

Construction of Evaluative Meanings by Kurdish-Speaking Learners of English: A Comparison of High- and Low-Graded Argumentative Essays

Alireza Jalilifar

Associate Professor, Shahid Chamran University of Ahvaz, Iran

Ali Hemmati

Ph.D. Candidate, Shahid Chamran University of Ahvaz, Iran

Received: May 2, 2013; **Accepted:** November 3, 2013

Abstract

Academic writing ability is an important goal that learners of English as a Second Language (ESL) or English as a Foreign Language (EFL) try to attain. While ESL students' academic writings have been widely explored, owing to few studies investigating appraisal resources in EFL students' argumentative writing, the gap still exists about EFL students' academic writing. This study aimed to see how Kurdish-speaking learners of English employ appraisal resources in their writings. It further aimed to explore whether the appraisal framework can be utilized as an assessment scale for evaluating the students' argumentative writing. To this end, the study investigated the argumentative essays written by 15 bilingual Kurdish-Iranian graduates of English within the framework of the appraisal theory. The instruments applied in this study consisted of a modified rating scale for assessing the essays in terms of the macrostructures exploited in them and the framework for the analysis of appraisal resources. Quantitative findings revealed high-graded essays employ more attitudinal items and fewer monoglossic resources than low-graded ones. Qualitatively, the high-graded essays articulated attitudinal values in nominal forms and sometimes in a backgrounded manner while these values were mostly presented by surge of feelings and in a foregrounded way in the low-graded essays. Regarding engagement, unlike the high-graded essays, the low-graded ones were poor in recognizing other voices and alternative positions. Inspired by the strength of the appraisal model evaluating writing, results suggest that high-graded essays are successful in positioning readers attitudinally and clarifying the ethical message to readers.

Keywords: argumentative writing, appraisal theory, high- and low-graded essays, Kurdish

Authors' emails: ar.jalilifar@gmail.com; ali_hemati5363@yahoo.com

INTRODUCTION

As one of the best predictors of academic success (Gaiser & Studely, 2001), academic writing ability is one of the most important goals that learners of English as a Second Language (ESL) or English as a Foreign Language (EFL) try to attain. In addition to conveying disciplinary knowledge (Nordquist, 2011), academic writing “carries a representation of writer” (Hyland, 2002a, p. 1092). In fact, qualified academic writing enhances a writer’s interaction with the potential readers by taking a special voice, exploiting interpersonal meanings and delivering a sound argument so that the readers are persuaded to take a voice as the writer’s. In this way, ESL/EFL learners can “make words say what they want to say” (Kisting, 2007, p. 4).

Among the different rhetorical modes of description, narration, exposition and argumentation (Connor, 1996; Richards & Schmidt, 2002; Schultz, 1991), the latter tends to be applied more in academic writing because the essence of academic writing is argument (Irvin, 2010). However, as academic writing attainment is a crucial and demanding task for ESL/EFL learners, robust and well-grounded evaluation of written texts is equally important and even sometimes a dilemma for instructors. As Weigle (2002, cited in Umair, 2011, p. 230) states, “writing is not only a mirror of one’s thought but it contributes newness to established information”. Thus, since academic writing is an interactive undertaking (Hyland, 2000) and requires sound judgment, it should be evaluated and assessed based on a rigorous linguistic model which cannot only comprehensively scrutinize language itself but also evaluate how writers exploit interpersonal meanings and show their attitude, identity, and voice. However, it seems that the existing analytical scales are deprived of such criteria to evaluate interpersonal meanings in students’ academic writing. Among different analytical scales for assessing students’ writing performance, four scales are widely used which, according to Hawkey and Baker (2004, p. 123), comprise Writing Profile (Jacobs, Zingraf, Wormuth, Hartfiel, & Hughey, 1981), Cambridge Main Suit Examination, International English Language Testing system (IELTS), and Weir’s (1983) Test of English for Educational Purposes (TEEP). Although each of these scales has provided a touchstone for assessing written texts, they are language proficiency tests and they may not explicitly and specifically deal with interpersonal resources that may be exploited in academic writing. Given the use of the modified

Composition Profile (Jacobs et al., 1981) for assessing the essays in this study, we later present a brief description of the profile.

LITERATURE REVIEW

The Composition Profile divides writing into five components with various percentages: content (30%), organization (20%), vocabulary (20%), language (25%), and mechanics (5%). According to this scale, each component has a set of criteria which range from *excellent* to *very poor* with a particular range of scores (Jacobs et al., 1981). Based on this profile, content refers to linguistic features corresponding to the relevance of the essay to the assigned topic and development of the thesis statement. Organization refers to how writers state and buttress their position thoroughly. Vocabulary criteria are germane to word choice, word form mastery, and appropriate register. Language use determines how sentences are constructed and to what extent word order, article, and tense are correctly applied. Finally, mechanical aspects are concerned with punctuation, capitalization, and spelling (Jacobs, et al., 1981). Similarly, according to this scale, the highest possible score for overall proficiency in English composition is 100. Although this scale analytically assesses ESL writing and assigns a separate score for different aspects of writing, it fails to “consider constraints specific to writing related to a given purpose” (Connor & Mbaye, 2002, pp. 272–273). For instance, regarding content, this scale lacks the criteria for “the effectiveness of addressing the audience and appeals used” (Connor & Mbaye, 2002, p. 273). Moreover, in terms of organization, this scale provides little direction regarding clear evaluation of the elements of an argument. That is, the moves that should be followed in an argument are not completely explicated in this scale and, consequently, it may fail to effectively address the macro-structural organization of argumentative essays.

Generally, the important roles of evaluative language and interaction in academic writing have been emphasized by scholars (e.g., Hood, 2004, Hunston, 2000, Hyland, 2002b). Due to the importance of academic writing ability for ESL/EFL students, some studies have specifically investigated students’ argumentative writing (e.g., Lee, 2006, 2008; Liu, 2013; Liu & Thompson, 2009; Wu, 2007, 2008). The studies most pertinent to the present one have generally found a close relationship between the number and types of appraisal resources exploited in essays

and the assessing of the essays as high- and low- graded (Lee, 2006, 2008; Liu, 2013).

Although ESL students' academic writings have more recently been explored and outstanding results have been achieved, owing to few studies (e.g., Liu, 2013; Liu & Thompson, both on Chinese EFL learners), as far as the researchers know, investigating appraisal resources in EFL students' argumentative writing, the gap in the literature still exists in relation to EFL students' academic writing. Particularly, the studies on how Kurdish Iranian EFL learners, whose second language is Persian learning English as a third language (L3), exploit appraisal resources in their argumentative writing are regrettably rare. Furthermore, all the aforementioned studies on ESL/EFL students' academic writing have investigated undergraduate essays. Thus, past research has provided little direction regarding how master's students exploit appraisal resources.

PURPOSE OF THE STUDY

The present study aimed at investigating how Iranian EFL learners (MA students) whose native language is Kurdish and whose second language is Persian, learning English as an L3, employ appraisal resources in high- and low-graded argumentative writing. We assume that the appraisal theory (Martin, 1997, 2000; Martin & Rose 2003; Martin & White, 2005) can play a role in evaluating and assessing argumentative writing. Since this theory is an offshoot of Systemic Functional Linguistics (Halliday, 1994, 2004; Halliday & Matthiessen, 2004) and takes into account delicate interpersonal meanings and evaluative resources, it may present a useful way of assessing students' academic essays and exploring the interpersonal meanings employed in them. To this end, two questions were put forward.

1. In what ways are the main categories of the appraisal theory—Engagement, Attitude, and Graduation and their sub-categories—distributed in the high- and low-graded essays?
2. Can the appraisal theory be utilized as an assessment scale for evaluating the students' argumentative writing?

METHOD

Participants

Seventeen postgraduate (MA) university students majoring in TEFL voluntarily participated in this study. They studied at Razi University of Kermanshah, Iran, and were in the second or third semester at the time when this study was conducted. All the participants had passed at least a minimum of four courses in writing including Grammar, Paragraph Development, and Essay Writing. Our initial survey revealed that participants had not received special instruction on argumentative writing prior to the study. Moreover, to meet their homogeneity, participants' educational background rather than their language proficiency was of paramount importance since there is consensus that by running a single language proficiency test, the students' level of proficiency cannot be fully tapped. In fact, some scholars believe that researchers are occasionally misled by language proficiency tests to divide students into more and less proficient ones (Abedi, Lord, & Plummer, 1997; Artiles, Rueda, Salazar, & Higareda, 2005; LaCelle–Peterson & Rivera, 1994; MaCswan & Rolstad, 2006). As the study aimed to collect data from the Kurdish-speaking participants whose second language is Persian learning English as an L3, two participants whose native language was Azerbaijani Turkish were excluded from the study (N=15). The uniqueness of our participants is that their academic writing experience deals with their second rather than their first language. In fact, Kurdish is only utilized for oral communication among native speakers of Kurdish.

Instrumentation

The instruments which were applied in this study consisted of a modified rating scale, which was adapted from two other scales originally developed by Jacobs et al. (1981) and Ramage, Bean, and Johnson, (2012) for assessing the essays in terms of the macrostructures exploited in them, and the framework for the analysis of appraisal resources.

Rating the Essays

The rating scale employed for assessing the essays was a modified version which was a combination of the Composition Profile (Jacobs et al., 1981) and the Organization Plan for an Argument (Ramage et al.,

2012). The former scale comprises five components of content, organization, vocabulary, language use, and mechanics. To compensate for the shortcomings of the composition Profile, the Organization Plan for an Argument consisting of “introduction, presentation of writer’s position, summary of opposing views, response to opposing views, and conclusion” (Ramage, et al., 2012, p. 59) replaced the organization component of the Composition Profile. In addition, the content category of this scale was modified by adding the “audience reader awareness” as part of the criteria to assess the content category (Connor & Mbaye, 2002, see Appendix for the modified rating scale). It should be noted that although some parts of the Composition Profile were modified, the criteria and the ranges of scores identified by Jacobs, et al. (1981) were followed for rating the essays. Therefore, the highest feasible score for overall proficiency can be 100 and the lowest score can be 34 (Jacobs, et al., 1981). Similarly, the individual scales for each component (content, organization, vocabulary, language use, mechanics) and the total summed scale were categorized as four master levels: *excellent to very good* (83–100 points), *good to average* (63–82 points), *fair to poor* (52–62 points), and *very poor* (34–51 points). The clear-cut point for separating poor and good writers was 63 (Jacobs, et al., 1981). Given this division, essays that were graded 63 and higher were identified as high-graded and those below 63 (between 34 and 62) as low-graded.

The Framework for Appraisal Analysis

The appraisal theory concerned with the interpersonal meanings at the level of grammar and discourse (Martin & White, 20005) can uncover how a writer takes a voice in academic writing. To this aim, we investigated the possible application of this model (see Figure 1) for assessing EFL students’ argumentative writing.

According to Martin and White (2005, p. 35), appraisal consists of three categories of Engagement, Attitude, and Graduation. Engagement covering resources that provide additional voices for discourse by using projection consists of two categories namely, *Monogloss* and *Heterogloss*. The former does not allow any space for other voices and opinions while the latter makes reference to alternative positions and voices (Martin & White, 2005; Nakamura, 2009). The sub-categories of *Heterogloss* are bifurcated into *contract* and *expand* and these two sub-categories in turn are branched out into a few sub-categories as shown in

Figure 1 (Martin & White, 2005). It should be noted here that, following the examples by Martin and White (2005), as well as the classification provided by other scholars (Ho, 2011, Liu, 2013), in this study, *entertain* was sorted out into *probability*, *appearance*, *opinion*, *modal verb*, *conditional verb*, *rhetorical question* and *hearsay* which for space brevity only three of them are displayed in Figure 1.

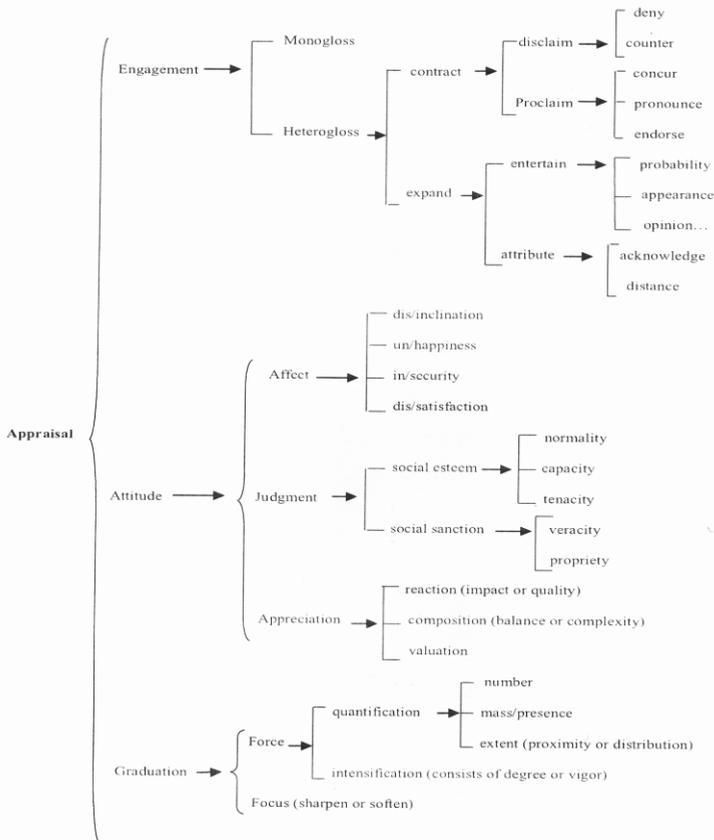


Figure 1. The appraisal system (adapted from Martin & White, 2005)

Attitude is concerned with emotive responses, judgment of behavior and evaluations of natural phenomena (Martin & White, 2005, pp. 42–43). Attitudinal values can be divided into *inscribed* (explicit) and *token/invoked* (implicit) types. Inscribed evaluation is explicitly presented by means of a lexical item whereas invoked evaluation relies on ideational meaning to imply evaluation (Martin, 2004; Martin & White, 2005). The heart of the semantic system of Attitude is *Affect* because it is concerned with positive and negative feelings based on which later development of *Judgment* and *Appreciation* occurs (Martin & White, 2005, p. 42). Judgment values dealing with the resources by which we make our negative or positive evaluation of people and their behavior, according to Martin and White (2005, p. 52), are split into *social esteem* (values “critical to the formation of social network”) and *social sanction* (values germane to “civic duty and religious observance”). Appreciation is concerned with evaluation of either natural or man-made and abstract or concrete processes and phenomena (Martin & White, 2005; Martin, 2000; White, 2005).

Graduation consists of prime branches of *force* and *focus* for grading. Force deals with assessing of amount (*quantification*) or intensity (*intensification*). The extent, as a sub-category of quantification, is split into *proximity* and *distribution* both of which are concerned with time and space (Martin & White, 2005). *Focus* is concerned with grading the values on a scale between core (sharpen) and marginal (soften) membership in a category (Lee, 2006; Martin & White, 2005; Nakamura, 2009).

Within the framework of the appraisal theory, appraisal analysis requires three points namely, *reading position*, *top-down or bottom-up analysis* and *double-coding analysis* to clarify how the data were read and analyzed.

Reading Position: Attitudinal values may be realized explicitly or implicitly. When attitudinal meanings are indicated through explicit lexical items, they can be clearly judged based on the appraisal coding. However, in some cases, attitudinal values can be evoked in an implicit way, and this might cause “subjectivity” in the appraisal analysis (Martin & White, 2005, p. 62). Thus, in order to eschew subjectivity the reading position that a researcher follows should be clarified. In this regard, Martin and White (2005, p. 62) classify three types of reading positions

namely *compliant*, *resistant*, and *tactical* reading and describe them as follows:

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 by the co-selection of meanings in a text, while compliant readings subscribe to it.

Following other studies conducted on EFL learners’ academic writing (Liu, 2013; Liu & Thompson, 2009), in this study, one of the researchers, being a native-speaker of Kurdish and thus culturally and ideationally sharing many values with the study participants, took a compliant reading position towards the participants’ essays for appraisal analysis.

Top-down or Bottom-up Analysis: The appraisal analysis can be done in two ways: top-down and bottom-up. According to Martin and White (2005, p. 70), “this means starting with prosodies and working down to their realizations (top-down) or starting with realizations and working back to the ‘mood’ of a text (bottom-up)”. Both top-down (e.g., Lee, 2006) and bottom-up (e.g., Liu & Thompson, 2009) have been utilized for the appraisal analyses. In this study, the bottom-up analysis was employed.

Double-coding Analysis: In attitudinal analysis of texts, double-coding is sporadically required because attitudinal items are semantically interconnected and sometimes can simultaneously be interpreted in two different ways. The need for double-coding has been voiced by scholars in the field. Martin and White (2005, p. 58) assert that double-coding is allowed for the borders among attitude items. Similarly, Martin and Rose (2003, p. 38) emphasize “bordering of ‘character’ (Judgment) and ‘value’ (Appreciation)”. Table 1 illustrates the double-coding of Attitude.

Table 1: The double-coding of attitude

Appraisal items	Item(s) appraised	Attitude		
		Affect	Judgment	Appreciation
Even things <u>lose</u> their value in his eyes and ...	Things	-sat		-val

sat=satisfaction; val=valuation

As demonstrated in Table 1, the two values coding the appraisal items (lose their value in his eyes) are both from *Attitude*, but they concurrently belong to two dissimilar subcategories of *Affect* and *Appreciation*. Similarly, double-coding is sometimes apposite to different categories of the appraisal theory such as the example illustrated in Table 2.

Table 2: The double-coding of attitude and graduation

Appraisal items	Item(s) Appraised	Attitude			Engagement	Graduation
		Affect	Judgment	Appreciation		
Poor children are <u>better</u> prepared for the problems	poor children		+cap			Intens> quality
...						

cap=capacity; intens= intensification

Accordingly, the value coding the appraisal items of *better prepared* belongs to social esteem which is one of the sub-systems of *Judgment*. However, the appraisal item *better* (or better than) can simultaneously be analyzed in terms of *Graduation*. Throughout the appraisal analysis of the data, wherever the double-coding was distinguished both values were taken into account. It should be noted here that during the appraisal analysis of the data, there was sometimes a dilemma in terms of some appraisal resources to be put under the right category. For example, the lexical item *just* can act either as sharpen (*Graduation* > focus > sharpen) or as counter (contract > disclaim > counter). In this study, when *just* conveyed the meaning of *exactly*, it was considered as sharpen, and when *just* acted as a counter-expectational adjunct, as attested by scholars (e.g., Charles, 2009; Ho, 2011), it was coded as counter.

Data Collection Procedure

In order to examine interpersonal meanings in students' argumentative writing, the participants were required to write an essay of about 250 words within the timeframe of 50 minutes. The topic of the essay was chosen from one of the IELTS exams (IELTS, 2006) entitled *children who are brought up in families that do not have large amount of money are better prepared to deal with the problems of adult life than children*

brought up by wealthy parents. The participants were required to compare and contrast or evaluate the topic and offer reason(s) for their answers. The topic was of general interest, so most of the students were assumed to have some information on it.

The site for data collection for this study was Razi University of Kermanshah, Iran. Since one of the researchers cooperated with the English Language Department of the University, the lecturers who taught advanced writing and other courses to MA students of TEFL were accessible. Following a specific session with a lecturer who taught advanced writing, the main researcher followed the instructor into the class. The topic of the essay along with the instructions and three questions about participants' first language, their second language, and their prior knowledge on argumentative writing which had been prepared on a sheet was administered to the students. Then, the concise descriptive guidelines regarding argumentative essays and the elements of an argument such as *claim, ground, warrant*, etc. were orally provided to the participants in about ten minutes. The participants were announced that they were not permitted to use a dictionary. In addition, to meet ethical considerations, the students were assured of confidentiality of the information that they imparted and of their writing being only utilized for research purposes. By the same token, they were told that their right of remaining anonymous would be assured. Thus, code numbers (1–15) substituted for the students' names. They were also announced that while their participation would be highly appreciated, they still had the option not to take part.

Data Analysis

Altogether, 17 essays were collected out of which two essays written by two participants whose first language was a language other than Kurdish were excluded from the corpus. The next step was the identification of high- and low-graded essays. To do this, the essays were assessed and rated by three raters. The raters were provided with the essays and a check-list of modified rating scale for correction criteria (see Appendix). Likewise, a session was separately held with each rater and any ambiguous point was negotiated before they commenced to assess the argumentative essays. All the three raters were the university lecturers who taught writing courses at the BA and MA levels and also presented IELTS writing classes for a minimum of five years.

Following the assessment of the essays by the three raters, three scores were assigned to each essay. The assignment of the scores to the essays furnished identification of the high- and low-graded essays, and then the process was wound up by computing inter-rater reliability of the scores to verify or reject the consistency of scores. For this purpose, the SPSS software (version 19) was exploited. Pearson Correlation Coefficient obtained the value of 0.91, showing that the consistency of scores is significantly acceptable (Bachman, 1990; Mousavi, 1999). Following this phase, the high- and low-graded essays were differentiated based on the scores given by the raters. Out of 15 graded essays, seven essays were identified as high-graded and eight as low-graded. Finally, the study adopted a mixed method, and the essays were analyzed quantitatively for evaluative markers to find how appraisal resources were exploited in them. Then, out of the 15 essays, two were chosen as the representatives of high- and low-graded essays for bottom-up analysis in order to identify the possible patterns of appraisal exploited in them.

RESULTS

General Analysis

To answer the first research question, results from the analysis of the overall distribution of appraisal categories revealed differences between high-graded (Henceforth HGEs) and low-graded essays (LGEs). As shown in Table 3, although the number of HGEs (N=7) was less than that of LGEs (N=8), a high proportion of attitude was chosen in HGEs (296) in comparison with LGEs (217). However, engagement markers were more frequent in LGEs due to the rise in the exploitation of monoglossic resources in LGEs (N=77) as compared to the same resources in HGEs (N=47) whereas the difference in graduation resources between LGEs (189) and HGEs (184) was marginal. Attitude analysis indicated that both HGEs and LGEs tended to employ more *judgment* and *appreciation* devices than *affect*, which can be justified as one characteristic of argumentative essays (Lee, 2006). Compared to LGEs (N=116; N=77), engagement analysis showed lower proportions of heteroglossic (N=114) and monoglossic (N=47) values in HGEs.

Table 3: Deployment of appraisal categories in each essay

HGEs and LGEs	Attitude			Engagement		Graduation		Total
	Affect	Judgment	Appreciation	Heterogloss	Monogloss	Force	Focus	
HGE 1	3	12	18	14	0	10	3	60
HGE 2	10	15	15	26	2	18	5	91
HGE 3	8	22	16	11	8	23	5	93
HGE 4	5	27	13	9	12	25	8	99
HGE 5	6	17	15	14	7	27	8	94
HGE 6	11	28	21	22	7	21	8	118
HGE 7	2	21	11	18	11	20	3	86
Total	45	142	109	114	47	144	40	
LGE 8	7	7	15	22	14	30	3	98
LGE 9	2	22	9	7	1	33	2	76
LGE 10	2	9	12	24	9	19	7	82
LGE 11	2	19	7	10	6	21	4	69
LGE 12	9	20	11	15	12	18	4	90
LGE 13	5	0	5	20	9	11	6	66
LGE 14	1	6	7	9	14	12	2	51
LGE 15	11	7	11	9	12	17	0	67
Total	39	100	78	116	77	161	28	

Bottom-up Analysis

To see how students articulate themselves in their writings, in the next phase of the study we carried out a more detailed investigation into high- and low-graded essays. To this end, two essays (essay 6, henceforth HGE and essay 12, henceforth LGE), one from each group, were selected for more in depth analysis of how appraisal resources are usefully incorporated and how these devices may help students construct disciplinary identity. It should be noted here that regardless of the number of appraisal resources, LGE 8 was not selected as the representative of LGEs because the force values exploited in this LGE were egregiously higher than LGE 12 and this skewed the results. The number of different categories exploited in HGE and LGE is 118 and 90 respectively (see Table 3). In what follows, the results of using these categories are presented.

Attitude Analysis

Affect Analysis: As demonstrated in Table 4, although similar in the total number of Affect values, HGE produced more in/security and dis/satisfaction to evaluate the consequence of being poor and its disadvantages (see examples 1 & 2). By contrast, LGE utilized

dis/inclination and un/happiness to backup the stated thesis, that is, children in poor families are better prepared to deal with the problems of adult life (see example 3 & 4):

1. *For example, they may show some signs of lack of confidence [affect > -security] and anxiety [affect > -security] with respect to some challenging aspects of life.*
2. *Another disadvantage is that sometimes they feel less self-satisfied [affect > +satisfaction] and, as a result, may do some wrong-doings.*
3. *They can tolerate many bad situations that may be a wealthy person cannot, they like [affect > +happiness] many things in life... so they try [affect > +inclination] to reach them.*
4. *Some of them continue education to reach everything they want [affect > +inclination].*

Table 4: Deployment of sub-categories of affect

	In/Security		Dis/Satisfaction		Dis/Inclination		Un/Happiness	
	H	L	H	L	H	L	H	L
Positive	4	0	3	1	1	6	0	2
Negative	3	0	0	0	0	0	0	0
Explicit	7	0	3	0	1	6	0	2
Implicit	0	0	0	1	0	0	0	0

H= HGE; L= LGE

Likewise, HGE sometimes expressed affectual items in terms of nominalized items (e.g., confidence, anxiety) presenting them in a foregrounded way while LGE mostly articulated these values by surge of feelings (e.g., like, try, want) and in a backgrounded manner. Moreover, in the last paragraph of HGE, the writer considered himself as being poor or wealthy and articulated this emotion by commencing the sentence with *we* and connecting it with strong obligation, that is, *must*, as illustrated in examples 5 and 6:

5. *As a poor child, we must not be anxious about not having money...*
6. *As a wealthy child, we must not be pleased very much for having money ...*

By articulating these sentences, the student writer aimed to emotionally advise the potential readers by incorporating the plural pronoun (*we*) which is a characteristic of Kurdish in giving advice. However, the absence of this trend in the other essays warrants further research in different topics with a larger number of Kurdish-speaking participants.

Analysis of judgment markers: As demonstrated in Table 5, despite both HGE and LGE opting for more social esteem (40) than social sanction values (8), differences were revealed regarding the sub-categories of judgment. HGE utilized *capacity* more than twice as often as LGE whereas LGE employed *normality* almost twice and *tenacity* more than twice as often as HGE.

Table 5: Exploitation of sub-categories of judgment

	Social Esteem				Social Sanction					
	Normality		Capacity		Tenacity		Veracity		Propriety	
	H	L	H	L	H	L	H	L	H	L
Positive	3	6	7	6	2	5	0	0	4	0
Negative	2	3	6	0	0	0	0	0	4	0
Explicit	5	9	13	6	2	5	0	0	8	0
Implicit	0	0	0	0	0	0	0	0	8	0

H= HGE; L= LGE

Regarding social sanction, whereas HGE deployed one sub-category of social sanction, propriety, to evaluate and judge the behavior of children in poor and wealthy families (as exemplified in 7 & 8), this value was not incorporated in LGE to make a judgment.

7. ... *they are kind [judgment > +propriety] to other people and sympathize [Judgment > +propriety] more willingly and wholeheartedly with weak people of their own society.*

8. ... *a poor family may virtually immunize a child against spoiling [Judgment > -propriety] and heighten his/her power*

However, the representative essays were similar to each other in explicitly encoding all judgment values and in employing them to evaluate human beings (in these essays children in poor and wealthy

families) as shown in the following examples (example 9 was taken from HGE and example 10 from LGE).

9. ... *these people would be capable of collaborating [judgment > + capacity] with other people*
10. *A person who has lived in a poor family has more motivation and more tolerance [judgment > +capacity]*

Appreciation Analysis: Differences cropped up among the three sub-categories of appreciation as well as between HGE and LGE. That is, generally, in most cases, appreciation values were encoded as valuation in both HGE and LGE which implied that both essays employed appreciation as things which were significant and worthwhile or insignificant and worthless (Table 6). Likewise, all these encoded valuations were explicitly articulated and expressed in terms of advantages and disadvantages (except for one case) of being brought up in poor and wealthy families in HGE more than twice as often as LGE. Note examples of 11 and 12.

11. *One of its major and important [appreciation > +valuation] advantages [appreciation > +valuation] is that children of poor families are more confident and stronger....*
12. *For example, they may show some signs of lack of confidence and anxiety with respect to some challenging [appreciation > - valuation] aspects of life. Another disadvantage [appreciation > - valuation] is that*

Table 6: employment of sub-categories of appreciation

	Reaction		Composition		Valuation	
	H	L	H	L	H	L
Positive	0	2	1	1	12	5
Negative	0	0	0	0	8	3
Explicit	0	2	1	1	19	8
Implicit	0	0	0	0	1	0

H= HGE; L= LGE

Similarly, LGE also encoded valuation in terms of children's motivation and positive capacities and lack of motivation in children brought up in rich families. Note example 13.

13. *The children who are brought up in a poor family have more motivation and tolerance ability in front of any problem [appreciation > +valuation]*

In terms of other sub-categories of appreciation, both HGE and LGE encoded values as composition whereas they differed in using sub-types of composition, that is, HGE encoded this as positive *balanced* while LGE as positive *complexity* as depicted in examples 14 and 15.

14. *... but it is we as human beings who should firmly learn to solve our problems in an appropriate [appreciation > +composition > balance] way.*
15. *... but about persons brought up in a poor or wealthy families is an intricate [appreciation > + composition > complexity] story.*

Finally, regarding reaction, only LGE encoded two instances of appreciation as reaction using the sub-system of *quality* showing that LGE sometimes relied on emotion to construct appreciation as illustrated in example 16.

16. *... they can tolerate many bad [appreciation > -reaction > quality] situations that maybe a wealthy person cannot.*

Engagement Analysis

Though approximately alike (HGE=29 and LGE=27), LGE exploited more monoglossic values than HGE, showing that LGE left less room for other voices and alternative positions. The fewer number of monoglossic values employed by HGE is compatible with Wu's (2007) study although in Wu's research the participants wrote academic essays in geography. By contrast, HGE deployed more heteroglossic values (HGE=22; LGE=15). Further analysis revealed that whereas contract was preferred in LGE (examples 17 and 18), contract and expand were approximately equally exploited in HGE (examples 19, 20, and 21).

17. ... *he/she does not [disclaim > deny] have any motivation.*
 18. *I say [proclaim > pronounce] money is so important in life.*
 19. *They may [entertain > modal verb] show some signs of lack of confidence and anxiety....*
 20. ... *they feel less self-satisfaction and, as a result, may [entertain > modal verb] do some wrong-doings.*
 21. ... *but [disclaim > counter] it is we as human beings who should learn to solve our problems*

Table 7: deployment of sub-categories of engagement

	<u>Monogloss</u>				<u>Heterogloss</u>											
					<u>Contract</u>							<u>Expand</u>			<u>attribute</u>	
	<u>disclaim</u>		<u>proclaim</u>		<u>entertain</u>											
	de	coun	con	pron	end	pro	app	op	mv	cm	rq	hsay	ack	dist		
H	7	5	3	1	1	0	1	1	0	10	0	0	0	0		
L	12	4	2	0	5	1	1	1	0	0	0	0	1	0		

H= HGE; L= LGE; de= deny; coun= counter; con= concur; pron= pronoun; end= endorse; pro= probably; app= appearance; op= opinion; mv= modal verb; cm= conditional modal; rq= rhetorical question; hsay= hearsay; ack= acknowledge; dist= distance

Graduation Analysis

As demonstrated in Table 8, although, in general, HGE employed more grading items (except for Intensification) in comparison with LGE, both were roughly alike in exploiting grading resources.

Table 8: sub-categories of graduation in HGE and LGE

	<u>Force</u>				<u>Intensification</u>		<u>Focus</u>	
	<u>Quantification</u>		<u>extent</u>		<u>quality</u>	<u>process</u>	<u>Sharpen</u>	<u>Soften</u>
	number	mass	pro	dist				
H	10	2	0	2	5	2	5	3
L	8	0	0	2	7	1	3	1

H= HGE; L= LGE; pro= proximity; dist= distribution

However, a further analysis of HGE and LGE demonstrated that grading resources exploited in HGE were mostly integrated with attitudinal resources as illustrated in the following examples:

22. *One of its major and important advantages is that children of poor families are more [force > intensification > quality] confident [affect > +security] and stronger [judgment > +capacity] [force > intensification > quality].*
23. *Another disadvantage is that sometimes [force > intensification > process] they feel less [force > intensification > quality] self-satisfaction [affect > +satisfaction]*

In example 22, the priority of children in poor families over children in wealthy families was positively and explicitly evaluated and graded through high-value intensification, that is, *quality*. Similarly, in example 23, positive evaluation of the affectual item was down-graded by using low-value intensification (*quality*) but this case was emphasized to happen occasionally by employing the low-value intensification (*process*).

Finally, compared to HGE, although LGE sometimes deployed grading items in association with attitudinal resources (example 24), it primarily exploited grading items in isolation (example 25).

24. *The children who are brought up in a poor family have more [force > intensification > quality] motivation and tolerance [judgment > +capacity]*
25. *Some [force > quantification > number] of them continue education to reach everything they want [affect > +inclination].*

DISCUSSION

The pattern of deploying appraisal categories found in this study is compatible with Wu's (2007) study in which HGEs exploited fewer numbers of monoglossic items than LGEs. Regarding graduation, though, in general, the difference was minimal, the examination of the sub-categories of graduation showed that HGEs applied more items of focus (N=40) and fewer values of force (N=144) than LGEs (force=161; focus=28). Regardless of the proficiency level, marked differences were observed between the essays within each group showing that the essays were not completely homogeneous.

Two reasons may explain this heterogeneity. First, in assessing the essays as high- and low-graded, more important than the number of appraisal markers exploited in essays was how these resources were

employed. For example, as the in-depth qualitative analysis of the two essays showed, contrary to LGE, HGE integrated grading items with affectual devices. The second reason relates to the cut-off point (that is, 63 out of 100) that differentiated high- and low-graded essays. For instance, HGE 1 was assigned 63 and thus distinguished as high-graded while LGE 11 was assigned 61 and consequently as low-graded. In fact, such minimal differences may show the hidden deficiencies of the existing rating scales, in particular, Jacobs et al.'s (1981) ESL Composition Profile.

The findings on attitude analysis are in line with Lee's (2006) and Liu and Thompson's (2009) studies which in the former she found that high-graded essays coded affect resources by nominalized and abstract items and in the latter, as a case study, the researchers observed that their EFL Chinese student expressed affectual items by behavior surges and in a foregrounded manner.

As to judgment analysis, although the prevalence of capacity by HGEs and LGEs is reported in other studies (Lee, 2008; Liu, 2013), a somehow contrasting result was obtained in this study. That is, only HGE used capacity as the main value for evaluating behavior, but LGE applied normality and tenacity instead for evaluating children's behavior in poor and wealthy families as illustrated in the following examples.

26. ... a person who has lived in a poor family has more motivation and more tolerance as well as is patient [judgment > +tenacity], and hardworking [judgment > +tenacity] and adaptable [judgment > +tenacity].

27. They are familiar [judgment > +normality] with hardship

The fact that human beings were the explicit targets of judgment in the LGE essays is consistent with the previous studies (e.g., Lee, 2008, Liu, 2013). However, regarding HGE, the results of this study are in contrast with Lee's (2008, p. 50) successful argumentative writing in terms of the "nominalized abstraction" without explicitly referring to human beings. The scores assigned to HGEs in this study may reflect the inadequacy of the existing rating scales, in particular ESL Composition Profile, for lack of a criterion for this aspect of academic writing.

Regarding appreciation analysis, the deployment of reaction by LGE is compatible with Lee's (2008) and Hood's (2004) findings based on which low-rated essays and consequently "poor writers employ relatively

more frequent constructions of appreciation as reaction” (Lee, 2008, p. 53). However, the reason that LGE in this study exploited more appreciation items as valuation rather than reaction may be justified by the participants’ study level in this research (Master’s students) and Lee’s (2008) study (undergraduate students).

Based on the findings from engagement analysis, some marked differences can be teased out about heteroglossic values employed in HGE and LGE. According to the findings, pronoun values tilted the balance in favor of LGE which somehow reduplicates Wu’s (2007) and Liu’s (2013) studies both of which reported low-graded essays employing more pronoun items. It seems that self-citation enabled the student writer of LGE to get an authorial self and take responsibility about the text. Similarly, LGE relied more on contractive than expansive resources to construct authorial voices whereas HGE adopted a more balanced approach to exploit contractive and expansive resources. However, one sub-category of expand, attribute, also called external voice or extra-vocalizing (Martin & White, 2005), was not applied in the essays. Finally, concerning graduation analysis, the interconnection between grading items and attitudinal values buttresses the essay to be more persuasive as verified in Liu’s (2013) study.

To sum up this section, despite heterogeneous cases, quantitative analysis of high- and low-graded essays showed that, in most cases, those essays in which more appraisal resources, especially attitudinal categories were exploited, were rated as high-graded essays. This may be interpreted that the appraisal theory can be applied as an assessment model for rating ESL/EFL students’ argumentative writing. Considering academic writing as an interactive setting (Hyland, 2000), the appraisal theory can uncover how a writer creates a successful interaction with the potential readers. Because the appraisal theory deals with interpersonal meanings at the level of grammar and discourse (Martin, 2000; Martin & White, 2005), it can meticulously delineate how a writer, as a human being, conveys his/her feelings and specifies his/her voice in writing. However, the question is whether the usual criteria that assessment tools (e.g., Jacobs, et al., 1981; Weir, 1983) take into account are seriously taken if the appraisal theory is applied as the assessment scale. In this case, at least, the appraisal theory can be utilized along with the other analytical scales for evaluating argumentative writing. The influence of appraisal devices on raters in rating essays as high- or low-graded can also be noticed in qualitative analysis of the essays. The number and the

types of the appraisal resources in HGE partly showed how a successful writer uses appraisal resources and integrates attitudinal items with grading resources to make more persuasive writing.

CONCLUSION AND IMPLICATIONS

This study, as part of a larger research, explored the deployment of the appraisal framework in argumentative writings of Kurdish Iranian learners of English. The analysis of HGEs and LGEs revealed important findings. Both quantitative and qualitative findings showed that HGEs favored devices (in particular, in/security and dis/satisfaction), in nominalized forms, to position readers attitudinally. Likewise, in line with Lee (2008), HGEs primarily exploited judgment>capacity and judgment>propriety to furnish the potential readers with the message. Our quantitative results showed that the difference between HGEs and LGEs in using engagement was large and noticeable. Compared to LGE, the qualitative findings revealed that HGE applied a higher frequency of heteroglossic and a lower frequency of monoglossic resources. In addition, HGE followed a balanced application of contractive and expansive resources whereas LGE preferred only contractive items to construct authorial voices. Furthermore, the obtained patterns of graduation revealed that HGE utilized grading items to make a connection and an interaction between grading values and attitudinal resources.

Irrespective of the nature of the topic, though the current study assessed the argumentative writings of masters' students of English as their third language, the effect of their linguistic background was not seriously explored in this study. Accordingly, further research is called for to compare students with various linguistic backgrounds and to investigate whether writers' first or second language is effective in exploitation of appraisal devices in their writings.

Finally, the analysis developed here suggested the pivotal role of explicit evaluative strategies on the raters to assign the essays as high-graded. Thus, if argumentative essays are to be evaluated as high-graded, the observed features of the HGEs like the exploitation of more judgment > capacity, appreciation > valuation and the integration of grading items with affectual ones underscore the values upon which judgments of an appropriate argumentative writing are made. However, it should be noted that these characteristics can only be seen as tendencies rather than as

established facts, and they do not stifle creativity which can be the essence of every piece of writing.

Although this study investigated appraisal resources exploited in students' argumentative writing, quantitatively, the population was not large enough to be generalized to other situations and qualitatively only two essays were analyzed. In addition, this study only examined the argumentative writing of Kurdish speaking Iranian MA students learning English as an L3. Thus, the issue warrants further research into a larger number of academic essays written by students with other native languages to see how they employ appraisal markers and whether patterns of L1 or L2 or even both languages are transferred into L3. Finally, more studies are needed to substantiate the application of the appraisal theory as an analytical scale for assessing students' writing.

Bio-data

Alireza Jalilifar is associate professor of Applied Linguistics at Shahid Chamran University of Ahvaz, Iran. He has published more than 70 papers in local and international journals as well as three books in discourse analysis. His main interests are second language writing, genre analysis, and academic discourse.

Ali Hemmati is currently a Ph.D. candidate of TEFL at Shahid Chamran University of Ahvaz, Iran. His areas of interest include discourse analysis, (academic) writing, and testing.

References

- Abedi, J., Lord, C., & Plummer, J. (1997). *Language background as a variable in NAEP mathematics performance (Tech. Rep. No. 429)*. Los Angeles: University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Students Testing.
- Artiles, A. J., Rueda, R., Salazar, J. J., & Higaeda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children*, 71(3), 1-17
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Charles, M. (2009). Stance, interaction and the rhetorical patterns of restrictive adverbs: Discourse roles of only, just, simply and merely. In M. Charles,

- D. Pecorari & S. Hunston (Eds.), *Academic writing: An interface of corpus and discourse* (pp. 152-169). London: Continuum.
- Connor, U. (1996). *Contrastive Rhetoric: Cross-cultural aspects of second-language writing*. Cambridge: Cambridge University Press.
- Connor, U., & Mbaye, A. (2002). Discourse approaches to writing assessment. *Annual Review of Applied Linguistics*, 22, 263-278.
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd edition). London: Edward Arnold.
- Halliday, M. A. K. (2004). *The language of early childhood*. London: Continuum.
- Halliday, M. A. K., & Matthiessen, C. (2004). *An introduction to functional grammar*. London: Arnold.
- HawKey, R., & Barker, F. (2004). Developing a common scale for the assessment of writing. *Assessing Writing*, 9(2), 122-159.
- Ho, V. L. (2011). *Non-native argumentative writing by Vietnamese learners of English: A contrastive study*. Unpublished PhD thesis. Washington, DC.
- Hood, S. (2004). *Appraising research: Taking a stance in academic writing*. Unpublished PhD thesis. University of Technology, Sydney.
- Hunston, S. (2000). Evaluation and the planes of discourse: Status and value in persuasive tests. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 176-208). Oxford: Oxford University Press.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London: Longman.
- 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.
- IELTS (2006). IELTS scores explained. Retrieved from: <http://www.ielts.org>
- IELTS (2006). Research note, Issue 23. Retrieved from: <http://www.ielts.org>
- Irvin, L. L. (2010). What is academic writing? In C. Lowe & P. Zemliansky (Eds.), *Writing spaces: Readings on writing* (Vol. 1), (pp. 3-17). Indiana: Parlor Press LLC.
- Jacobs, H., Zingraf, S., Wormuth, D., Hartfiel, V. F., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. MA: Newbury House Publishers.
- Kisting, W. (2007). *Writing effective essays: A guide to college-level writing*. North Charleston, US: Createspace Independent Publishing Platform.
- LaCelle-Peterson, M. W., & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review*, 64(1), 55-75.

- Lee, S. H. (2006). *The use of interpersonal resources in argumentative/persuasive essays by East-Asian ESL and Australian tertiary students*. Unpublished PhD thesis. University of Sydney.
- Lee, S. H. (2008). Attitude in undergraduate persuasive essays. *Prospect*, 23(3), 43-58.
- Liu, X. (2013). Evaluation in Chinese university EFL students' English argumentative writing: An appraisal study. *Electronic Journal of Foreign Language Teaching*, 10(1), 40-53.
- 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.
- MacSwan, J., & Rolstad, K. (2006). How language proficiency tests mislead us about ability: Implications for English language learner placement in special education. *Teachers College Record*, 108(11), 2304-2328.
- Martin, J. R. (1997). Analysing genre: functional parameters. In F. Christie & J. R. Martin (Eds.), *Genre and Institutions: Social process in the workplace and school* (pp. 3-39). London: Cassell.
- Martin, J. R. (2000). Beyond exchange: Appraisal system in English. In S. Hunston & G. Thompson (Eds.), *Evaluation in text* (pp.145-175). Oxford: Oxford University Press.
- Martin, J. R. (2004). Sense and sensibility: Texturing evaluation. In J. Foley (Ed.), *Language, education, and discourse* (pp. 270-304). London: Continuum.
- Martin, J. R., & Rose, D. (2003). *Working with discourse: Meaning beyond the clause* (1st ed.). London: Continuum.
- Martin, J. R., & White, P. P. R. (2005). *The language of evaluation: Appraisal in English*. New York: Palgrave Macmillan.
- Mousavi, S. A. (1999). *A dictionary of Language testing* (2nd edition). Tehran: Rahnama Publications.
- Nakamura, A. (2009). *Construction of evaluative meanings in IELTS writing: An intersubjective and intertextual perspective*. Unpublished PhD thesis. University of Wollongong, Wollongong.
- Nordquist, R. (2011). Academic writing. Retrieved from: <http://grammar.about.com/od/g/academicwritingterm.htm>
- Ramage, J. D., Bean, J. C., & Johnson, J. (2012). *Writing arguments: A rhetorical with readings* (9th edition). New Jersey: Pearson Education, Inc.
- Richards, J. C., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics* (3rd ed.). Pearson Education Limited.
- Schultz, J. M. (1991). Writing mode in the articulation of language and literature classes: Theory and practice. *The Modern Language Journal*, 75, 411-417.

- Umair, N. (2011). Problems of multi-ability academic English writing classes in Arab countries. *Arab World English Journal*, 2(2), 230-242.
- White, P. R. R. (2001). An introductory tour through appraisal theory. Retrieved from: <http://www.Grammatics.Com/appraisal/index.html>.
- White, P. R. R. (2005). Update on engagement. Retrieved from: <http://www.grammatics.Com/appraisal/Engagement.Latest.doc>
- Wu, S. M. (2007). The use of engagement resources in high- and low-graded undergraduate geography essays. *Journal of English for academic Purposes*, 6(3), 254-271.
- Wu, S. M. (2008). Investigating the effectiveness of arguments in undergraduate essay from an evaluation perspective. *Prospect*, 23(3), 59-75.
- Yumin, C. (2009). *Interpersonal meaning in the textbooks for teaching English as a foreign language in China: A multimodal approach*. Unpublished PhD thesis, University of Sydney.

Appendix: The Modified Assessment Scale for Essays (adopted from Jacobs, et al., 1981; Ramage, et al., 2012)

Score	Range	Content
	30-27	EXCELLENT TO VERY GOOD: knowledgeable • substantive • thorough development of thesis • relevant to assigned topic • audience/reader awareness
	26-22	GOOD TO AVERAGE: some knowledge of subject • adequate range • limited development of thesis • mostly relevant to topic, but lacks detail • somewhat the awareness of audience/reader
	21-17	FAIR TO POOR: limited knowledge of subject • little substance • inadequate development of topic • inadequate audience/reader awareness
	16-13	VERY POOR: does not show knowledge of subject • non-substantive • not pertinent • OR not enough to evaluate • no awareness of audience/reader
		Organization plan for an argument
	20-18	EXCELLENT TO VERY GOOD: presents the complete introduction (attention grabber, writer's thesis) • presents the reason(s) supporting the claim • summarizes the views differing from writer's • discusses for and against the views • brings essay to closure
	17-14	GOOD TO AVERAGE: somewhat introduces the thesis •

		presents and supports the reason(s) to some extent • loosely presents the main body of essay • limited discussion for and against the views • loosely sums up argument
	13-10	FAIR TO POOR: incomplete thesis • disconnected reasons • incomplete discussion of views • ambiguous and incomplete summary of argument
	9-7	VERY POOR: no organization of thesis, reason(s) and conclusion
		Vocabulary
	20-18	EXCELLENT TO VERY GOOD: sophisticated range • effective word/idiom choice and usage • word form mastery • appropriate register
	17-14	GOOD TO AVERAGE: adequate range • occasional errors of word/idiom form, choice, usage but meaning not obscured
	13-10	FAIR TO POOR: limited range • frequent errors of word/idiom form, choice, usage • meaning confused or obscured
	9-7	VERY POOR: essentially translation • little knowledge of English vocabulary, idioms, word form • OR not enough to evaluate
		Language Use
	25-22	EXCELLENT TO VERY GOOD: effective complex constructions • few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions
	21-18	GOOD TO AVERAGE: effective but simple constructions • minor problems in complex constructions • several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions but meaning seldom obscured
	17-11	FAIR TO POOR: major problems in simple/complex constructions • frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, deletions • meaning confused or obscured
	10-5	VERY POOR: virtually no mastery of sentence construction rules • dominated by errors • does not communicate • OR not enough to evaluate
		Mechanics
	5	EXCELLENT TO VERY GOOD: demonstrates mastery of conventions • few errors of spelling, punctuation, capitalization, paragraphing
	4	GOOD TO AVERAGE: occasional errors of spelling,

		punctuation, capitalization, paragraphing but meaning not obscured
	3	FAIR TO POOR: frequent errors of spelling, punctuation, capitalization, paragraphing • poor handwriting • meaning confused or obscured
	2	VERY POOR: no mastery of conventions • dominated by errors of spelling, punctuation, capitalization, paragraphing • handwriting illegible • OR not enough to evaluate