2000).

Though addressed from divergent approaches, previous studies on evaluation and interaction (e.g. Hood, 2004; Hyland, 2002b; Hyland & Anan, 2006; Thompson, 2001) have all shown that the proper management of interpersonal language uses is essential in constructing a critical voice and building persuasion in argumentative writing. In recent years, increasing attention has been given to the study of interpersonal management in ESL students' academic writing (e.g. Lee, 2006, 2008; Wu, 2007, 2008), but little research has been conducted on how EFL university students deal with this challenge. Particularly, few studies have been conducted on evaluative language use in Chinese university EFL students' English writing in a systematic and comprehensive way. This study of Chinese university EFL students' argumentative writing, based on the APPRAISAL Theory (Martin & White, 2005), is to address this paucity of linguistic studies on interpersonal meanings in EFL students' English writing.

## 2 APPRAISAL theory

APPRAISAL Theory (Martin, 2000; Martin & Rose, 2003; Martin & White, 2005) is a recently developed analytical framework within Systemic Functional Linguistics (SFL) (Halliday & Matthiessen, 2004), which aims to describe various ways of linguistic realization of interpersonal meanings in language use. It has three subsystems, namely ATTITUDE, ENGAGEMENT and GRADUATION, and each subsystem has certain subcategories. Figure 1 gives an overall view of the whole framework.

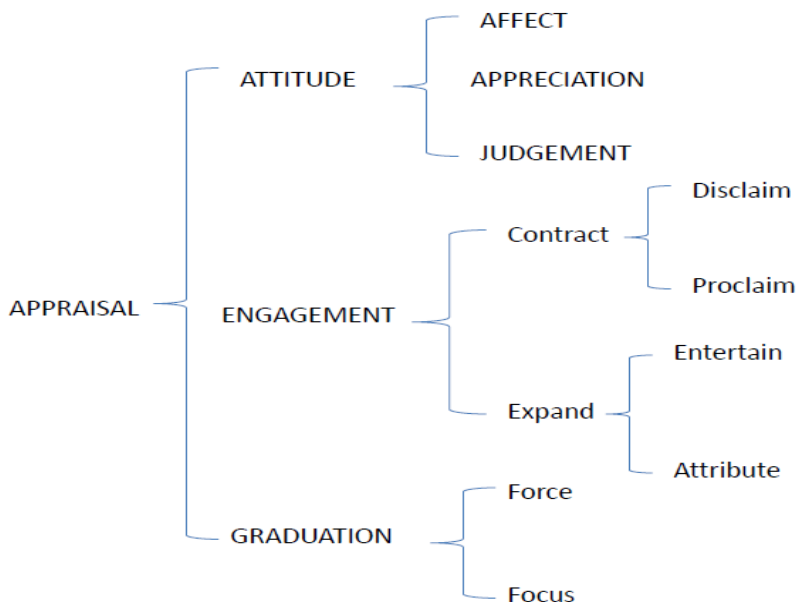


Fig. 1. Overview of APPRAISAL systems (based on Martin & White, 2005)

ATTITUDE is probably the most studied subsystem within APPRAISAL and is the umbrella term for evaluative language in attitudinal positioning in texts (see Fig. 2). It consists of three subsystems: namely AFFECT, which deals with emotional responses, such as *happy*, *frightened*; JUDGEMENT, which refers to the evaluation of human beings and/or their behaviours, such as *capable*, *honest*; APPRECIATION, which concerns the evaluation of entities, such as *complex*, *important*. It is clear from Figure 2 that the three subsystems have their own subcategories.

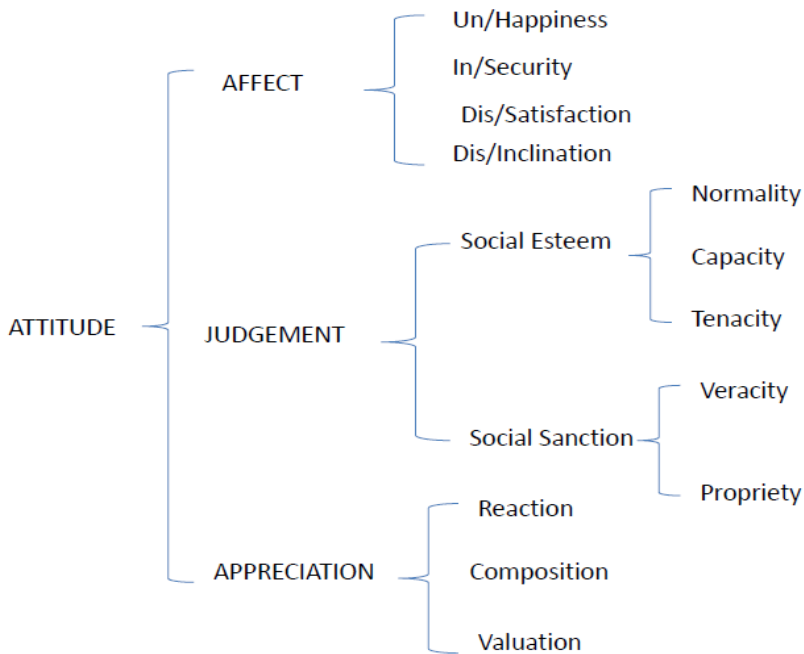


Fig. 2. Overview of ATTITUDE subsystems

Two important features particular to the ATTITUDE system need to be mentioned. First of all, Attitude values can be positive or negative and explicit or implicit/invoked. For example, to be *happy* is a kind of positive affect while to be *afraid* refers to the feeling of insecurity, thus coded as negative. Under the APPRAISAL coding practice, positive and negative values are normally indicated by “+” and “-” respectively. An implicit expression refers to the inscription of Attitude values not through the use of explicit attitude lexis. For example, “he is a good man” is an explicit expression of Judgement, while we cannot easily tell the evaluative meaning encoded in the sentence “he spends 100 pounds per day.” In the latter example, for a thrifty person, it may connote a negative reading of a lavish way of living. An implicit inscription of Attitude values is indicated by “t” in APPRAISAL coding. One point worth noting here is that while coding invoked Attitude values, it is essential for the analyst to state his/her reading position as Martin and White (2005) have pointed out: “When analyzing invoked evaluation it is certainly critical to specify one’s reading position as far as possible.” (p. 62) They classified three types of reading positions: compliant, resistant or tactical reading, and explained them in the following way:

By a tactical reading we refer to a typically partial and interested reading, which aims to deploy a text for social purposes other than those it has naturalized; resistant readings oppose the reading position naturalized by the co-selection of meanings in a text, while compliant readings subscribe to it. (Martin & White, 2005, p. 62)

For the current study, the author takes a compliant reading of students’ texts. As a native Chinese speaker, the author has shared culture and value systems with these EFL students. As an English teacher who has taught English at the tertiary level for many years, the author is aware of the challenges Chinese EFL learners are experiencing and understands the way they are thinking when writing in English.

Another feature is that AFFECT, JUDGEMENT and APPRECIATION constitute an interconnected and interactive system of evaluation. They are all motivated by affectual responses with JUDGEMENT institutionalizing affectual positioning with respect to human behaviours and AP-

PRECIATION institutionalizing affectual positioning with respect to product and process. This feature results in many bordering cases in coding Attitude values (Martin & White, 2005, p. 58).

Besides the above-mentioned ATTITUDE system, APPRAISAL Theory has two other subsystems, that is, ENGAGEMENT and GRADUATION. The ENGAGEMENT system contains a set of resources by which writers adjust and negotiate the arguability of their propositions and proposals, and dialogically engage with potential readers. This intersubjective positioning is informed by Bakhtin's now widely influential notions of *heteroglossia* and *intertextuality* (Bakhtin, Emerson, Holquist, & McGee, 1986; Bakhtin & Holquist, 1981). This diverse range of linguistic resources to realize ENGAGEMENT include various values widely discussed in previous literature under headings such as attribution, modality, polarity, concession, evidentiality, hedging and metadiscourse.

From the perspective of dialogical space between writers and readers, there are two broad subsystems: *Monogloss* and *Heterogloss*. Monoglossic statements are traditionally regarded as facts and bare assertions which ignore dialogical potentials. Heterogloss (see Fig. 3) acknowledges the dialogical divergences by either contracting or expanding the dialogical space with potential readers of the text.

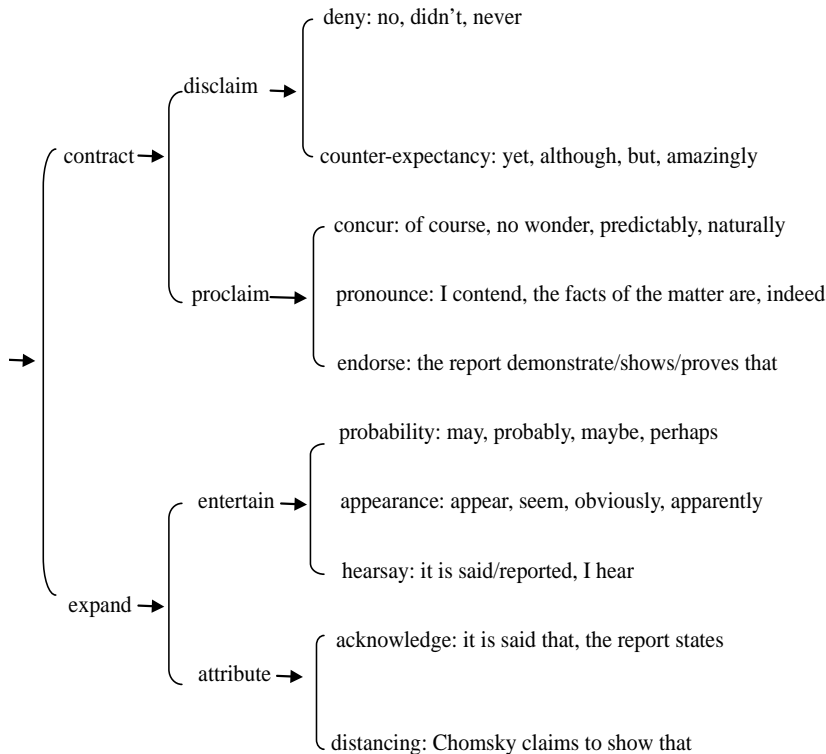


Fig. 3. Heteroglossic system adapted from Martin and White (2005)

The GRADUATION system has two sets of resources: *Force* and *Focus*. Focus refers to resources which indicate to what extent the evaluated item resembles the prototypicality, or is close to the trueness or preciseness of the category boundaries drawn (Martin & White, 2005). Under Focus, values are graded on a scale between “core” and “marginal” membership and the scaling is realized through the semantics of this category membership. It has two subsystems of resources which up-scale or sharpen specifications revealing prototypicality, or down-scale or soften specifications indicating the marginal membership (Martin & Rose, 2003; Martin & White, 2005).

Force refers to the system of resources scaling the intensity of meanings from low to high or “turning the volume up and down” (Martin & Rose, 2003, p. 38). It has two main subcategories, namely *Intensification* and *Quantification*, which indicate the degree of intensity and amount, respectively.

### 3 The study: Student writers and the writing task

This study is taken from a larger project, which aims to examine the writer-reader interaction in the English and Chinese writing by Chinese university EFL students. In the project, over 30 third-year English-majors from an intact class in a Chinese university attended the English writing task on the following topic.

Some people say that the Internet provides people with a lot of information and much convenience. Others think access to so much information creates problems and brings potential troubles. What is your opinion? You are given 40 minutes to write a 250-word argumentative essay with specific reasons and examples to support your opinion.

After establishing satisfactory inter-rater reliability, two Chinese ELT colleagues rated the English scripts independently according to the official writing rating rubric for the Test for English Majors Band 4 (TEM4). The rating rubric enables raters to evaluate the writing holistically instead of focusing on particular linguistic phenomenon. The score for each piece of writing is the average of the two raters' scores. For the purpose of this study, only two English essays, which received the highest and lowest scores, were chosen. Out of the total score of 15, the high-rated essay had a score of 14, while the low-rated essay had a score of 7.

### 4 Contrastive analysis

This section will focus on the comparative analysis of appraisal resources in the high- and low-rated English essays. In the following tables, H stands for the high-rated essay and L for the low-rated one.

#### 4.1 *ATTITUDE analysis*

Table 1 shows that the inscription of attitudinal values in both essays shows a similar overall tendency, that is, among the three subsystems, both essays contained many more Appreciation and Judgement items than Affect ones. This pattern in the use of attitudinal resources is regarded as being characteristic of the argumentative genre (Lee 2006; Liu & Thompson, 2009). The resultant rhetorical effect of the predominance of Appreciation values is to make the text sound more appreciative than emotional or judgemental (Hood, 2004). In other words, less Affect and Judgement mean that there is less disclosure of personal emotion and the avoidance of direct ethical or moral evaluations.

**Table 1. Overview of ATTITUDE subsystem**

	ATTITUDE					
	AFFECT		JUDGEMENT		APPRECIATION	
	H	L	H	L	H	L
Explicit	4	0	5	4	34	19
Implicit	1	0	4	3	1	6
Positive	5	0	7	3	24	15
Negative	0	0	2	4	11	10

One marked difference in the inscription of ATTITUDE between the high- and low-rated essays is that no Affect values were encoded in the low-rated essay. On the contrary, the high-rated essay had three occurrences of Authorial-affect (see Table 2), which indicates that the writer takes responsibility for the attitudinal value assessment.

**Table 2. Affectual values in the high-rated essay**

	In/Happiness	Dis/Satisfaction	In/Security	Dis/Inclination
Explicit	0	3	0	1
Implicit	0	1	0	0
Positive	0	4	0	1
Negative	0	0	0	0
Authorial	0	2	0	1

In the high-rated essay, the emergence of the Internet and the provision of quick and ample information were evaluated as significant (see Example 1), toward which the writer explicitly expressed her satisfaction. The chain of explicit expression of positive appreciation and personal satisfaction helps prosodically to consolidate the thesis in the first paragraph (see Example 2), which also contained an Authorial-affect value, serving to pose a strong personal voice or evaluative stance.

**Example 1**

Taking SARS, earthquake in Sichuan, the H1N1 at present for examples, we are deeply impressed [+affect: satisfaction: authorial] by the interaction cooperation and humanitarian aids which play vital roles in those events.

**Example 2**

As for use, especially judged from my own experience, I'm in favor of its use [+affect: satisfaction: authorial] as long as human have a certain limit or bottom line to it.

In addition, the three Authorial-affect values were realized through behaviour surges or surges of feelings, such as "impress" "in favor of" and "be aware of." Unlike nominalised state of feelings, whose agents are unclear, the evaluated agents in these behaviour surges or surges of feelings are present and foregrounded. This kind of foregrounded affect has the potential to project a clear writer identity (Hyland, 2002a) and to position the potential audience attitudinally.

**Table 3. Overview of APPRECIATION subsystem**

	Reaction		Composition		Valuation	
	H	L	H	L	H	L
Explicit	3	0	2	0	29	19
Implicit	0	0	0	0	1	6
Positive	2	0	1	0	21	15
Negative	1	0	1	0	9	10

Table 3 shows that Appreciation values were predominantly encoded as Valuation in both the high- and low-rated essays (85.7% and 100% respectively). This means that in the two essays, the majority of Attitudinal values were devoted to the explicit evaluation of the significance of things or events. In the high-rated essay (see Example 3), changes brought by Internet and online information were all deemed to be a positive part of human progress and also regarded as contributing to various aspects of human life in the rest of the writing.

#### Example 3

First and foremost, it can't be denied that the computerization [+appreciation: valuation] and the advancement [+appreciation: valuation] of IT industry even the Information Age [+appreciation: valuation] are essential [+appreciation: valuation] and inevitable [+appreciation: valuation] stages the history of human progress must go through as to move forward further.

Similar to the high-rated essay, at the beginning of the low-rated essay (see Example 4), the Internet was evaluated as being "the most important" and the "impetus" for social progress which contributed to building the thesis in a rather objective way.

#### Example 4

Though some people consider it creates problems [-appreciation: valuation] and troubles [-appreciation: valuation], I hold my idea that the huge [force: quantification] benefits [+appreciation: valuation] from the Internet is the most important [+appreciation: valuation]. It's the impetus [+appreciation: valuation] for our era.

Table 4 shows that as far as JUDGEMENT is concerned, both the high- and low-rated essays employed more Social Esteem values (66.7% and 85.7% respectively) than Social Sanction values (33.3% and 14.3% respectively). It means that in both essays, more Social Esteem values were encoded to show the writer's evaluation of people's intellectual capacity and behaviour. Hence, it implies that an ethical and legal judgement of people and/or people's behaviours was not the concern in the two essays.

**Table 4. Overview of JUDGEMENT subsystem**

	Social Esteem				Social Sanction					
	Normality		Capacity		Tenacity		Veracity		Propriety	
	H	L	H	L	H	L	H	L	H	L
Explicit	0	0	2	1	0	2	0	0	3	1
Implicit	0	0	4	2	0	1	0	0	0	0
Positive	0	0	5	2	0	1	0	0	2	0
Negative	0	0	1	1	0	2	0	0	1	1

It is also clear from Table 4 that both the high- and low-rated essays showed the same pattern of distribution of Judgement values, namely the predominance of Capacity (66.7% and 42.9% respectively), which is consistent with previous studies (Lee, 2008; Wu & Allison, 2003). A further examination of the high- and low-rated essays shows a difference in the way Judgement values were encoded in the texts. As Example 5 illustrates, in the high-rated essay, the targets of Judge-

ment evaluation were either semiotic things (“those fields”) or not mentioned at all, while in the low-rated essay (see Example 6), the targets were often explicitly specified as human beings (“people”). What’s more, in the high-rated essay, Judgement values were most often nominalised (“achievements” in Example 5). This nominalized expression of Judgement values without explicit targets is deemed as characterizing successful argumentative academic writing which tends to sound impersonal and thus maintains a certain level of formality (Lee, 2008, p. 50).

Example 5

And thus, the compound achievements [+appreciation :valuation][+judgement: capacity] from all those fields stimulate the speed of civilization as well as the society.

Example 6

People can talk [+judgement: capacity] with each other any time any place with the help of the Internet.

## 4.2 ENGAGEMENT analysis

Table 5 shows the distribution of Engagement items in the high- and low-rated essays. To make the numbers comparable, all occurrences of items have been normalized to an article with 100 words. Overall, after normalization, the low-rated essay contained 14.7 Engagement items, more than two times of those in the high-rated essay which had 7.3. The resultant rhetoric effect is that the low-rated essay sounds more dialogic than the high-rated one. In other words, the low-rated essay was more likely to foreground and promote heteroglossic diversities than the high-rated essay.

As far as the Monogloss system is concerned, Monoglossic values accounted for 19.2% of total Engagement values while in the low-rated one, they only accounted for 4.8%. Since Monoglossic values do not overtly refer to other voices or recognize alternative positions (Martin & White, 2005) compared to the low-rated essay, the high-rated essay foregrounded authorial voice and thus rhetorically sounded more affirmative and authoritative.

However, this pattern is different from that in Wu’s (2007) study, in which the high-rated scripts have fewer occurrences of Monoglossic resources and thus rhetorically stated hypotheses in a dialogically expansive and inviting manner. This difference might be attributed to the different nature of the writing tasks. In Wu’s (2007) study, students wrote academic essays for the specific subject of geography and their writing was favourably evaluated through “crafting claims that neither overstated nor understated by exploring the various options” more in a Heteroglossic system (Wu, 2007, p. 267).

**Table 5. Overview of ENGAGEMENT subsystem**

Mono-gloss		Heterogloss															
		Contract								Expand							
		Disclaim				Proclaim				Attribute		Acknowledge					
		Counter		Deny		Concur		Pronounce		Endorse		Entertain		Acknowledge		Distance	
H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L
5	2	3	7	3	6	9	4	1	4	0	3	5	14	0	1	0	1
1.4	0.7	0.8	2.5	0.8	2.2	2.5	1.4	0.3	1.4	0	1.1	1.4	5.0	0	0.4	0	0.4

*NB: the two rows of numbers in this table are the occurrences of Engagement items before and after normalization respectively.*

In the current study, however, students wrote on a general topic and they were expected to articulate in an explicit manner their viewpoints on issues at risk. Thus, a more affirmative and au-



thoritative statement of opinions through Monoglossic expressions may be more favourably received.

Table 5 also shows some differences in the use of Heteroglossic resources between the two essays. Firstly, it is evident from the table that the low-rated essay had many more Disclaim resources (Counter and Deny) than the high-rated one. Similar to Lee's (2006) findings, this low-rated essay contained many more Deny resources (2.5) and particularly it employed many double negations as illustrated in Example 7.

Example 7

We can't [denial] ignore [denial] it.

Meanwhile, the interplay between Counter, Deny and other Contracting resources such as Pronounce (see Example 8) made the authorial voice too strong, which rendered arguments less persuasive. This frequent use of Disclaiming resources (Counter and Deny) made the low-rated essay carry a negative coloring and thus rhetorically ignored possible refutation from a resistant audience by a strong textual voice, which probably characterized the poor writing (Lee, 2006).

Example 8

Though [disclaim: counter] some people consider it creates problems and troubles, I hold my idea that [proclaim: pronounce] the huge benefits from the Internet is the most important.

Another marked difference is found in the use of Pronounce resources. It is evident from Table 5 that the low-rated essay employed more Pronounce items (1.4) than the high-rated essay (0.3). This phenomenon is different from Lee (2006), which found that high-rated essays relied more on Pronounce to construct authorial voice, but is similar to Wu's (2007) findings that low-rated scripts had a higher frequency of Pronounce resources.

Through Pronounce, the writer may "interpolate themselves directly into the text as the explicitly responsible source of the utterance" and this practice "increases the interpersonal cost of any rejection/doubting of their utterance, rendering such a direct challenge to the author's dialogic position" (White, 1998, p. 89). In the low-rated essay in this study, two Pronounce items (see Examples 9 and 10 below) were realized through the explicit interpolation of the authorial self. The resultant rhetorical effect is to "construe the writer as authoritative in relation to the views put forward," foregrounding the "subjectivity of the argument" (Lee, 2006, p. 304). However, the fact that this essay was rated as low may possibly suggest that the writing with explicit interpolation of authorial self might not be favorably rated. This phenomenon might probably be attributable to the Chinese culture of advocating modesty rather than being explicitly subjective or aggressive (Carson & Nelson, 1994; Matalene, 1985; Shen, 1989).

Example 9

I hold my idea that [proclaim: pronounce] the huge benefits from the Internet is the most important.

Example 10

I can say that [proclaim: pronounce] the world even can't has its daily life any more.

Instead of classifying ENGAGEMENT subsystems as dialogically contractive and expansive resources, they can also be interpreted from the angle of the sources of voices, namely whether the responsibility for intersubjectivity is attributed to external voice (extra-vocalizing) or remains the authorial one (intra-vocalizing). According to White (2005), Heteroglossic subcategories, namely Endorse, Acknowledgement and Distancing are classified as extra-vocalizing resources. Then, one distinct difference could be noticed between the two essays: the low-rated essay had extra-vocalizing resources on all subcategories (Endorse, Acknowledgement and Distancing) while the high-rated essay had none. This is different from Wu (2007), who found that the high-rated scripts had a higher frequency of Endorse values and that these writers showed alignment with external voices to support their claims. Though the employment of extra-vocalizing resources helps intro-

duce external voices to present examples and arguments, and thus make the text sound authoritative and persuasive, this rhetoric effect could not be achieved, if the extra-vocalizing resources do not work harmoniously with surrounding evaluative resources or in the proper textual locations. The low-rated essay has this problem. For example, at the very beginning of the low-rated essay, there were two Endorse values (see Example 11) which serve to align potential readers to their propositions. However, readers were not properly navigated due to the lack of consistency in meanings between the two sentences.

Example 11

As everyone knows [proclaim: endorse], 21st century is an information century. An authoritative data shows [proclaim: endorse] that nearly 81% of the exotic information is transmit by the Internet.

Another occurrence of Endorse item (see Example 12) does not seem to appear in the appropriate position in the essay. In this example, a popular saying was presented, which should function as the topic sentence of the second paragraph, but the actual topic sentence was postponed and did come until in the middle of the paragraph.

Example 12

An old byword says [proclaim: endorse] everything has two sides.

This brief analysis supports Lee's (2006) suggestion that an analysis of Engagement resources needs to take into consideration the co-text in which they occur. In other words, the co-patterning between prosody (the rhetorical effect by evaluative language) and periodicity (information flow) better navigates readers through discourse phasing (Martin & Rose, 2003). A global examination of the low-rated essay shows that though there were more Heteroglossic resources, they did not work harmoniously with surrounding evaluative resources to achieve the co-patterning between prosody and periodicity. Here, two examples from the second and third paragraph in the low-rated essay are taken to further illustrate this point.

The second paragraph started with the elaboration of the advantages of the Internet and online formation (see Example 13), but it was mixed with counter-arguments about the disadvantages at the end of the paragraph which might be better put in the third paragraph. The highly frequent inclusions of Heteroglossic alternatives interwoven with strong authorial voice made the flow of ideas full of countering and negations. The rhetorical effect of the interplay between interrupted periodicity and varying prosody is to make the text sound disharmonious. Readers would feel lost by being navigated by interruptive periodicity and too frequent changes in aligning or dis-aligning with various propositions.

Example 13

If [condition] some people think [attribute: distancing] ... I can say that [pronounce] ... even [counter] can't [entertain][denial] ... So [result] ... In my shoes [entertain] ... By the way, we can't [entertain][denial] ignore [denial] ... The International affairs and even [counter] the Eight Diagrams ... can [entertain]

The same problem existed in the third paragraph (see Example 14). The topic sentence was not introduced right at the beginning but was postponed to come in the middle after the presentation of specific examples. The last part of the paragraph did not see further arguments for the topic but provided irrelevant explanations. The frequent and congruent uses of Heteroglossic resources were co-articulated with the twisted periodicity which failed to produce a harmonious resonance effect.

Example 14

On the other hand [counter], really [concur] ... An old byword says [endorse] ... We can't [probability] [denial] ignore [denial] it . But [counter] ... can be [entertain] ... May be [entertain] ... just [counter] don't [denial] ...

### 4.3 GRADUATION analysis

An examination of Table 6 shows two distinct differences between the low-rated and high-rated essays regarding the distribution of Graduation resources and graduating strategies.

Force				Focus			
Intensification		Quantification		Sharpening		Softening	
H	L	H	L	H	L	H	L
15	8	12	24	6	0	2	1
4.2	2.9	3.4	8.6	1.7	0	0.6	0.4

**Table 6. Overview of GRADUATION subsystem**

First, the high-rated essay had more balanced uses of Graduation resources across all subcategories, while in the low-rated essay, there was a disproportionately high number of Quantification items (72.7%), whereas there was not a single occurrence of Sharpening items.

Second, the high-rated essay's strategic deployment of Graduation resources constituted a dynamic interplay among Graduation resources and between Graduation and Attitude which resulted in a rhetorical prosody resonance. In Example 15, social changes brought by the Internet were positively evaluated as "easy" and "effective" which has been graduated with high-value Intensification or Quantification and Sharpening. In Example 16, the positive Authorial-affect "impressed" was intensified and followed by a positive ethical Judgement:Propriety of human behavior. In this example, both attitudinal evaluation and graduation were attributed to the internal authorial voice, which contributed to aligning readers and enhanced the arguability. In both examples, the resultant rhetorical effect by this harmonious distribution and interplay among Graduation values and between Graduation and Attitude is to enrich prosody and build up persuasion.

#### Example 15

School learnings, scientific studies, business news and specific technologies are much [force: quantification] easier [force: intensification][+appreciation: composition] and more [force:intensification] effective [+appreciation: valuation] then ever [focus: sharpening] before.

#### Example 16

Taking SARS, earthquake in Sichuan, the H1N1 at present for examples, we are deeply [force: intensification] impressed [+affect: satisfaction: authorial] by the interaction cooperation [+judgement: propriety] and humanitarian aids [+judgement: propriety]

On the contrary, in the low-rated essay, the majority of Graduation resources were used either in an isolated manner or to graduate experiential meanings, as shown in Examples 17 and 18. Rhetorically, due to this monotonous distribution of Graduation resources and the lack of interaction between Graduation and Attitude, there was less space for the arguability and thus persuasion was weak.

#### Example 17

People can [entertain] talk with [+judgement: capacity] each other any [force: quantification] time any [force: quantification] place with the help [+appreciation: valuation] of the Internet.

#### Example 18

A part of [force:quantification] them is illegitimate [-appreciation: valuation]. But [counter] other large [force: quantification] part of them can be [entertain] changed by the endeavor [force: intensification][+judgement: tenacity] of us.

## 5 Limitations and suggestions for future studies

The current study has a few limitations and the results of the investigation are better interpreted as a prelude to large-scale studies of significant areas. Firstly, future studies based on a larger corpus of students' texts are needed to verify the above-mentioned findings and generate a new understanding of evaluative language use in EFL students' English writing. Secondly, a cross-cultural comparison of the use of evaluative language – such as the examination of appraisal values in the English and Chinese writings by Chinese EFL students (e.g. Liu & Thompson, 2009) – might also provide interesting and useful results. Finally, as L1 literacy and the teaching of L2 writing play an important role in forming EFL/ESL students' L1/L2 writing practice (Kubota, 1998; Liebman, 1992), it would be also revealing to investigate the relation between “small culture” (Connor, 2004) or the culture of learning (Cortazzi & Jin, 1996) and EFL/ESL student writers' textual features. In all, an Appraisal analysis of a larger cohort of students' writings from multiple perspectives is needed to better reveal EFL/ESL students' use of evaluative language in their L1/L2 writings.

## 6 Conclusion

This study applied the complete appraisal framework in Chinese university EFL students' English argumentative writing. After analyzing the high- and low-rated English essays, several patterns of differences and similarities in the use of appraisal resources were found. Firstly, the investigation found that within the ATTITUDE subsystem, both the high- and low-rated essays had predominantly more Appreciation resources than Judgement and Affect ones, which made their writing sound more appreciative than personal and emotional. This might be related to the nature of the topic, namely the discussion of the Internet and online information instead of human behaviour. Therefore, more studies are needed to examine the interaction between different topics and the employment of attitudinal resources. Secondly, unlike the low-rated essay which had no Affect values, the high-rated essay had successfully employed the Affect values through behaviour surges or surges of feelings to position readers attitudinally. Thirdly, though both essays had a similar tendency to use more Capacity values, the high-rated essay displayed formality through nominalised Judgement values and implicit targets of evaluation. The fourth important finding is that in the high-rated essay, there was a lower frequency of Engagement occurrences but a higher frequency of Monoglossic resources. The resultant rhetoric effect makes the text sound less dialogically expansive, but rather affirmative and authoritative. Finally, the study found that the high-rated essay had a more balanced use of Graduation resources and more importantly deployed them strategically to constitute a dynamic interplay between Graduation and Attitude values.

The important pedagogical implication gained from the above findings for English writing instruction in EFL/ESL context is that instead of primarily focusing on the correctness of grammar use, EFL/ESL writing instruction and learning needs to pay more attention to evaluative meanings conveyed through linguistic constructions. In SFL terms, apart from ideational and textual metafunctions, the interpersonal metafunction needs to be taken into consideration in EFL/ESL writing pedagogy. EFL/ESL students find argumentative writing difficult and challenging (Connor & Lauer, 1988; Hirose, 2003; Lee, 2006) and are often blamed for a lack of critical stance and personal voice (Johns, 1997). Just as this study shows that a proper use of evaluative language helps establish personal voice and position readers, thus achieving critical arguments, EFL students need to be given the practice of analyzing their writing from an interactional perspective and training in the use of proper evaluative constructions.

## References

- Bakhtin, M., Emerson, C., Holquist, M., & McGee, V. (1986). *Speech genres and other late essays*. Austin: University of Texas Press.

- Bakhtin, M., & Holquist, M. (1981). *The dialogic imagination: Four essays*. Austin: University of Texas Press.
- Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*, 9, 93–124.
- Carson, J. G., & Nelson, G. L. (1994). Writing groups: Cross-cultural issues. *Journal of Second Language Writing*, 3(1), 17–30.
- Connor, U. (2004). Intercultural rhetoric research: Beyond texts. *Journal of English for Academic Purposes*, 3(4), 291–304.
- Connor, U., & Lauer, J. (1988). Cross-cultural variation in persuasive student writing. In A. C. Purves (Ed.), *Writing across languages and cultures: Issues in contrastive rhetoric* (pp. 138–159). Newbury Park, CA: London Sage.
- Cortazzi, M., & Jin, L. (1996). Cultures of learning: Language classrooms in China. In H. Coleman (Ed.), *Society and the language classroom* (pp. 169–206). Cambridge: Cambridge University Press.
- Crismore, A. (1989). *Talking with readers: Meta-discourse as rhetorical act*. New York: Peter Lang.
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed.). London: Arnold.
- Hirose, K. (2003). Comparing L1 and L2 organizational patterns in the argumentative writing of Japanese EFL students. *Journal of Second Language Writing*, 12(2), 181–209.
- Hood, S. (2004). Appraising research: *Taking a stance in academic writing* (Unpublished Ph.D. Dissertation). University of Technology, Sydney, Australia.
- Hunston, S., & Thompson, G. (2000). *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press.
- Hyland, K. (1999). Talking to students: Metadiscourse in introductory coursebooks. *English for Specific Purposes*, 18(1), 3–26.
- Hyland, K. (2002a). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34(8), 1091–1112.
- Hyland, K. (2002b). Directives: Argument and engagement in academic writing. *Applied Linguistics*, 23(2), 215–239.
- Hyland, K., & Anan, E. (2006). Teachers' perceptions of error: The effects of first language and experience. *System*, 34(4), 509–519.
- Johns, A. M. (1997). *Text, role, and context: Developing academic literacies*. Cambridge: Cambridge University Press.
- Kubota, R. (1998). An investigation of L1-L2 transfer in writing among Japanese university students: Implications for contrastive rhetoric. *Journal of Second Language Writing*, 7(1), 69–100.
- Lee, S. H. (2006). *The use of interpersonal resources in argumentative/persuasive essays by East-Asian ESL and Australian tertiary students* (Unpublished Ph.D. Dissertation). University of Sydney, Sydney, Australia.
- Lee, S. H. (2008). Attitude in undergraduate persuasive essays. *Prospect*, 23(3), 43–58.
- Liebman, J. D. (1992). Toward a new contrastive rhetoric: Differences between Arabic and Japanese rhetorical instruction. *Journal of Second Language Writing*, 1(2), 141–165.
- Liu, X., & Thompson, P. (2009). Attitude in students' argumentative writing: a contrastive perspective. In L. J. O'Brien & D. S. Giannoni (Eds.), *Language studies working papers* (Vol. 1; pp. 3–15). Reading: University of Reading.
- Martin, J. R. (2000). Beyond exchange: APPRAISAL systems in English. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 142–177). Oxford: Oxford University Press.
- Martin, J. R., & Rose, D. (2003). *Working with discourse: Meaning beyond the clause*. London: Continuum.
- Martin, J. R., & White, P. R. R. (2005). *Language of evaluation: Appraisal in English*. New York: Palgrave Macmillan
- Matalene, C. (1985). Contrastive rhetoric: An American writing teacher in China. *College English*, 47(8), 789–808.
- Shen, F. (1989). The classroom and the wider culture: Identity as a key to learning English composition. *College Composition and Communication*, 40(4), 459–466.
- Thompson, G. (2001). Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics*, 22(1), 58–78.
- Wu, S. M. (2007). The use of engagement resources in high- and low-rated undergraduate geography essays. *Journal of English for Academic Purposes*, 6(3), 254–271.

- Wu, S. M. (2008). Investigating the effectiveness of arguments in undergraduate essays from an evaluation perspective. *Prospect*, 23(3), 59–75.
- Wu, S. M., & Allison, D. (2003). Exploring appraisal in claims of student writers in argumentative essays. *Prospect* 18(3), 71–91.