Relationship between grammatical and pragmatic competence in EFL of Spanish learners through a computer adaptive test

Relación entre competencia gramatical y pragmática en EFL de estudiantes españoles a través de una prueba adaptativa por computadora

Explorar la relació entre els nivells de competència pragmàtica i domini de l'anglès dels estudiants espanyols d'ALE a través d'un test adaptatiu per ordinador

Beatriz Martín Marchante*

Universidad Politécnica de Valencia, Valencia, España
*Corresponding author: beamarm2@upvnet.upv.es (Beatriz Martín Marchante)

ABSTRACT: The work at hand is part of a wider study by Marchante (2015), the objective of which was to study the way pragmatics assessment is approached in the Oxford Online Placement Test (OOPT) which is, currently, one of the most profusely administered computer adaptive tests (CAT). This paper aims, now, to find out the degree of difficulty of the pragmatic items in the OOPT and how much importance is given to pragmatic competence in this test. Furthermore, we analyze the relationship between grammatical and pragmatic competence in EFL of a group of Spanish undergraduate students, who took this test. To this end, a descriptive analysis and a multiple linear regression analysis was carried out. The results showed that the quality of the pragmatic items in the OOPT could be improved. Also, we have found that the pragmatic items were somehow related to the student's final score but not so much as other items of the test, such as those related to grammar and reading comprehension. Additionally, the results indicated that there was no correlation between the grammatical and the pragmatic competence of this group of students.

KEYWORDS: pragmatics; English as a foreign language (EFL); computer adaptive test

RESUMEN: Este trabajo es parte de un estudio más amplio (Marchante, 2015) cuyo objetivo era, entre otros, analizar los ítems de pragmática en el Oxford Online Placement Test (OOPT) para llegar a algunas conclusiones sobre su grado de calidad. El presente artículo tiene, ahora, como objetivo averiguar qué peso tiene la competencia pragmática en la evaluación global del OOPT, ya que esta información no es accesible, a pesar de se trata de una de las pruebas de lengua adaptativas que con más profusión se administran actualmente. Además, analizamos la relación entre el nivel de dominio y el nivel de competencia pragmática en EFL de los mismos examinandos, un grupo de estudiantes universitarios españoles. Con este fin, se realizó un análisis de regresión lineal múltiple y los resultados mostraron que los ítems de pragmática estaban de alguna manera relacionados con la puntuación final del estudiante, pero no tanto como otros ítems de la prueba como, por ejemplo, los relacionados con la gramática y la comprensión lectora. Además, los resultados indicaron que no había relación entre el nivel de dominio y la competencia pragmática de este grupo de estudiantes.

PALABRAS CLAVE: pragmática; Inglés como lengua extranjera (ILE); pruebas de lengua adaptativas

RESUM: El treball que ens ocupa és part d'un estudi més ampli de Marchante (2015), l'objectiu del qual va ser estudiar com s'aborda l'avaluació pragmàtica en l'Oxford Online Placement Test

(OOPT), que és actualment un dels programes informàtics més utilitzats en les proves adaptatives (PA). Aquest article té l'objectiu de conèixer el grau de dificultat dels ítems pragmàtics en l'OOPT i quina importància es dóna a la competència pragmàtica en aquesta prova. A més l'article analitza la relació entre competència gramatical i pragmàtica en ALE d'un grup d'estudiants espanyols de pregrau que van realitzar aquesta prova. Per fer-ho es va fer una anàlisi descriptiva i una anàlisi de regressió lineal múltiple. Els resultats van mostrar que la qualitat dels ítems pragmàtics en l'OOPT podria millorar-se. A més, hem trobat que els ítems pragmàtics estaven relacionats d'alguna manera amb la puntuació final de l'estudiant, però no tant com amb altres ítems de la prova, com els relacionats amb la gramàtica i la comprensió lectora. D'altra banda, els resultats van indicar que no hi havia correlació entre la competència gramatical i pragmàtica d'aquest grup d'estudiants.

PARAULES CLAU: pragmàtica; Anglés com a llengua estrangera (EFL); proves de llengua adaptatives

---

Practitioner Notes

What is currently known about the object of this research?

- Despite the importance of pragmatic competence in the field of communicative teaching of second languages, the evaluation of this competence is still a neglected area in many of the commercial and computer adaptive tests (CAT), such as the Oxford Online Placement test (OOPT), and in some cases it is even non-existent.
- Furthermore, we have noticed that only a small number of studies have examined the relation between the level of proficiency in English and the pragmatic competence of EFL learners. The studies on this issue can be grouped into two blocks, on the one hand, those that result in a positive correlation, and on the other hand, those that show that the level of linguistic competence is not a sufficient condition to determine the level of pragmatic competence.

What does this article contribute as original?

- According to the results in this study, we cannot confirm that there is a relationship between the level of linguistic competence and the level of pragmatic competence of the participants in this research. This contradicts the results obtained in some previous research. Moreover, the results in our study also indicate that the weight of pragmatic competence in the configuration of the final score in the OOPT has been proved to be low.

What are the implications of this work for practice and future policy?

- High stake tests have the capacity to influence the programs and methodologies applied in the classroom, consequently, an effective systematic teaching of pragmatics throughout the different stages of education depends to a great extent on the way pragmatic items are dealt with in these influential types of tests. Hence, further research and work needs to be done which focus on the way pragmatics is being taught in the classroom and also on the improvement of construct validity of pragmatic items in high stake EFL tests, and more specifically, in CATs.

---

1. INTRODUCTION

In this study we focus on the analysis of the pragmatic subtest of the OOPT, which is administered by the Faculty of Teacher Training (University of Valencia, Spain) to undergraduate students who need to certify their level of proficiency in English as a foreign language. After scoring a group of test takers, some uncertainties arose regarding the quality of the items in the test, since many examinees complained about their difficulty, and they disagreed with the scores obtained. We got the impression that those items which the students found more difficult were the ones corresponding to the pragmatics part of the test. In view of this, our objectives were the following: to analyze the pragmatics items validity; to see how much importance is given to pragmatic competence in the OOPT, and finally, we aimed to explore the relationship between grammatical and pragmatic competence of the participants in this study. To

present the answer to these questions, this paper is organized as follows. Firstly, the theoretical background considered in this study is offered. Secondly, the students involved in the project, the materials and the method used are explained. Thirdly, the results extracted from our analysis are shown and finally, some conclusions are drawn.

2. PRAGMATIC COMPETENCE IN COMPUTER ADAPTIVE TESTS

Pragmatics, the systematic study of communicative effectiveness, as defined by Reyes (1990), is also a discipline that "analyzes the difference between what is codified and what is transmitted by the speaker" (Yule, 1996). In addition, as Escandell (2004) emphasizes, pragmatics helps describe the rules and principles, which are in force when speakers communicate although they are usually unaware of. There are, therefore, other guidelines in addition to grammar rules which determine the adequacy of the linguistic use. Gutiérrez (2004), states that pragmatics is presented as an integrating discipline that explains concepts such as politeness, speech acts or implicit meaning, and evince the inferential process in a new communicative dimension. In the field of second language learning and teaching, pragmatic competence is at present considered as a sub-competence within linguistic competence (Bachman, 1990; Bachman & Palmer, 1996; Purpura, 2004). In addition, numerous researchers (Allami & Naeimi, 2011; Carrió-Pastor, 2016; Kasper & Rose, 2001; Martínez-Flor & Usó-Juan, 2010; Safont-Jordá, 2005; A. Soler, E., Flor, & A., 2008; E. A. Soler & Martínez-Flor, 2005; S. Takahashi, 2010) show the possibilities of teaching pragmatics in the English as a Foreign Language (EFL) classroom.

Taking this into consideration, it would not be wrong to think that the evaluation of pragmatic competence should, consequently, be included in all EFL domain tests.

But despite the importance of pragmatic competence in the field of communicative teaching of second languages, the evaluation of this competence is still a neglected area in many of the commercial and computer adaptive tests, and in some cases, it is even non-existent.

Furthermore, some of these tests do not contemplate the interactive nature of the language or its purpose. In the same way, contextualization is often poorly defined. It is also noticeable that the only commercial CATs which contain specific pragmatic items are the Test of English as a Foreign Language (TOEFL) in its reading comprehension section, and the Oxford Online Placement Test (OOPT), with a pragmatic subtest (Marchante, 2015). Another characteristic of the tests previously detailed is that they contain closed, multiple-choice and completion type items. This type of items has been called into question by some authors (Bachman, 1990; Carroll, 1980a; Fulcher, 2015; Morrow, 1981; Yamashita, 2008) (among others) since, according to them, these items do not measure the genuine communicative performance, but rather something that resembles it and is still artificial. Furthermore, there are some concerns about the construct validity in the field of pragmatics assessment. It is thought to be underrepresented since only the following components of the pragmatic competence are usually chosen as the framework for the validation of the construct (Grabowski, 2009; Messick, 1989; Roever, 2006):

1. Knowledge of speech acts and their strategies.

2. Interpretation of implicatures.

3. Recognition of formulaic expressions and routines.

So, some other aspects such as production and recognition of speech styles, contextualization cues, discourse structure, sequence organization or the effect on interlocutor are hardly ever the object of studies in this field (Roever, 2011).

So far, we have seen some issues related to the treatment of pragmatic competence in the most widely administered CATs today, as well as some aspects that affect the validity of this construct. However, we must add to this some other facets that affect the design of the EFL domain tests. One of these aspects is their psychological effect on the examinees since this type of tests are often seen as a threat rather than an opportunity to learn and improve (Linn, 2000; Martínez-Rizo, 2008; Shepard, 2006). Furthermore, the unexpected negative effects of the use of high-stake tests as well as the necessary rendering of accounts are often more important than the positive effects that are intended.

Despite everything, in the educational environment, teaching and language tests are aspects of the same basic problem according to Oller (1979), who states that apparently both fields are quite different although beyond this appearance there is a very important similarity, actually they share identity and purpose. On the other hand, language tests can be a source of information on the effectiveness of learning and teaching and serve to evaluate the effectiveness of different approaches in language teaching. Some scholars (Shepard et al., 2005) go beyond this fact and claim that external tests can reshape the curriculum and can also have profound effects on classroom practices. This is known as the *back-wash effect* or the influence of testing on teaching and learning (Bachman & Palmer, 1996) . Likewise, they maintain that between the external tests and the evaluation in the classroom should, in principle, there be a coherent link with the same fundamental model of learning. This approach finds its theoretical basis in what Carroll (1980b) coined as *curriculum triangle* (see Figure 1) to show the relationship established between the language programs and the measurement system and how they are derived from an analysis of learner´s communication needs.
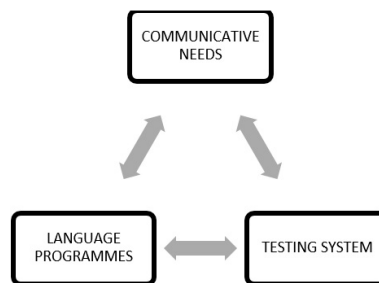

Figure 1. Curriculum triangle (Carroll, 1980b)

The Council of Europe project for adult language learning is an example of this systematic approach since it considers the purposes for which the members of the European community will most likely communicate since courses and evaluation are developed from them. However, the Common European Framework of Reference for Languages (CERF) (EU, 2001) does not specify how to include relevant topics of pragmatics in the curriculum. Examples of this are the questions formulated in the CEFR (EU 2001, 154) concerning the development of pragmatic competences:

The development of the learner's **pragmatic competences** should be

1. Assumed to be transferable from education and general experience in the mother tongue (L1)? Or facilitated:

2. By progressively increasing the complexity of discourse structure and the functional range of the texts presented to the learner?

3. By requiring the learner to produce texts of increasing complexity by translating texts of increasing complexity from L1 to L2?

4. By setting tasks that require a wider functional range and adherence to verbal exchange patterns?

5. By awareness-raising (analysis, explanation, terminology, etc.) in addition to practical activities?

6. By explicit teaching and exercising of functions, verbal exchange patterns and discourse structure?

Users of the Framework may wish to consider and where appropriate state

- To what extent sociolinguistic and pragmatic competences can be assumed or left to develop naturally.

- What methods and techniques should be employed to facilitate their development where it is felt to be necessary or advisable to do so. (p.154)

The CEFR points to pragmatics as part of communicative competences, dedicating several sections to it in the document but leaves it out of the evaluation in the ninth chapter. This lack of precision is due to the fact that the Council of Europe has no power to impose anything on any member of state and that it is only a point of reference and "not a means for coercing teachers, nor even a basis for measures of accountability", as Alderson (2007) notes. This author (ibid.) highlights that one of the objectives of the CEFR is to broaden the understanding of what language teaching and learning entails and to gather thought and research on languages under one umbrella and to contribute to the improvement of understanding about what teaching, learning, and evaluating a foreign language is. If we ponder the aforementioned, investigating how the evaluation of pragmatic competence is carried out in the main high-stake tests can give us some clues about the importance that is currently given to the evaluation of pragmatic competence in the classroom and also about the importance given to it in the curricula.

Another aspect worth considering is the one that draws our attention in this study, which is the degree of relation between the level of grammatical and pragmatic competence of EFL learners. We noticed that only a small number of studies have examined this correlation (Carrió-Pastor & Casas-Gómez, 2015; Carrió-Pastor & Marchante, 2018; Tsutagawa, 2013). The studies on this issue can be grouped into two blocks, on the one hand, those that result in a positive correlation, and on the other hand, those that show that the level of linguistic competence is not a sufficient condition to determine the level of pragmatic competence. Norris (2001), when studying the correspondence between the correct use of the forms of treatment and the ability in the language of a group of students, confirms that such correspondence does not occur. Roever (2006), on the other hand, discovers that what defines a higher pragma-linguistic competence is the level of proficiency, rather than the time of exposure to the English language. Jian-Da (2007); Liu (2006), in line with Norris (2001), shows that students with high scores in the TOEFL do not seem to have a corresponding high pragmatic competence in their interlanguage. Therefore, according to the author, it is not surprising that some students who score more than 670 points on a traditional TOEFL test cannot communicate well in English. However, Grabowski (2009), who uses Purpura's model (Purpura, 2004), confirms in his results that, there is a relation between the level of proficiency in EFL and the score in the test of pragmatic competence. The study carried out by Ifantidou and Tzanne (2012) also shows that pragmatic competence correlates positively with the level of competence in English. Yet, in different scenarios, motivation can invalidate the domain or vice versa, since it seems that more motivated learners with a higher level of proficiency may be superior in pragmatic awareness than those with less motivation and mastery. As stated by Xiao

(2015), many factors can underlie the differing results of the above-mentioned studies, for instance, the nature of target pragmatic features to be measured such as types of speech acts, modalities of pragmatic performance (comprehension and production), social variables involved in task situations (social status or distance), power relationship, or length of stay in the target language community.

In the following sections, the method, the participants, and the procedure which have been carried out are explained.

3. METHOD

Our interest in studying the OOPT emerged when a group of students complained after taking the test and obtaining their results. They found some items very difficult and they disagreed with the scores obtained. Regarding this, our hypothesis was that those items were the ones corresponding to the pragmatic part. Consequently, the following research questions arose:

- Do the pragmatic questions in the OOPT hold an outstanding frequency of wrong answers which could indicate that further analysis of tests items is needed?

- What is the contribution of the pragmatic part to the configuration of the final score?

Moreover, another question led the main objective of this study:

- What is the magnitude of the correlation between the grammatical and pragmatic competence in English of the learners involved in this study?

3.1. *Subjects*

The participants in this study were 34 Spanish undergraduate students at the Faculty of Teacher Training (University of Valencia, Spain). The sample was homogeneous in terms of age, between 18 and 25 years old, and in terms of proficiency in English since all the examinees ranged from A2 and B1 levels acquired through formal education. Four of the participants were men and 40 were women, so we can say that in this study the anthropometric variables have little variability. Only samples of the students whose results ranged between A2 and B1 levels of proficiency in English were selected for the analysis of the test because B1 was the expected level to be reached by those who did not have obtained it by the end of the first academic year.

None of the students had stayed in the target language community, so, they were provided with culture-rich, interactive materials to foster spoken interaction and spoken production. The teaching methodology implemented throughout the course was based on a pragmatic approach which supported learning awareness rising, in such a way that the students were aware of the importance of the context in the verbal communicative act and the significance of the inferential process in the decoding of implicit meanings.

3.2. *Materials*

The OOPT is an adaptive test that assesses the level of proficiency in English and it is distributed by Oxford University Press. The main objective of the OOPT is not only to measure the grammatical or lexical competence, but also the ability of the examinee to understand many grammatical forms and the meanings that they transmit

in different contexts. It also measures to what extent learners can use these language resources to communicate in English language situations (Purpura, 2009). The OOPT consists of two main parts, one focused on the use of English and another focused on the oral comprehension. The first contains up to 30 questions and evaluates the vocabulary, grammar and reading comprehension. In addition, a pragmatic subtest is included in this part, which contains 12 items. As the OOPT is adaptive, the number of items in each part of the test can vary depending on the examinees. It should be remembered that, due to this fact, it was not possible to perform an item analysis through which to discern the coefficients of difficulty and discrimination. But instead, a descriptive analysis was carried on. All the items in the OOPT are *Written Discourse Completion Task* (WDCT) *and Multiple-Choice Completion Task* (MCCT). Prior research by Carrió-Pastor and Marchante (2018) showed that the categories under which the pragmatic items were classified were mostly lexicalized trope-inferences, implicatures and indirect speech acts. The following are a few examples:

Lexicalized trope-inferences[1]

Man: It's high time our son got down to doing his schoolwork.

Woman: Don't hold your breath. He hasn't shown any interest in school for months now.

What does the woman mean?

A It's going to be very difficult to make him do any schoolwork.

B I don't think you should try to make him do his schoolwork.

C There is no chance that he will do any schoolwork.

Implicature[2]

Woman: I just don't know what's the matter with me. My boss has invited me to a Paris fashion show and I'm struggling to get excited about it.

Man:You'd normally go without a second thought.

What does the man mean?

A I think you should go.

B It's odd that you should feel like that.

C You should think about this more seriously.

Indirect speech act[3]

Man: It's a nice day for a drive in the countryside.

Woman: Yes, but I'm a bit busy this afternoon.

The man is . . .

A suggesting a trip to the countryside.

B agreeing to drive the woman to the countryside.

---

[1]Lexicalized Trope Inferences are groups of fixed words the meaning of which meaning is usually known by any competent speaker and usually does not come exclusively from the isolated meanings of the words that constitute them (Gallardo-Paúls, 2005)

[2] Gallardo-Paúls (1995) defines implicatures as a kind of nonconventional inferential meaning which is based not on the use of certain words, but on the application of certain communicative norms. They have social character and are outside the structures of language (inferences). According to this author (Gallardo-Paúls, 1995, 2005) , an implicature is the meaning that is not said but inferred from a statement.

[3]Speech acts are, according to Austin, "[...] functional units of communication that have prepositional or locutionary meaning (the literal meaning of the utterance), illocutionary meaning (the social function of the utterance), and perlocutionary force (the effect produced by the utterance in a given context." (Austin 1962in Cohen 1996). When there is a direct relationship between structure and function, we have direct speech acts, but when this relation is indirect, and the transmission of meaning does not coincide with the illocutionary act, we find ourselves facing indirect speech acts.

C finding out about the weather in the countryside.

As stated by Purpura (2009), it is essential to know the grammatical and pragmatic capacity of students if we intend to help them improve in these skills. For this reason, a part that measures students´ knowledge about pragmatic meanings encoded in different types of interaction is included in the test.

3.3. *Data analysis*

To address the data analysis, the first drawback we had to tackle with was the impossibility to carry out an item analysis to get the discrimination and difficulty indexes of the pragmatic items due to the adaptive format of the OOPT. So, the first research question had to be formulated as follows: "Does the pragmatic part in the OOPT hold an outstanding frequency of wrong answers?" To resolve this question a descriptive analysis of the erroneous answers by the examinees in each of the test items was carried out. Then, the results obtained were described, both, globally and by parts. Four blocks or parts totaling 44 items have been distinguished in the analysis, the first corresponds to grammar items, the second corresponds to pragmatics, the third one is devoted to writing skills and the fourth is the reading comprehension part. The second and third research questions were: "What is the contribution of the pragmatic part to the configuration of the final score?" and "What is the magnitude of the correlation between the grammatical and pragmatic competence of the participants in this research?" Therefore, to give an answer to these questions the relative weight of each block in the configuration of the final score was quantified to find out the degree of contribution of the pragmatic items in the OOPT final score.

The analysis, carried out to identify the blocks that most contribute to the configuration of the final score, consisted in the estimation of a multiple linear regression model with dependent variable, the score of the test, and independent variables, the number of failures (or equivalent, the percentage of failures) in each block of items. The selection of variables was carried out by means of a method of successive steps. The value of the coefficient of determination was deemed as an indicator of the degree of adjustment achieved with the model. The possible collinearity between the independent factors was evaluated, considering the condition and decomposition of the variance indexes.

A preliminary exploratory analysis of the correlation between parameters was carried out using the Pearson or Spearman correlation coefficient, according to the observed sampling distribution and the result of the Kolmogorov-Smirnov test (Martínez, Sánchez, Toledo, & Faulín, 2006). The Kolmogorov-Smirnov test analyzes whether the values of a parameter follow a normal distribution. In our case, we applied Spearman because the Kolmogorov test rejected the normality hypothesis, and the sample (n = 34) was not large enough.

The estimations of the regression coefficients were accompanied by the 95% confidence interval. The theoretical hypotheses of the model were validated, that is, the results met a series of hypotheses: they followed a normal distribution, they had a constant variability (homoscedasticity), and they were uncorrelated between them (this is seen through the Durbin-Watson test). After checking these points, we saw that the model was robust.

4. RESULTS

The 44 items of the test are grouped into four large blocks: the first block comprises the items that evaluate the grammatical form or use of English with multiple-choice questions; the second block is made up of multiple-choice pragmatic items; the third

block is composed of fill-in-the-blank items that assess writing, and the fourth block includes reading comprehension items.

4.1. Results about the first research question

The first research question asked which part of the OOPT held the highest frequency of wrong answers.

The results obtained in the descriptive analysis of failures in the four blocks of questions that make up the OOPT are presented below. The percentage of errors on the total number of items in each block has been calculated (Table 1):

Table 1. *Percentage of wrong responses by blocks and total*

	n	mean	Standard deviation	Minim	Maxim	Median
GRAMMAR (10 items)	34	36,2	11,6	10,0	60,0	40,0
PRAGMATICS (12 items)	34	45,8	16,6	8,3	75,0	50,0
WRITING (5 items)	34	47,1	29,1	,0	100,0	60,0
READING COMPREHENSION (17 items)	34	44,8	11,5	23,5	70,6	47,1
TOTAL (44 items)	34	43,4	8,0	27,3	61,4	43,2

In the grammar or 'use of English' block (Table 1) it is seen that each student fails, on average, 36.2% of the items (s.d. $\pm$ 11.6). The one who fails the least does it in 10% of the items and the one who fails the most, does it in the 60%. Half of the students fail less than 40% of the items.

For a better understanding of the sampling distribution of the parameter 'percentage of errors', in Figure 2, it is observed that in the block of writing, the percentage range is greater (from 0% to 100%). There are only 6 possible values: 0% when the subjects fail the 5 items, 20% if they fail 1, 40%, if they fail 2, 60%, if they fail 3, 80%, if they fail 4, and 100% if they fail the 5 items. According to the graph, in the sample we find all the possibilities. This block corresponds to the highest median (60.0%). That is, half of the examinees fail at least 60% of the items. The pragmatics block corresponds to the second highest median since 50% of the examinees fail at least 50% of the items. Observe that the dispersion is quite broad. It could be interpreted that for this concept there is an extreme variety of levels of competence. There are students who have a great aptitude, compared to others with a very low level. In the reading comprehension block, the median is similar (47.1). However, the range of values is narrower; no such heterogeneous level of response capacity is found.

Finally, the grammar items are revealed as the easiest to answer since only 50% of subjects fail more than 40% (the median) of the items in the block. It is noteworthy that the range of 'normal' values is the narrowest of all blocks. The participants present a more comparable level in terms of the results obtained. Now, there is a peculiar characteristic, it is the only block where atypical cases are identified, one above and one below.

4.2. Results about the second and third research questions

The second and third questions addressed the contribution of the pragmatic part to the configuration of the final score and the relationship between the grammatical and the pragmatic competence of the participants in this research.

A multiple linear regression analysis was performed to determine the relationships between the (final) score, and the number of failures recorded in each block, and the relation between the grammatical, reading and writing levels of competence and the
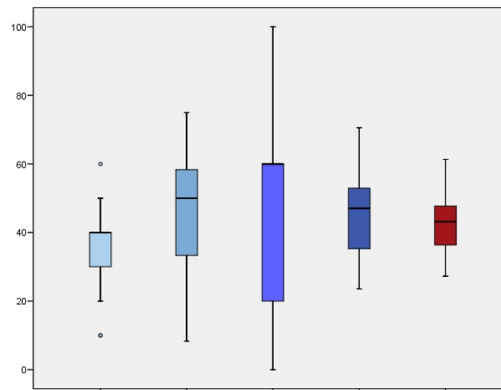
Figure 2. *Grammar Pragmatics Writing Reading comprehension Total. Sample distribution of the parameter 'percentage of wrong responses'.*

pragmatic level. Basically, the issue here is to discern how the final score is configured from the partial ones. To control the effect of the different number of questions in each block, we chose to use the wrong answers percentages rather than the absolute number as predictors. Table 2 shows the matrix of correlations between the complete set of variables, the score, and the predictors. The Spearman's Rho coefficient has been estimated once it has been previously verified that the independent factors do not conform to normality (p-values 0.001, 0.007, 0.019 and 0.113 for the blocks in the usual order, by means of the Kolmogorov test, according to Martínez et al., 2006).

Table 2 shows that the correlation between the score and the block indicators is statistically significant for the block of grammar and reading, showing a strong tendency also for the pragmatic block. However, no type of association with the result of the writing block is noticed. It should be noted that the coefficients are negative; the relationship is inverse because the block indicators are 'failure percentages', that is, the more failures, the lower final score.

The result of the regression model is presented in Table 3. This contains the values of the so-called 'coefficients' of the model that represent the impact on the final score of the changes in the results of the different blocks. Finally, the confidence interval is a kind of 'clamp' for the estimated coefficient.

The model (Table 3) has included the result of reading comprehension and grammar as the only sections that significantly contribute, to the configuration of the final grade. That is, knowing the percentage of errors in these two blocks of questions means having enough information to predict the score with the highest possible reliability. In a first step, the percentage of failures in the reading comprehension block has been introduced as the most decisive aspect, although in the final model it shares a level of importance utterly similar to that of grammar. Note how an increase of one percentage point in the level of failures in the "reading comprehension" block (that is, going from failing x% of the 17 questions to one (x + 1) %) means that the final score is reduced by -0.627 points. This reduction can be bounded with a 95% confidence between -0.874 and -0.380. (Table 3). For the "grammar" factor, the impact of increasing one percentage point the failures on the final score is -0.601. From the table of correlations presented above, we know that the pragmatic section also exhibits a strong association with the score. The current model should not conclude that the real influence of pragmatics is of lesser importance, it is simply that in the presence of the levels of failure in reading comprehension and grammar, addressing the result of pragmatic items would not bring anything new (the explicability of the score would not improve). The data in Table 4 show how the pragmatic block has been excluded from the model by a very narrow margin (p = 0.075).

Table 2. Matrix of correlations

	Correlations		Score	Gramm.	Pragmatics	Writing	Reading
Rho Spearman	SCORE	Correlation coefficient	1,000	-,471(**)	-,299	,012	-,516(**)
		Sig. (bilateral)	.	,005	,086	,945	,002
		N	34	34	34	34	34
	GRAMMAR	Correlation coefficient	-,471(**)	1,000	-,340 (*)	-,029	,040
		Sig. (bilateral)	,005	.	,049	,870	,822
		N	34	34	34	34	34
	PRAGMATICS	Correlation coefficient	-,299	-,340(*)	1,000	-,011	,269
		Sig. (bilateral)	,086	,049	.	,952	,125
		N	34	34	34	34	34
	WRITING	Correlation coefficient	,012	-,029	-,011	1,000	-,025
		Sig. (bilateral)	,945	,870	,952	.	,888
		N	34	34	34	34	34
	READING COMPREH.	Correlation coefficient	-,516(**)	,040	,269	-,025	1,000
		Sig. (bilateral)	,002	,822	,125	,888	.
		N	34	34	34	34	34

** The correlation is significant at the 0.01 level (bilateral).
* The correlation is significant at the 0.05 level (bilateral).

Table 3. Coefficients of the regression model.

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.	95% confidence interval for B	
		B	Standard error	Beta	Lower limit	Upper limit	Lower limit	Upper limit
1	(Constant)	61,705	7,315		8,435	,000	46,804	76,606
	READING	-,717	,158	-,625	-4,528	,000	-1,039	-,394
2	(Constant)	79,443	6,581		12,071	,000	66,021	92,865
	READING	-,627	,121	-,547	-5,179	,000	-,874	-,380
	GRAMMAR	-,601	,121	-,527	-4,986	,000	-,847	-,355

Table 4. *Sections of the test that are excluded depending on whether the model is 1 or 2.*

		Beta IN	t	Sig.	Partial correlation
1	GRAMMAR	-,527(1)	-4,986	,000	-,667
	PRAGMATICS	,014(1)	,094	,926	,017
	WRITING	-,018(1)	-,127	,900	-,023
2	PRAGMATICS	-,207(2)	-1,841	,075	-,319
	WRITING	-,039(2)	-,365	,717	-,067

1.0 Predictor variables in the model: (Constant), READING
2.0 Predictor variables in the model: (Constant), READING,GRAMMAR

The R2 or coefficient of determination achieved with the model obtained takes the value 0.662. That is, up to 66% of all the variability of the score can be explained from the percentage of failures made in the blocks of reading and grammar. Some students have high scores, and others have low ones. That is, the score is presented in the data with a dispersion or variability that can be explained by different factors. Among those that we study here, the results of the different blocks that serve to predict the score or assume part of the responsibility of that final score. The interpretation is that the pragmatic items have a certain relationship with the final score of the student, but they do not improve the prediction that the grammar part and the reading part alone provide. On the other hand, the pragmatic and grammar part exhibit an inverse correlation, that is, if a student is good at grammar, we roughly know that he/she is not so good at pragmatics; therefore, it is not necessary to explicitly incorporate the pragmatic score in the model to better predict the final score. The results of a forced introduction model of all the factors are shown in Table 5.

Table 5. *Coefficients of the four sections of the test.*

	Non standardized coefficients		Standardized coefficients	t	Sig.	95% confidence interval for B	
	B	Standard error	Beta	Lower limit	Upper limit	Lower limit	Upper limit
(Constant)	87,354	7,883		11,081	,000	71,231	103,477
GRAMMAR	-,685	,126	-,600	-5,425	,000	-,943	-,427
PRAGMATICS	-,165	,091	-,207	-1,818	,079	-,351	,021
WRITING	-,018	,046	-,040	-,392	,698	-,113	,077
READING	-,548	,126	-,478	-4,342	,000	-,807	-,290

The situation is totally stable. The R2 of this model is 0.698, that is, with two more factors we hardly increase the explicability of the final score, so the previous, more simplified solution is preferable.

5. CONCLUSIONS

The starting hypothesis at the beginning of this study was that the type of pragmatic items in the OOPT could be negatively influencing the results obtained by the 34 students who took this test, given that, after listening to the students' opinions, the items seemed to be of rather high difficulty. Then, the first research question arose: "Does the pragmatic part in the OOPT hold an outstanding frequency of wrong answers?" To resolve this question a descriptive analysis of wrong answers by the examinees in the test was completed and it was found that the 'writing' part of the test was

the one holding the highest mean of wrong answers, followed by the pragmatics part. Therefore, our hypothesis is null since, although the difficulty of the items is moderate, they are not the most difficult items of the test.

As for the second research question, "what is the contribution of the pragmatic part to the final score", a regression analysis was applied to solve it, which reveals that the pragmatic items have a certain relationship with the final score of the students but do not improve the prediction of the final score that the grammar part and the reading comprehension part alone provide. Therefore, the results do not confirm our starting hypothesis and the null hypothesis must be accepted since we verify that these items hardly influence the final score.

As for the third research question, "What is the magnitude of the correlation between the grammatical and pragmatic competence of the participants in this research?", we can conclude that the part of the test corresponding to the evaluation of pragmatic competence and the grammar part have an inverse correlation, that is, if a student is good at grammar, it can be inferred that he/she is not so much at pragmatics. According to these results, we cannot confirm that there is a relationship between the level of language or grammatical competence and the level of pragmatic competence of this group of students. This contradicts the results obtained in some previous studies which have been discussed in this work which demonstrate the existence of such agreement and note that examinees with greater command of the language obtain better results in a pragmatic test than those with a lower level. Also, other works (Ameriks, 2011; Grabowski, 2009; Liao, 2009) reveal that grammatical competence is a strong predictor of the student's ability to communicate in a pragmatically correct way. However, several investigations (Bardovi-Harlig & Hartford, 1991, 1993; Jian-Da, 2007; Omar, 1990; T. Takahashi & Beebe, 1987) seem to demonstrate the opposite, verifying disparity between grammatical development and pragmatic development of learners. Our work is aligned with the latter ones. The disparity in the results may be due to different factors such as the type of evaluation, research method, different type of items, or to the combination of all. In addition, as Adams (2002) points out, in the field of the evaluation of pragmatic competence, an evolutionary or development-based approach remains conflictive due to complex social interaction; pragmatic functions that vary according to context and audience; to the linguistic and cognitive capacity; to the individual styles of communication, as well as to the cultural influence. For this reason, it is necessary to examine the relative contribution of each of these factors in pragmatic linguistic awareness, as indicated by S. Takahashi (2010). The relationship between language learners' grammatical competence and pragmatic competence has been a hot issue in the literature because determining their relationship helps to know the development itinerary of foreign language learning and the primary emphasis of foreign language teaching.

On the other hand, the lack of correlation between the level of grammar and the participant´s pragmatic competence in this study could be because, pragmatics is not being systematically or sufficiently taught, neither explicitly nor implicitly, unlike grammar. Taking this into consideration, the results are informing us of the need for pragmatics to be taught in a more resourceful and efficient way in the classrooms since, as teachers, we cannot expect this competence to be developed and acquired without any type of instruction, at the expense of the learning and acquisition of grammatical competence. It may be necessary to undertake the teaching of pragmatics more thoroughly, similarly to the new approaches to vocabulary teaching and by adopting, in our methodology; the contributions of the studies that are being carried out in the field of lexical pragmatics "a research field that tries to give a systematic and explanatory account of pragmatic phenomena that are connected with the semantic under specification of lexical items" (Blutner, 1998). Lexical pragmatics investigates the processes involved when the literal meaning of words is modified in use. Examples

of these processes are the conceptual reduction (narrowing), the approximation, and the metaphorical extension. This theory rejects the traditional distinction between literal meaning and figurative meaning and defends that neither metaphor, nor approximation, nor does hyperbole require interpretive mechanisms other than those required by literal or common expressions. On the contrary, they are the result of a single pragmatic process that refines the details of the interpretation of almost every word (Carston, 1997, 2002).

This work has also revealed two facts that should be noticed. On the one hand, it is verified that the incorporation of pragmatic items in the EFL tests is generally deficient, since it has been seen that only two of the numerous current commercial high-stake tests include pragmatic items. This indicates that pragmatics is a competence to which EFL tests still fail to give enough importance. On the other hand, the weight of pragmatic competence in the configuration of the final score in the OOPT has been proved to be low compare to the weight allocated to the grammar part. This fact is paradoxical if one thinks that within the communicative approach, which is currently being applied in most EFL classrooms, students are intended to understand and express themselves adequately in the target language according to different communicative and socio-cultural contexts. We are aware that this research may have some limitations, for instance, the sample size and the nature of pragmatic features analyzed may have slightly influenced the outcome. Yet, the results indicate that it is worth asking whether there is a truly consistent relationship between language programmes and high stake, standardized measurement systems and to what extent they derive from an analysis of the learner's communicative needs. There does not seem to be a coherent link between the external tests and the classroom evaluation with the fundamental model of learning. High stake tests have the capacity to influence the curriculum and methodologies applied in the classroom. This is due to the influential power of this type of tests and their active role in shaping education policies. Consequently, an effective systematic teaching of pragmatics throughout the different stages of education depends, not only, but to a great extent on the way pragmatic items are dealt with in these tests. Hence, further work needs to be done which focus on the way pragmatics is being taught in the classroom and on the improvement of construct validity of pragmatic tests in EFL.

REFERENCES

Adams, C. (2002). Practitioner review: The assessment of language pragmatics. *Journal of child psychology and psychiatry*, *43*(8), 973–987. https://doi.org/10.1111/1469-7610.00226

Alderson, J. C. (2007). The Challenge of (Diagnostic) Testing. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner, & C. Doe (Eds.), *Language Testing Reconsidered.* Ottawa: University of Ottawa Press.

Allami, H., & Naeimi, A. (2011). A cross-linguistic study of refusals: An analysis of pragmatic competence development in Iranian EFL learners. *Journal of Pragmatics*, *43*(1), 385–406. https://doi.org/10.1016/j.pragma.2010.07.010

Ameriks, Y. S. (2011). *Investigating validity across two test forms of the Examination for the Certificate of Proficiency in English (ECPE): A multi-group structural equation modeling approach* (Unpublished doctoral dissertation). Teachers College, Columbia University.

Austin, J. L. (1962). *How to do Things with Words* (J. O. Urmston, Ed.). London: Oxford University Press.

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests.* Oxford: Oxford University Press.

Bardovi-Harlig, K., & Hartford, B. S. (1991). Saying "no" in English: Native and nonnative rejections. *Pragmatics and language learning*, *2*, 41–57.

Bardovi-Harlig, K., & Hartford, B. S. (1993). Learning the rules of academic talk: A longitudinal study of pragmatic change. *Studies in Second Language Acquisition*, *15*(3), 279–304.

Blutner, R. (1998). Lexical pragmatics. *Journal of Semantics*, *15*(2), 115–162. https://doi.org/10.1093/jos/15.2.115

Carrió-Pastor, M. L. (2016). Introduction. In M. L. Carrió-Pastor (Ed.), *New Challenges for Language Testing: Towards Mutual Recognition of Qualifications* (pp. 7–11). London: Cambridge Scholars Publishing.

Carrió-Pastor, M. L., & Casas-Gómez, M. (2015). The Use of Corpora to Identify the Pragmatic Knowledge Associated with Different Levels of Language Proficiency. In S. Gesuato, F. Bianchi, & W. Cheng (Eds.), *Teaching, Learning and Investigating Pragmatics* (pp. 207–230). Cambridge: Cambridge Scholars.

Carrió-Pastor, M. L., & Marchante, B. M. (2018). Analysis of pragmatic items in an ESL online adaptive placement test. *International Journal of English Studies*, *18*(2), 97–117. https://doi.org/10.6018/ijes/2018/2/322581

Carroll, B. J. (1980a). *Testing Communicative Performance. An Interim Study.* London: Pergamon Press.

Carroll, B. J. (1980b). *Testing Communicative Performance. An Interim Study.* Oxford: Pergamon Press.

Carston, R. (1997). Enrichment and loosening: complementary processes in deriving the proposition expressed? In E. Rolf (Ed.), *Pragmatik. Linguistische Berichte (Forschung Information Diskussion)* (Vol. 8, pp. 103–127). https://doi.org/10.1007/978-3-663-11116-0_7

Carston, R. (2002). Linguistic Meaning, Communicated Meaning and Cognitive Pragmatics. *Mind & Language: Special Issue on Pragmatics and Cognitive Science*, *17*(1), 127–148. https://doi.org/10.1111/1468-0017.00192

Cohen, A. D. (1996). Developing the ability to perform speech acts. *Studies in Second Language Acquisition*, *18*(2), 253–267. https://doi.org/10.1017/S027226310001490X

Escandell, M. V. (2004). Aportaciones de la Pragmática. In J. S. Lobato & I. S. Gargallo (Eds.), *Vademecum para la formación de profesores de profesores. Enseñar español como segunda lengua (L2)/lengua extranjera (LE)* (pp. 179–198). Madrid: SGEL.

EU. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Retrieved from https://rm.coe.int/16802fc1bf

Fulcher, G. (2015). *Re-examining language testing: A philosophical and social inquiry.* London: Routledge.

Gallardo-Paúls, B. (1995). El sobreentendido. *Pragmalingüística*, *3*(4), 351–381.

Gallardo-Paúls, B. (2005). El lenguaje y la mente. In A. López & B. Gallardo-Paúls (Eds.), *Conocimiento y lenguaje.* Valencia: PUV.

Grabowski, K. C. (2009). *Investigating the construct validity of a test designed to measure grammatical and pragmatic knowledge in the context of speaking* (Unpublished doctoral dissertation). Columbia University, Columbia USA.

Gutiérrez, S. (2004). La subcompetencia pragmática. In J. S. Lobato & I. S. Gargallo (Eds.), *Vademecum para la formación de profesores de profesores. Enseñar español como segunda lengua (L2)/lengua extranjera (LE)* (pp. 533–552). Madrid: SGEL.

Ifantidou, E., & Tzanne, A. (2012). Levels of pragmatic competence in an EFL academic context: A tool for assessment. *Intercultural Pragmatics*, *9*(1), 47–70. https://doi.org/10.1515/ip-2012-0003

Jian-Da, L. (2007). Assessing EFL learners' interlanguage pragmatic knowledge: Implications for testers and teachers. *Reflections on English Language Teaching*(5), 1–22.

Kasper, G., & Rose, K. R. (2001). *Pragmatics in language teaching.* Cambridge: Cambridge University Press.

Liao, S. (2009). Variation in the use of discourse markers by Chinese teaching assistants in the US. *Journal of Pragmatics*, *41*(7), 1313–1328. https://doi.org/10.1016/j.pragma.2008.09.026

Linn, R. L. (2000). Assessments and Accountability. *Educational Researcher*, *29*, 4–16. https://doi.org/10.3102/0013189x029002004

Liu, J. (2006). *Measuring interlanguage pragmatic knowledge of EFL learners.* Frankfurt: Peter Lang.

Marchante, B. M. (2015). *La evaluación de la competencia pragmática en lengua extranjera a través de una prueba adaptativa* (Doctoral dissertation, Universidad Politécnica de Valencia). Retrieved from https://doi.org/10.4995/Thesis/10251/59437

Martínez, M. A., Sánchez, A., Toledo, E., & Faulín, J. (2006). *Bioestadística amigable.* Madrid: Díaz de Santos.

Martínez-Flor, A., & Usó-Juan, E. (2010). *Speech act performance: Theoretical, empirical and methodological issues.* London: John Benjamins Publishing. https://doi.org/10.1075/lllt.26

Martínez-Rizo, F. (2008). *Proyecto de creación del INEE.* México: Instituto Nacional para la Evaluación del Educación. Retrieved from https://www.inee.edu.mx/wp-content/uploads/2019/01/P1C226.pdf

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (pp. 13–103). New York, NY: Macmillan.

Morrow, K. (1981). ELT Documents. 111 Issues in Language Testing. In J. C. Alderson & A. Hughes (Eds.), *Communicative language Testing: revolution or evolution* (pp. 9–25). London: British Council.

Norris, J. M. (2001). Use of address terms on the German Speaking Test. In K. R. Rose (Ed.), *Pragmatics in Language Teaching* (pp. 248–282). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139524797.018

Oller, J. W. (1979). *Language Tests at School.* London: Longman.

Omar, A. S. (1990). *How Learners Greet in Kiswahili: A Cross Sectional Survey.* Retrieved from http://files.eric.ed.gov/fulltext/ED326061.pdf

Purpura, J. E. (2004). *Assessing Grammar.* Cambridge: Cambridge University Press.

Purpura, J. E. (2009). Retrieved from https://www.oxfordenglishtesting.com/uploadedfiles/6_New_Look_and_Feel/Content/oopt_measure.pdf

Reyes, G. (1990). *La pragmática lingüística.* Barcelona: Editorial Montesinos.

Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing*, *23*(2), 229–256. https://doi.org/10.1191/0265532206lt329oa

Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, *28*(4), 463–481. https://doi.org/10.1177/0265532210394633

Safont-Jordá, M. P. (2005). *Third language learners: Pragmatic production and awareness.* London: Multilingual Matters.

Shepard, L. A. (2006). Classroom assessment. *Educational measurement*, *4*, 623–646.

Shepard, L. A., Hammerness, K., Darling-Hammond, L., Rust, F. W., Snowden, J. B., Gordon, E., ... Pacheco, A. (2005). Assessment . In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world. What teachers should learn and be able to do* (pp. 169–200). San Francisco, CA: Wiley.

Soler, A., E., Flor, M., & A. (Eds.). (2008). *Investigating pragmatics in foreign language learning, teaching, and testing. London: Multilingual Matters.*

Soler, E. A., & Martínez-Flor, A. (2005). Editors' introduction to pragmatics in instructed language learning. *System*, *33*(3), 381–384. Retrieved from https://dx.doi.org/10.1016/j.system.2005.06.002   10.1016/j.system.2005.06.002

Takahashi, S. (2010). Assessing learnability in second language pragmatics. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (pp. 391–422). De Gruyter Mouton. https://doi.org/10.1515/9783110214444.3.391

Takahashi, T., & Beebe, L. M. (1987). The development of pragmatic competence by

Japanese learners of English. *JALT journal*, *8*(2), 131–155. Retrieved from https://jalt-publications.org/sites/default/files/pdf-article/jj-8.2-art3.pdf

Tsutagawa, F. S. (2013). Pragmatic knowledge and ability in the applied linguistics and second language assessment literature. *Working Papers in TESOL & Applied Linguistics*, *13*(2), 1–20.

Xiao, F. (2015). Proficiency effect on L2 pragmatic competence. *Studies in Second Language Learning and Teaching*, *5*(4), 557–581. https://doi.org/10.14746/sllt.2015.5.4.3

Yamashita, S. O. (2008). Investigating interlanguage pragmatic ability: what are we testing? . In E. A. Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching, and testing* (pp. 201–223). London: Multilingual Matters.

Yule, G. (1996). *Pragmatics*. Oxford: Oxford University Press.