

Gravity of EFL Learners' Grammatical Errors: A Survey-Based Study of Teachers' Perception

Musa Nushi¹, Vida Khazaei²

¹ Shahid Beheshti University, Tehran, Iran

E-mail: m_nushi@sbu.ac.ir

ORCID: <https://orcid.org/0000-0003-1917-5372>

² Shahid Beheshti University, Tehran, Iran

E-mail: vida.khpro1989@gmail.com

ORCID: <https://orcid.org/0000-0002-9567-4336>

DOI: 10.26907/esd.19.2.04

EDN: GMXMWV

Submitted: 8 August 2023; Accepted: 17 May 2024

Abstract

The present study investigates 110 EFL teachers' perception of the gravity of ten types of grammatical errors made by EFL learners in terms of acceptability. Moreover, it examines the relationship of age, gender, academic degree, years of teaching experience, and the highest level taught with the teachers' judgements. Results revealed that the teachers' evaluations form a hierarchy in which errors are placed at different gravity levels in accordance with their level of acceptability. Moreover, the three variables of academic degree, years of teaching experience, and the highest level taught had a positive correlation with the teachers' evaluations. This study suggests that teachers should make their evaluations systematic, treat errors in accordance with their priority, and become aware of the factors that contribute to evaluations of grammatical errors.

Keywords: EFL teachers, EFL learners, grammatical errors, error gravity, acceptability.

Серьезность грамматических ошибок, совершаемых учащимися курсов английского языка как иностранного (EFL): Опрос мнения преподавателей

Муса Нуши¹, Вида Хазаи²

¹ Университет имени Шахида Бехешти, Тегеран, Иран

E-mail: m_nushi@sbu.ac.ir

ORCID: <https://orcid.org/0000-0003-1917-5372>

² Университет имени Шахида Бехешти, Тегеран, Иран

E-mail: vida.khpro1989@gmail.com

ORCID: <https://orcid.org/0000-0002-9567-4336>

DOI: 10.26907/esd.19.2.04

EDN: GMXMWV

Дата поступления: 8 августа 2023; Дата принятия в печать: 17 мая 2024

Аннотация

В настоящем исследовании 110 преподавателей EFL анализировали серьезность десяти типов грамматических ошибок, допущенных учащимися. Оценки педагогов образовали иерархию, в которой ошибки располагались на разных уровнях в соответствии со степенью их прием-

лемости. Авторами исследовалась связь возраста, пола, ученой степени, стажа и уровня преподавания – с оценками преподавателей. Результаты показали, что три переменные – ученая степень, стаж и уровень преподавания – положительно коррелируют с оценками преподавателей. Полученные данные предполагают, что оценивание должно носить систематический характер, рассматривать ошибки в соответствии с их приоритетом и учитывать факторы, которые способствуют адекватной оценке грамматических ошибок.

Ключевые слова: преподаватели EFL, учащиеся EFL, грамматические ошибки, серьезность ошибок, приемлемость.

Introduction

Although errors are indicators of learning progress, it does not mean that they should not be corrected. In fact, error correction is a key practice in developing second language (L2) writing skills (Ellis, 2009; Ferris, 2002; Shirazizadeh & Amirfazlian, 2019). However, what poses a major problem for teachers in the process of error correction is the evaluation of errors. Sheorey (1986, p. 306) believes “evaluation is particularly challenging because of the lack of well-established guidelines for evaluating the significance of different types of errors”. Although researchers such as James (1977, p. 124) have argued that there is not an absolute answer to the question ‘How serious is such-and-such a grammatical error?’, many studies have tried to provide an answer to the that question. In general, error gravity “is characterized by the attempt to identify those errors which are perceived to be most serious and/or distracting to readers and/or listeners, together with the factors that tend to influence such judgements” (Endley, 2016, p. 1).

One of the major problems with the previous studies on error gravity is that most of them have focused on the comparison between native and non-native teachers. These studies could be categorized into two groups. The first group, such as the one by Sheorey (1986), are focused on the differences existing between native and non-native teachers’ perception of the error gravity. Their main reason for doing so is to make EFL teachers’ evaluative criteria closer to those of native teachers. The basic assumption of these studies is that the native teacher “not only knows what is correct, but also what is appropriate for each context of use” (Ellis, 1986, p. 78). The point being ignored in this type of studies is that native teachers’ perception of the error gravity cannot necessarily be extended to the contexts where English is taught and learned as a foreign language. In these contexts, non-native teachers are more concerned with “the basicness of the rules infringed,” in contrast to native teachers who are mainly concerned with the intelligibility of the message being communicated (Hughes & Lascaratou, 1982, p. 177).

The second group of studies seek the similarities that exist between native and non-native teachers’ perception of error gravity in an effort to reach an agreement on the severity of different error types. The main problem with this type of studies is that they have not led into any consistent results to enhance the quality of “native/nonnative interaction” (Rifkin & Roberts, 1995, p. 512). Lennon (2008) argues this lack of consistency among teachers’ perception is caused by their reliance on different evaluative criteria. According to Davies (1983), teachers who have the same first language (L1) as learners would not normally base their evaluations on the degree of intelligibility of utterances, since most of the utterances seem intelligible to them because of their ability to recognize L1 interference.

The purpose of the present study was to investigate EFL teachers’ perception of the gravity of different types of grammatical errors made by EFL learners to see how they would evaluate the errors in terms of the criterion of acceptability, defined as the extent to which an error deviates from the target language norms. This study specifically focuses on the written grammatical errors at the sentence level. The main reason for focusing on grammatical errors, not lexical errors, is these errors occur more frequently in the

learners' writings. In addition, the study investigates the possible correlations of age, gender, academic degree, years of teaching experience, and the highest taught level with the teachers' judgements. Most of the previous error gravity studies ignored the possible impact of these variables on teachers' evaluations.

Literature Review

Error Gravity

Error gravity refers to the extent to which a listener or reader perceives a language error to be serious or distracting (Rifkin & Roberts, 1995). One of the factors that influence teachers' perception of the error gravity is the teaching method. In general, those methods that have their origin in behavioural psychology (e.g., the Audiolingual Method) find errors undesirable and are strict about correcting them (Lennon, 2008). Proponents of the communicative approaches (e.g., Communicative Language Teaching), however, believe that teachers should "be tolerant of learners' errors as they indicate that the learner is building up his or her communicative competence" (Richards & Rodgers, 2014, p. 95). It should be noted that determining the seriousness of errors based on their perceived gravity is a subjective process. This means that teachers rely on different criteria when they evaluate an error, although "they do not explicitly formulate these criteria" (James, 1977, p. 116). Brown (1991) reached the same conclusion and indicated that although teachers might assign written errors the same scores, it is possible that they approach the task of evaluation from different perspectives.

Error Gravity Criteria

Intelligibility and acceptability are the two frequently used criteria to assess the gravity of language errors. Intelligibility refers to the extent to which an error can impede effective communication, and acceptability relates to the degree to which an error appears irritating to the listeners or readers (Rifkin & Roberts, 1995). According to Burt (1975), *global errors* can highly affect the intelligibility of a message. Therefore, to improve the communicativeness quality of the sentences produced by learners, teachers should pay more attention to global errors. The acceptability of an error, on the other hand, depends on the extent to which it deviates from the L2 norms (Rifkin & Roberts, 1995). The norms can be either competence-based or performance-based: Competence-based errors refer to those errors that result from not following the fixed rules of the L2 (e.g., word order), and are the same among all the speakers. On the other hand, performance-based errors happen when the learners do not follow "some standard view" of the L2 grammar (p. 523).

Factors Affecting Teachers' Evaluations of Error Gravity

Vann et al. (1984) investigated the perceived seriousness of written language errors and the factors that might affect the judges' evaluation of those errors. The researchers found that among the predefined factors of the study, only age and academic discipline of faculty members appeared to have an effect on their evaluation of the gravity of the errors. Hyland and Anan (2006) examined the ability of raters to identify errors in the writing of a Japanese EFL student. The investigation utilized data gathered from a correction task and a questionnaire, focusing on the beliefs and practices of three distinct groups consisting of 16 participants each: native English-speaking (NES) EFL teachers, Japanese-speaking EFL teachers, and educated native English-speaking non-teachers. The participants were instructed to identify and rectify errors in an authentic text written by the student, assess which errors they considered the most serious, and provide justifications for their choices. The findings indicated that non-native English-speaking

(NNES) teachers tend to be more stringent in grading errors and place greater emphasis on rule violations rather than comprehensibility when evaluating severity. Furthermore, it was discovered that Japanese teachers were more inclined to perceive stylistic variations as errors, whereas NES teachers demonstrated sensitivity towards formal features and academic appropriateness. The researchers contend that these disparities arise from the participants' diverse experiences.

Rao and Li (2017) also investigated NES and NNES English language teachers' perceptions of the gravity of errors, along with the factors that could impact their perception. The results indicated that factors such as "cultural belief, educational background, teaching style, and English proficiency" influenced the teachers' evaluations (p. 51). Rao and Liu (2020) examined the assessment criteria used by NES and NNES teachers when evaluating writing in a Chinese educational context. The study employed both quantitative and qualitative methods to explore not only how these two groups of teachers holistically assess students' writing but also how they differ in their analytical justifications for their ratings. The findings indicated significant differences between NES and NNES teachers in holistic evaluation, with NES teachers assigning higher scores to all essays. In terms of analytic evaluation across ten specific categories, the two groups exhibited statistically significant differences in four categories. NES teachers tend to be more lenient in assessing 'grammar' and 'sentence structure,' while NNES teachers are less stringent in rating 'ideas' and 'arguments.' The results suggested that disparities in the assessment of students' English writing between NES and NNES teachers may stem from differences in their language learning experiences, teaching methodologies, and pedagogical beliefs.

Continuing that line of research, the present study focused on a group of Iranian EFL teachers to see how they would evaluate different types of grammatical errors in terms of the criterion of acceptability. In addition, it investigated the role of age, gender, academic degree, years of teaching experience, and the highest taught level on the teachers' judgements. To achieve the objectives of the study, the following questions were posed:

1. What are Iranian EFL teachers' perceptions of the gravity of written grammatical errors in terms of the criterion of acceptability?
2. Is there any significant relationship between teachers' age, gender, academic degree, years of teaching experience and the highest taught level they taught and their perceptions of error gravity in terms of the criterion of acceptability?

Method

Participants

The first group of participants were 33 Iranian upper-intermediate EFL learners studying English at a language institute in Tehran, Iran. The second group of participants were 110 Iranian EFL teachers. The main requirements for the second group of participants were having at least an undergraduate degree in one of the academic disciplines of English Literature, English Translation Studies, and Teaching English as a Foreign Language (TEFL) and having experience in teaching English to Iranian learners. The requirement ensured that the participating teachers were proficient enough in English to judge the gravity of the learners' linguistic errors. Moreover, Allwright (1998, as cited in Oliaei & Sahragard, 2013) contends that "inexperienced teachers, ..., have considerable difficulty when it comes to making judgments of acceptability". The participants were all selected through convenience sampling; they were asked to fill out a consent form and were assured of the anonymity of the data. Table 1 provides information on the five variables of the study.

Table 1. *Demographic information of the participants*

<i>Age</i>	<i>Number</i>	<i>Percentage</i>
20-25	17	15.5
26-31	25	22.7
32-37	23	20.9
38-43	19	17.3
44-49	13	11.8
50+	13	11.8
<i>Gender</i>	<i>Number</i>	<i>Percentage</i>
Female	45	40.9
Male	65	59.1
<i>Academic Degree</i>	<i>Number</i>	<i>Percentage</i>
Bachelor	30	27.3
Master	35	31.8
Ph.D.	45	40.9
<i>Years of Teaching Experience</i>	<i>Number</i>	<i>Percentage</i>
1-5	26	23.6
5-10	33	30
10+	51	46.4
<i>The Highest Taught Level</i>	<i>Number</i>	<i>Percentage</i>
Elementary	8	7.3
Intermediate	12	10.9
Upper-Intermediate	14	12.7
Advanced	39	35.5
Proficient	37	33.6

Instruments

The data for the study were obtained through two instruments. The first was a sample of IELTS Writing Task 1 and 2 from a retired IELTS test (Cambridge IELTS 12), completed by a group 33 Iranian EFL learners, which helped the researchers to collect grammatical errors. In Task 1, the learners were presented with a graph, chart, or diagram, and were required to analyze and present the information in 150-200 words. Task 2 required the learners to write a 250 to 300-word essay in response to a topic. The second instrument was an online seven-point Likert scale questionnaire that was used to collect the teachers' evaluations of the grammatical errors. The questionnaire was made of two sections. In the first section the respondents were asked to provide some information on their age, gender, academic degree, academic discipline, position, years of teaching experience, and the highest taught level. In the second section, they were asked to evaluate the acceptability of 20 errors embedded in 20 sentences on a seven-point Likert scale (1. perfectly acceptable, 2. acceptable, 3. slightly acceptable, 4. neutral, 5. slightly unacceptable, 6. unacceptable, and 7. totally unacceptable). Moreover, the teachers were encouraged to provide a brief

explanation for why they had evaluated an error as being totally unacceptable. It is noteworthy that the intelligibility criterion was excluded because, as stated by Davies (1983), teachers who share the same L1 with their students may be more lenient towards language deviations caused by L1 interference. This is because of their familiarity with and understanding of their students' L1. Another problem with the criterion of intelligibility is that the researcher cannot be sure that the meaning comprehended by the teachers is the same as the meaning intended by the writer (Khalil, 1985).

Before sending the questionnaire to the teachers, it was sent to five research experts to see how they would evaluate its face validity and content validity on a four-point scale (highly relevant, quite relevant, somewhat relevant, and not relevant). After calculating I-CVI (Item Content Validity Index) for each item, those items with values lower than 0.79 were modified or discarded from the study. The S-CVI (Scale Content Validity Index) of the questionnaire as a whole was above 0.85, which indicated that the average content validity index of the instrument was at an acceptable level. In the next phase, a pilot study was conducted to check the feasibility of the questionnaire. For this reason, the questionnaire was sent to 50 EFL teachers, and the collected responses were used to measure the construct validity and the reliability of the instrument through factor analysis and Cronbach's alpha respectively.

Data Collection Procedure

The data was collected in four phases: In the first phase, a group of 33 Iranian EFL learners were asked to write essays following IELTS Writing Task 1 and 2 instructions. In the second phase, the writings were examined for the most frequent errors. The errors fell into ten categories: preposition, article, plural, subject-verb agreement, conjunction, possessive, verb form (verb, to+verb, verb+ing), word order, pronoun, and tense. Twenty sentences out of 262 sentences, representing the aforementioned errors, were selected so that there were two sample sentences for each type of error. In the third phase, an online seven-point Likert scale questionnaire was developed by the researchers. To avoid different interpretations of the presented errors and to neutralize the negative effect of other erroneous aspects of the writings, the context for each error was limited to a sentence. Later, the sentences were corrected to have only one error in each so that the participants could focus on only one type of error in each sentence. Moreover, the sentences were presented randomly so that there were not two consecutive examples of the same type of error. In the final phase, the questionnaire was emailed to 110 Iranian EFL teachers. The teachers were given no information on the learners' language proficiency; they were asked to evaluate the acceptability of the errors, imagining that they were teaching English grammar for writing and the presented errors were made by their students in their writings.

Results

Reliability of the Questionnaire

The reliability coefficient of the questionnaire was measured through Cronbach's alpha, and the result was 0.93, which is a highly desirable value for the reliability of a scale. In general, values above 0.7 indicate that the reliability of an instrument is at an acceptable level (Shultz, et al., 2014, p. 72).

Construct Validity of the Questionnaire

The construct validity of the questionnaire was determined through exploratory factor analysis. The process of exploratory factor analysis was performed in three steps.

Checking Data Suitability

In the first step, the suitability of the data for factor analysis was examined. As the value of Kaiser-Meyer-Olkin (KMO) was above 0.6, the value of Barlett's test of sphericity was significant at 0.05, and most of the correlation coefficients in the Correlation Matrix table were above 0.3; thus, it was concluded that the data was suitable for factor analysis (Table 2).

Table 2. KMO and Barlett's Test

KMO	Barlett's Test
0.82	0.000

Factor Extraction

The main purpose of factor extraction is to identify the number of components on which the items load. In this study, the Principal Component Analysis was used as the extraction method, and to determine the number of desired factors to retain, Kaiser's criterion was used. To determine which components could be extracted in this step, it was needed to look for eigenvalues above 1. By looking at the values presented in Table 3, it was clear that there were five components that could be extracted. These components made 71.3% percent of the variance.

Table 3. Total Variance Explained

Component	Initial Eigenvalues		Cumulative %
	Total	% of Variance	
1	8.98	44.92	44.92
2	1.81	9.06	53.98
3	1.29	6.45	60.43
4	1.17	5.58	66.28
5	1.01	5.05	71.34

Sometimes Kaiser's criterion results in too many extracted components, therefore, one can check scree plot for a fewer number of components. Figure 1 shows that there is a break between components three and two. Considering the greatness of the breaks between the components, one can say that there are two components in this scale.

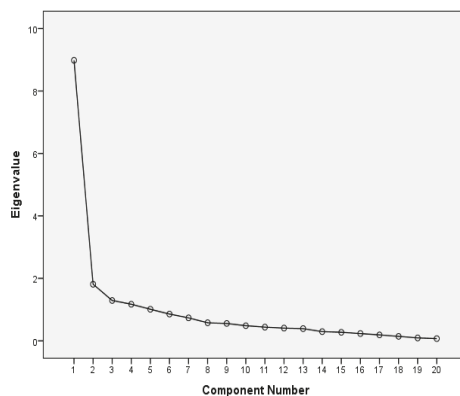


Figure 1. Scree Plot

To make the final decision about the number of components, the values of the Pattern Matrix were examined. Those component with fewer than 4 items loaded on were discarded. At the end two components with higher correlation values were retained. These two components were responsible for 53.977% of the variance.

Factor Rotation and Interpretation

To interpret the two retained components, one of the oblique techniques called “Direct Oblimin” was performed. The first thing needed to be checked about the two components was the strength of the relationship between them. As the values of correlations, presented in Table 4, were above 0.3, it was safe to say that the two components are strongly correlated.

Table 4. *Component Correlation Matrix*

<i>Component</i>	<i>1</i>	<i>2</i>
1	1.00	0.41
2	0.41	1.00

As the values of pattern coefficients indicated most of the items loaded substantially on component one. Moreover, the values of the structure coefficients indicated that most of the items had strong correlation with the component one. This can be an indication of construct validity since the underlying construct of the questionnaire could be legitimately represented via one component, accounting for approximately 45% of the variance.

Finally, the values of communalities indicated that there was no need for removing any of the items, since the variance explained in each item was greater than 0.32 (Table 5).

Table 5. *Pattern and Structure Matrix for PCA with Oblimin Rotation of Two Factor Solution*

<i>Item</i>	<i>Pattern Coefficients</i>		<i>Structure Coefficients</i>		<i>Communalities</i>
	<i>Component 1</i>	<i>Component 2</i>	<i>Component 1</i>	<i>Component 2</i>	
15	0.96	-0.31	0.83	0.09	0.77
9	0.86	-0.20	0.78	0.16	0.64
11	0.73	0.08	0.77	0.38	0.59
5	0.73	0.12	0.78	0.42	0.62
12	0.72	-0.25	0.62	0.05	0.34
10	0.69	0.08	0.73	0.37	0.54
8	0.67	0.25	0.77	0.53	0.65
19	0.66	0.25	0.76	0.52	0.63
17	0.65	-0.03	0.64	0.24	0.41
7	0.65	0.26	0.76	0.52	0.63
14	0.60	0.21	0.68	0.45	0.50
20	0.58	0.06	0.61	0.30	0.37
13	0.54	0.25	0.65	0.47	0.47
18	0.47	0.31	0.60	0.51	0.44
4	0.38	0.32	0.51	0.48	0.35
1	0.03	0.82	0.37	0.84	0.70
3	0.01	0.80	0.35	0.81	0.65
2	0.00	0.71	0.30	0.71	0.50
6	0.41	0.44	0.60	0.61	0.52
16	0.34	0.37	0.49	0.51	0.36

The values of the reliability index and the results of factor analysis warranted that the used instrument could serve as a basis for exploring Iranian EFL teachers' perception of the gravity of grammatical errors.

Research Question 1

The first research question asks about the Iranian EFL teachers' perception of the gravity of written grammatical errors in terms of the criterion of acceptability. To determine the general tendency of the respondents in their evaluations, the mean of the total scores, assigned by the participants, was compared with the median of the possible total scores. As the mean of the total scores (94.56) was above the median (80), it was concluded that there is a general tendency towards negative responses. Moreover, the choices "Slightly Unacceptable" (18.63%), "Unacceptable" (30.64%), and "Totally Unacceptable" (14.90%) were selected more than other options (Table 6).

Table 6. *Frequency of the Choices for each Item*

<i>Items</i>	<i>Perfectly Acceptable</i>	<i>Acceptable</i>	<i>Slightly Acceptable</i>	<i>Neutral</i>	<i>Slightly Unacceptable</i>	<i>Unacceptable</i>	<i>Totally Unacceptable</i>
Q1	5	26	22	6	22	21	8
Q2	6	18	16	5	26	28	11
Q3	5	11	18	8	19	30	19
Q4	3	17	12	10	24	31	13
Q5	4	14	8	4	25	31	24
Q6	2	16	12	7	10	42	21
Q7	3	12	12	4	28	36	15
Q8	2	9	8	9	16	45	21
Q9	2	11	9	5	12	47	24
Q10	5	9	13	3	16	45	19
Q11	7	15	18	5	17	34	14
Q12	8	7	11	5	20	39	20
Q13	8	10	14	11	25	28	14
Q14	8	20	17	12	26	19	8
Q15	1	10	11	3	29	40	16
Q16	11	18	24	8	18	23	8
Q17	1	5	10	5	13	44	32
Q18	4	10	12	2	17	40	25
Q19	4	30	19	10	28	14	5
Q20	8	16	15	4	19	37	11
<i>Total</i>	97	284	281	126	410	674	328

To determine the gravity of the errors, the total score and the mean score were calculated for each item. The highest mean score belonged to question 17, which asked the respondents to rate the acceptability of a type of verb form error, and the lowest mean score belonged to question 19, which sought the acceptability of a type of article error (Table 7).

Table 7. Descriptive Statistics for each Item

Question	Type of Error	Total Score	Mean	Standard Deviation
1: [Not only it would] provide them with financial support, but also it would teach them how to be a useful member of their society.	word order	439	3.99	1.79
2: Some believe that children's engagement in different types of paid [works] can be beneficial for them.	Plural	485	4.41	1.82
3: Japanese tourists showed an [interest to] Australia.	Preposition	521	4.74	1.81
4: Other [countries] share of the Japanese tourist market has been boosted.	Possessive	510	4.64	1.73
5: It encourages them to take more responsibilities, and as a result, [] will become more successful in their social life.	Pronoun	551	5.01	1.80
6: What kind of job [they are] trying to do?	word order	547	4.97	1.81
7: It has lots of effects on their self-confidence and [make] them responsible people.	subject-verb agreement	540	4.91	1.67
8: They can learn about the value of money by playing [or go] shopping with their parents.	parallel structure	577	5.25	1.59
9: The number of Japanese tourist [who they] want to travel to Australia is increasing.	Pronoun	581	5.28	1.68
10: in the year 1985, Japanese tourists did not travel to Australia, but as the years [go] by, they became more interested in the country.	Tense	557	5.06	1.76
11: The number of Japanese tourists [are] increasing.	subject-verb agreement	498	4.53	1.89
12: [It's] beautiful nature has made it a unique destination for tourists.	Possessive	549	4.99	1.81
13: There are some jobs that [require people being] in touch with other people.	verb form	505	4.59	1.79
14: Child labour includes [] variety of different activities.	Article	447	4.06	1.77
15: The percentage [] Japanese tourist travelling abroad has increased.	Preposition	563	5.12	1.53
16: Japan is rich in high technology [and a pioneer] in the field of computer science.	parallel structure	435	3.95	1.86
17: They need some practical skills that can [works] for their lifelong needs.	verb form	614	5.58	1.49
18: They were in a bad situation, and they [want] to find a way out of it.	Tense	568	5.16	1.77
19: In [] year 1994, Australia was visited by a large number of tourists.	Article	420	3.82	1.66
20: There were over 4 [millions] tourists.	Plural	495	4.50	1.89

*Errors are placed in brackets. Empty brackets mean something is missing.

In addition to the evaluation of the errors on the 7-point Likert scale, the teachers were asked to provide a brief explanation for why they had rated an error as totally unacceptable. Their comments are summarized as follows:

1. Sentence Intelligibility:

- “the error has affected the meaning”.
- “the error impedes communication”.
- “the error has caused misunderstanding”.
- “the error has caused different interpretations”.
- “the meaning is ambiguous”.

2. Sentence function:

- “the error has changed the function of the sentence”.

3. Basicness of the rule:

- “the correct form should have been mastered through repetition”.
- “the sources teaching this type of error are easily available”.
- “the error indicates lack of basic grammar knowledge”.

4. Typicality of the error:

- “as it is a common error among Iranian EFL learners, it deserves attention”.
- “this type of error has been fossilized”.

Research Question 2

The Spearman rank correlation coefficient was performed to see if there was any significant relationship between teachers' age and their perception of error gravity. The results indicated that there was no significant relationship between the two variables (Table 8).

Table 8. *The Correlation Between Teachers' Age and Their Perception of Error Gravity*

			Total Score	Age
Spearman's rho	Total Score	Correlation Coefficient	1.00	0.15
		Sig. (2-tailed)	-	0.12
		N	110	110

To investigate if there was any significant relationship between teachers' gender and their error gravity perception, the Pearson correlation coefficient was performed. The results indicated that there was no significant relationship between the two variables (Table 9).

Table 9. *The Correlation Between Teachers' Gender and Their Perception of Error Gravity*

		Total Score	Gender
Total Score	Pearson Correlation	1	-0.15
	Sig. (2-tailed)	-	0.12
	N	110	110

Another Spearman rank correlation coefficient was conducted to examine if there was any significant relationship between teachers' academic degree and their error gravity perception. The results indicated that there was a significant relationship between the two variables (Table 10). The strength and the direction of the relationship between the two variables were weak and positive respectively.

Table 10. *The Correlation Between Teachers' Academic Degree and Their Perception of Error Gravity*

			Total Score	Academic Degree
<i>Spearman's rho</i>	<i>Total Score</i>	<i>Correlation Coefficient</i>	1.000	0.221
		<i>Sig. (2-tailed)</i>	-	0.020
		<i>N</i>	110	110

To find out how much of the variance in the total scores was caused by its relationship to the teachers' academic degree Coefficient of Determination (r^2) was calculated. The results revealed that teachers' academic degree was responsible for 4.88% of the variance in the total scores.

The Spearman rank correlation coefficient was performed to determine if there was any significant relationship between teachers' years of teaching experience and their error gravity perception. The results indicated that there was a significant relationship between the two variables (Table 11). The strength and the direction of the relationship between the two variables were weak and positive respectively.

Table 11. *The Correlation Between Teachers' Years of Teaching Experience and Their Perception of Error Gravity*

			Total Score	Years of Teaching Experience
<i>Spearman's rho</i>	<i>Total Score</i>	<i>Correlation Coefficient</i>	1.000	0.240
		<i>Sig. (2-tailed)</i>	-	0.011
		<i>N</i>	110	110

To determine how much of the variance in the total scores was caused by its relationship to the teachers' years of experience, the coefficient of determination (r^2) was calculated. The results revealed that teachers' years of experience was responsible for 5.76% of the variance in the total scores.

To inspect if there was any significant relationship between the highest taught level and teachers' perception of error gravity, the Spearman rank correlation coefficient was performed. The results indicated that there was a significant relationship between the two variables (Table 12). The strength and the direction of the relationship between the two variables were moderate and positive respectively.

Table 12. *The Correlation Between the Highest Taught Level and Teachers' Perception of Error Gravity*

			Total Score	The Highest Taught Level
<i>Spearman's rho</i>	<i>Total Score</i>	<i>Correlation Coefficient</i>	1.000	0.328
		<i>Sig. (2-tailed)</i>	-	0.000
		<i>N</i>	110	110

To see how much of the variance in the total scores was caused by its relationship to the highest taught level, the coefficient of determination (r^2) was measured. The result

revealed that the highest taught level was responsible for 10.82% of the variance in the total scores.

To prioritize the errors in terms of their acceptability level, the total score for each pair of questions representing the same type of error was calculated. The results revealed that the teachers' evaluations of the grammatical errors built a hierarchy in which errors were placed at different levels in accordance with their level of acceptability (Figure 2). The errors in the pronoun category were evaluated as the least acceptable errors and the errors in the article category as the most acceptable ones (the higher the score, the less acceptable the error).

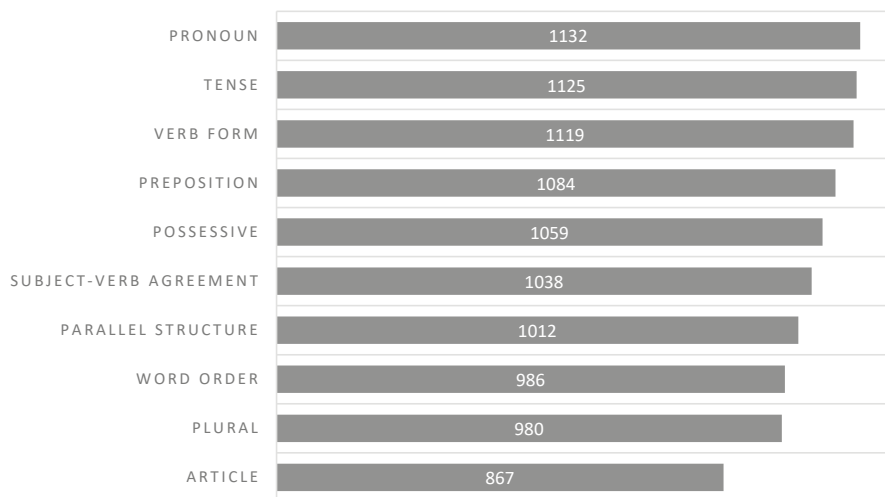


Figure 2. Grammatical Errors Prioritized Based on Teachers' Evaluations

Discussion

Considering the uniqueness of the EFL context and teachers' dependence on different criteria for their evaluation of error gravity, the present study focused on a group of non-native teachers, that is, Iranian EFL teachers and their perception of the gravity of written grammatical errors made by a group of Iranian EFL learners at the sentence level, in terms of the criterion of acceptability. Moreover, to see what factors might influence teachers' evaluations, it investigated the possible correlation of variables such as age, gender, academic degree, years of teaching experience, and the highest taught level with the teachers' judgements.

According to the results, it is possible to conclude that pronoun, tense, verb form, and preposition errors are the less tolerated types of errors. However, how to rank them from one to four, based on their level of acceptability, depends on the categories of errors, the criterion of evaluation, and the context of assessment. An interesting point about the responses submitted by the teachers, was the way that they treated the two examples of the same type of error. In fact, there were some errors, such as parallel structures (questions 8 and 16), for which the presented examples were rated considerably differently. For instance, question 16 "... rich in high technology and a pioneer ..." was rated more acceptable than question 8 "... by playing or go shopping ...". The difference between the scores assigned to these errors was 142. Such a drastic difference between the two

examples can be attributed to the typicality of the first error. This might have caused the teachers to have less severe attitude towards the acceptability of this type of error.

The sentences representing verb form errors (questions 13 and 17) experienced the same difference in scoring. In fact, question 13 "... require people being in touch ..." was rated 109 points below question 17 "... can works ...". This difference could be justified by the fact that errors such as question 17 are mostly considered an infringement of the basic rules that should have been mastered at early stages of language learning process. The same justification can be applied to the questions 6 and 1, which were two examples of word order errors. Since using inversion in question formation is acquired earlier than the type of inversion required by the correlative conjunction "not only ... but also", question 6 "What kind of job they are trying to do?" was rated 108 points above question 1 "Not only it would provide them with financial support ...". This finding is consistent with that of Hughes and Lascaratou's (1982) study where the non-native teachers based their evaluations on "the basicness of the rules infringed" (p. 177).

Factors Affecting Teachers' Perception of Error Gravity

Among the five variables of the study, the three variables of academic degree, years of experience, and the highest taught level had significant relationships with the teachers' perception of the gravity of grammatical errors. Although the present study did not find any relationship between teachers' age and their evaluations of the gravity of grammatical errors, researchers such as Vann et al. (1984) believed that age is one of the factors that might have an effect on the teachers' assessments. The results also indicated that the more experienced the teachers, the higher the scores assigned by them. This finding corroborates the result of Oliaei and Sahragard's (2013) study. In their study, "inexperienced teachers appeared to be more tolerant of error, both in theory and practice, than their experienced counterparts" (Oliaei & Sahragard, 2013, p. 55).

Among the three variables of academic degree, years of experience, and the highest taught level, the highest taught level had stronger relationship with the teachers' evaluations of the grammatical errors. This finding can be attributed to the teachers' "linguistic intuitions" (Davies, 1983, p. 307), meaning that there are some errors that teachers expect learners to master at early stages of language learning process; therefore, learners' failure to use the structures correctly at later stages is considered unacceptable. The descriptive statistics of the responses submitted by the teachers who taught the C2 proficiency level revealed that the teachers rated the failure to use the correct form of verb after modals, which is a type of verb form error, as the least acceptable type of error. This result was partly predictable since in C-level classes, students are expected to be able to use all forms of modal verbs flawlessly and not to make errors such as 'can works' in their writings.

Underlying Reasons Explaining the Teachers' Evaluations

In addition to the evaluation of the errors on a 7-point Likert scale, the teachers were asked to provide a brief explanation for why they had rated an error as totally unacceptable. Their comments are summarized as follow:

Sentence Intelligibility:

- "the error has affected the meaning".
- "the error impedes communication".
- "the error has caused misunderstanding".
- "the error has caused different interpretations".
- "the meaning is ambiguous".

Sentence function:

- “the error has changed the function of the sentence”.

Basicness of the rule:

- “the correct form should have been mastered through repetition”.
- “the sources teaching this type of error are easily available”.
- “the error indicates lack of basic grammar knowledge”.

Typicality of the error:

- “as it is a common error among Iranian EFL learners, it deserves attention”.
- “this type of error has been fossilized”.

These findings hold critical insights for EFL teachers. Analysis of the participant's comments uncovered instances where some teachers struggled to identify why certain marked errors in the questionnaire were classified as errors. For example, participants questioned the correctness of phrases like "... in different types of paid works ..." and "... showed an interest to ...". This lack of understanding could suggest a low proficiency. Of course, almost all of the errors committed by the learners were grammatical; therefore, only grammatical errors were selected for evaluation. Furthermore, the research specifically involved Iranian EFL teachers, limiting the generalizability of the results to a broader spectrum of EFL teaching contexts. Therefore, for future studies, interested researchers should consider expanding the scope of participants to include EFL teachers from a wider range of cultural and linguistic backgrounds. This broader perspective could provide a more comprehensive understanding of the challenges faced by EFL teachers globally and help in developing more inclusive teaching strategies. Additionally, future studies should incorporate a mix of error types beyond just grammar, such as vocabulary or pronunciation errors, to capture a more holistic view of language learning difficulties.

Conclusion

The results of the present study revealed that the general tendency of the teachers in their evaluations was towards negative responses. Moreover, their evaluations of the grammatical errors built a hierarchy in which errors were placed at different levels in accordance with their level of acceptability. In this hierarchy, pronouns were rated as the least acceptable and articles as the most acceptable types of errors. Among the five variables of the study, the three variables of academic degree, years of experience, and the highest taught level had significant relationship with teachers' perception of the gravity of the grammatical errors in terms of the criterion of acceptability. This meant that with an increase in one of the mentioned variables, the scores assigned by the teachers increased as well. Analysis of the comments submitted by the teachers revealed that in addition to the degree of deviation from the English language standards, other factors such as intelligibility of the sentence in which the error was embedded and the basicness of the grammatical structure had an effect on the teachers' evaluations of the gravity of grammatical errors in terms of the criterion of acceptability.

Moreover, the comments submitted by the teachers suggested that learners' level of language proficiency was the most determining factor in assigning the errors different levels of acceptability. Although the questionnaire did not specify any level for the learners who had committed the errors, there were teachers who tried to assume a level for them, since they believed that evaluating the acceptability level of an error without knowing learners' level of language proficiency would be impossible. Determining the acceptability level of an error based on learners' language level is not necessarily reliable. To determine the gravity of an error based on the criterion of acceptability, teachers should consider the degree to which a particular structure deviates from the target

language norms (Burt, 1975). If this deviation is not significant, teachers should not rate an error as totally unacceptable just because they believe that the structure should have been mastered earlier. For instance, the error in question seven “It has lots of effects on their self-confidence and make them responsible people.” is merely missing third person singular ‘s’; however, it was rated unacceptable by 36 teachers and totally unacceptable by 15 teachers. The most interesting comment submitted by very few of the teachers was that “error is an error”. The results of the present study showed that grammatical errors are of different levels of significance. Therefore, teachers should differentiate their approach to errors based on their level of importance and not apply the same level of strictness to every single language error.

References

- Brown, J. D. (1991). Do English and ESL faculties rate writing samples differently? *TESOL Quarterly*, 25(4), 587–603. <https://doi.org/10.2307/3587078>
- Burt, M. K. (1975). Error analysis in the adult EFL classroom. *TESOL Quarterly*, 9(1), 53–63. <https://doi.org/10.2307/3586012>
- Davies, E. E. (1983). Error evaluation: The importance of viewpoint. *ELT Journal*, 37(4), 304–311. <https://doi.org/10.1093/elt/37.4.304>
- Ellis, R. (1986). *Understanding second language acquisition*. Oxford University Press.
- Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, 1(1), 3–18. <https://doi.org/10.5070/l2.v1i1.9054>
- Endley, M. J. (2015). Perceptions of error gravity among native speaking and non-native speaking faculty at UAEU. In *Proceedings of the 21st TESOL Arabia Conference: Theory, Practice, Innovation, Deira, Dubai, 12-14 March 2016*. Dubai: TESOL Arabia Publications.
- Ferris, D. (2002). *Treatment of error in second language student writing*. The University of Michigan Press.
- Hughes, A., & Lascaratou, C. (1982). Competing criteria for error gravity. *ELT Journal*, 36(3), 175–182. <https://doi.org/10.1093/elt/36.3.175>
- Hyland, K., & Anan, E. (2006). Teachers' perceptions of error: The effects of first language and experience. *System*, 34(4), 509–519. <https://doi.org/10.1016/j.system.2006.09.001>
- James, C. (1977). Judgements of error gravities. *ELT Journal*, 31(2), 116–124. <https://doi.org/10.1093/elt/XXXI.2.116>
- Lennon, P. (2008). Contrastive analysis, error analysis, interlanguage. In S. Gramley, & V. Gramley (Eds.), *Bielefeld introduction to applied linguistics* (pp. 51–60). Aithesis.
- Oliaei, S., & Sahragard, R. (2013). Error treatment by experienced and inexperienced Iranian EFL teachers of writing. *Educational Research International*, 2(1), 55–75.
- Rao, Z., & Li, X. (2017). Native and non-native teachers' perception of error gravity: The effects of cultural and educational factors. *The Asia-Pacific Education Researcher*, 26(1-2), 51–59. <https://doi.org/10.1007/s40299-017-0326-5>
- Rao, Z., & Liu, F. (2020). An investigation into native and non-native teachers' assessment of Chinese students' English writing. *Southern African Linguistics and Applied Language Studies*, 38(2), 152–166. <https://doi.org/10.2989/16073614.2020.1805776>
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.
- Rifkin, B., & Roberts, F. D. (1995). Error gravity: A critical review of research design. *Language Learning*, 45(3), 511–537. <https://doi.org/10.1111/j.1467-1770.1995.tb00450.x>
- Sheorey, R. (1986). Error perceptions of native-speaking and non-native-speaking teachers of ESL. *ELT Journal*, 40(4), 306–312. <https://doi.org/10.1093/elt/40.4.306>
- Shirazizadeh, M., & Amirfazlian, R. (2019). Grammatical accuracy in L2 writing: Comparing different types of electronic corrective feedback. *Language Related Research*, 10(2), 1–24.
- Shultz, K. S., Whitney, D. J., & Zickar, M. J. (2014). *Measurement theory in action* (2nd ed.). Routledge.
- Vann, R. J., Meyer, D., & Lorenz, F. O. (1984). Error gravity: A study of faculty opinion of ESL errors. *TESOL Quarterly*, 18(3), 427–440. <https://doi.org/10.2307/3586713>