Analysis of pragmatic items in an ESL online adaptive placement test

MARÍA LUISA CARRIÓ-PASTOR∗ & BEATRIZ MARTÍN MARCHANTE
Universitat Politècnica de València (Spain)

ABSTRACT

The work at hand is part of a wider study the aim of which was to determine what kind of factors influence failures in pragmatic items of an online adaptive placement test taken by a group of 34 Spanish students in their first year at university. A preceding analysis (Carrió & Martín, 2016) showed the type of personal factors, such as lack of vocabulary, that caused the exam takers failures in the pragmatic part of the test according to their own perception. In this paper, we go deeper and analyze those specific items holding higher percentage of failures so that we can reach some conclusions about its content validity and also about the degree of relationship between these and the pragmatic categories dealt with in the most relevant research carried out in the field of pragmalinguistic testing and assessment.

KEYWORDS: pragmatics, ESL, language testing and assessment.

1. INTRODUCTION

In the Faculty of Teacher Training of the University of Valencia, Spain, students are offered the possibility of certifying their level of English through an external test, the Oxford Online Placement Test (OOPT). In a previous study (Carrió & Martín, 2016) it was observed, through a questionnaire, that the examinees found it more difficult to answer some questions than others, and that their greatest trouble was in understanding some phrases, dialogues and

_____

*∗**Address for correspondence**: María Luisa Carrió-Pastor. Departamento de Lingüística Aplicada. Universitat Politècnica de València, Spain. e-mail: lcarrio@upv.es

written and oral fragments of different length contained in certain items. The most frequently argued reason by the examinees was "misunderstanding of the vocabulary".

The aim of this work was to analyze the comprehension and interpretation of pragmatic aspects in a given context (language testing and assessment) by its users (students and examinees). More specifically, the objective was to check the content validity of the pragmatic part of the OOPT through the categorization of the items holding the highest percentage of failures and also by comparing them to those investigated in this field. In this way, we hope to contribute to the demarcation of pragmatic aspects that need to be improved in order to achieve the ultimate goal that foreign language teaching professionals aim at, i.e. to enable students´ progress in achieving greater fluency and mastery of a foreign language. We have found a research gap in the content validation of pragmatic items in online adaptive tests and we try to answer some research questions with our study.

The research questions that we aim to answer in this work are:

> What pragmatic categories do some problematic items in the OOPT correspond to?
> Are these categories used in other tests of pragmatic competence?

In order to answer these questions, this paper is organized as follows. Firstly, some key aspects of pragmatics that are used in the analysis are described: speech acts, implicatures and lexicalized tropic inferences (formulas, routines, etc.)[1] and the studies that focus on this referred to. Secondly, the evaluation of these pragmatic aspects is explored. Thirdly, the students involved in the project, the materials and the method used are explained and after that, the results extracted from our analysis are shown and discussed. Finally, some conclusions are drawn.

2. PRAGMATICS TEACHING AND ASSESSMENT

Usually, within the field of pragmatics testing and assessment, isolated aspects of pragmatic competence have been investigated. Most of the research has so far been done based within the traditional speech act theory (Austin, 1962; Brown & Levinson, 1987) with the exception of Walters (2004, 2007) and Roever (2005, 2006, 2011) tests.

However, pragmatics also encompasses different theoretical models such as relevance theory (Sperber, 1986, Wilson & Carston, 2007), or politeness theory (Brown & Levinson, 1987), and other areas of study, for example, enunciation and language functions (Halliday, 1979), inferences (Grice, 1975), modality and modalization (Payrató, 2010). Furthermore, in the light of recent research, there is now considerable concern about social theories of pragmatic learning, cognitive theories and complex dynamic theories (Taguchi, 2012).

From a general point of view, the phenomena studied by pragmatics ranges from very language context –dependent ones, such as deixis (person, space, time and social deixis, vocative), to less dependent on the structure of the language phenomena such as implicature, or presupposition. But in tests of L2 pragmatics, only the following components of the pragmatic competence are usually chosen as the framework for the validation of the construct (Messick, 1989; Rover, 2005):

1. Knowledge of speech acts and their strategies.
2. Interpretation of implicatures.
3. Recognition of formulas and routines.

These three points are explained in the following sub-sections.

2.1. Speech acts

At the core of speech act theory there is the idea that context impacts meaning and consequent action. What words can actually do and achieve in the real world depends on the context of an utterance. The following utterance, "He is a dangerous man", is a statement in the surface but it can also be interpreted as a warning. So, rather than a statement, this utterance is speech act (SA) since it uses language to achieve a real-world effect (Taguchi & Roever, 2017).

The theory of speech acts presupposes, according to Reyes (1990), that every linguistic act reveals an intention and is an exercise of sincerity. But she postulates, at the same time, that "[...] every speech has to be conventionalized, and use recognizable and perceptible formulas" (Reyes, 1990: 121). In other words: "[...] the minimal units of human communication are not linguistic expressions, but rather the performance of certain kinds of acts, such as making statements, asking questions, giving directions, apologizing, thanking, and so on" (Blum-Kulka, House & Kasper, 1989: 2).

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, 1962, in Cohen, 1996: 384). Indeed, a speech act is an action of language that produces effects in what surrounds us and in which three dimensions can be distinguished simultaneously: locutive, illocutionary or ilocutive[2] and perlocutionary.

Illocutionary speech acts can be direct or indirect. 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 before indirect speech acts.

The table below is a diagram showing the division of illocutionary speech acts and the direction of adaptation.

Speech act type	Direction of fit	S=speaker X= situation
Declarations	Words change the world	S causes X
Representatives	Make words fit the world	S believes X
Expressives	Make words fit the world	S feels X
Directives	Make the world fit the words	S wants X
Commissives	Make the world fit the words	S intends X

Table 1. Five general functions of speech acts (Yule, 1996: 55).

Table 1 shows the five general functions that speech acts can perform: declarative, representative, expressive, imperative and commissary. According to Yule (1996), when the speaker makes use of a statement, he/she tries to change the world through words. When the function used by the speaker is representative, he/she tries to fit the words in the world (world of beliefs). When the function is the expressive one, its objective is to adapt the words to the world (world of the feelings). If the speaker uses the imperative function, then what he/she intends to do is to adapt the world to the words through the listener, and when the function of the language is the commissary one, what he/she wishes is to adapt the world to the words but in this case through the speaker.

As mentioned before, the most common area of research in the field of language teaching and assessment is by far that of speech acts. Speech acts of request and apology have been the most widely studied as well as the rater variation in the assessment of speech acts (Blum-Kulka, 1982; Biesenbach-Lucas, 2007; Cohen & Shively, 2007; Felix-Brasdefer, 2007; Taguchi, 2006; Taguchi, 2011). Also, the appropriate levels of recognition and politeness (Liu, 2006) have commonly been of interest.

2.2. Implicatures

With regard to implicatures, Grice (1975) distinguishes between conventional and conversational. The latter are subdivided in turn into particularized and generalized ones[3]. According to Grice (1975), these are cases in which inferences can be done without the need for a special reference to the context, as in example (1):

(1) The other day Pilar entered a house

In this example, it is inferred that the place which Pilar entered was not her house.

The particularized conversational implicatures are characterized as cases in which the use of a certain form of words would normally carry (in the absence of special circumstances) one or the other implicature or type of implicature. These implicatures are inferred only because of a special context (Grice, 1989: 37), as can be seen in example (2):

> *(2) A: Can you tell me the time?*
> *B: Well, the milkman is here.*

In this second example, it is inferred that the hour must be the one the milkman usually arrives. Grice does not actually give a clear definition of implicature although it is understood that it is something that is not part of what it is said. He introduces the term implicature and uses it as a general term to avoid having to choose between words like "implying, suggesting or meaning" (1989). He points out that to involve is to perform a non-central speech act, whereas to say is to perform a central speech act. In other words, to say and to imply would be equivalent to an act of direct speech, and to an act of indirect speech respectively.

Gallardo (1995) defines implicatures as a kind of nonconventional inferential meaning not based on the use of certain words, but on the application of certain communicative norms. They have a social character and are outside the structures of language (inferences). According to this author (1995, 2005), the meaning that is not said but inferred from a statement is an implicature and she develops the following classification of implicit meaning from Grice (1975):

1. What is conventionally implied. In conventional implicatures the meaning depends on the concrete use of certain words or structures of the language and two types are distinguished between them:
   A. Presuppositions: They are meanings that are assumed to be true when using certain statements, and that are triggered by the use of certain triggers or presuppositional triggers.
   B. Lexicalized tropes or idioms: Locutions, phraseological units whose global meaning does not coincide with that of the sum of its parts.
2. What is not conventionally implied. They are meanings that are activated by virtue of certain socio-discursive norms, such as the Cooperative Principle (CP):
   A. Conversational, called implicatures by Grice (1975) related to the application of the conversational maxims that are derived from the CP (quality, quantity, manner and relevance) that in turn can be:
      - Generalized, that is, they are activated by respect to the conversational maxims.
      - Anomalous, that are activated when violating the conversational maxims.

B. Conventional: They are those that are based neither on the spoken words nor on discursive norms:

Implicit meanings can only be explained from the relationship between speaker and listener. Together with these inferences, which are activated with the syntactic-semantic discursive use, linguistics differentiates other types of textual nature which are more general and are called textual inferences.

3. Cultural implicit or assumed meaning: inferences that are activated by encyclopedic knowledge.

Grice's theory of implicatures (1975) is one of the most important aspects in pragmatics. However, in spite of the importance of this theory, there have been few studies on the teachability of Gricean pragmatics (Bromberek-Dyzman & Ewert, 2010; Keenan, 1976; Murray, 2010, 2011) and the discrepant results obtained in these studies evidence the need for further research.

On the other hand, Yamashita (2008) argues that it is necessary to expand the construct, and Brown (2008) suggests using the G theory and Decision studies (Tsutagawa, 2013: 11-13). Bouton (1988, 1994, 1999) evaluates knowledge about implicatures and the indirect use of English as a foreign language (EFL) by learners.

2.3. Lexicalized trope inferences

Included in the conventional implicit category are the lexicalized trope -inferences (LTI) or idioms, which are idiomatic expressions that are fixed in the language, that is, they have undergone a process that linguists call lexicalization or grammaticalization, since they are phrases that function as an inseparable lexical entity. They are groups of fixed words whose meaning is usually known by any competent speaker and usually does not come exclusively from the isolated meanings of the words that constitute them.

The semantic treatment of phraseological units is similar to that of other lexical units (words). It is important to distinguish between lexicalized trope-inference, converted into part of the lexicon of a language, and the figurative use as such, that a speaker can improvise, at a certain point in the discourse by making use of his/her creativity (Gallardo, 2005). Kerbrat-Orecchioni (1996) refers to lexicalized tropes as a type of conventional meaning that has crystallized in the language in the form of phrases, sayings, idioms, rhetorical questions, etc.

Also, routines or formula are, according to Dulay, Burt & Krashen (1982: 232-233) "whole utterances that are unusually error-free and show no transitional stages of development or systematic order of acquisition. They are learned as unanalyzed wholes,

much as one learns a single word". They are usually linked to more or less standardized situations and are highly conventionalized expressions.

Coulmas (1981: 2-4) defines them as "tacit agreements, which the members of a community presume to be shared by every reasonable co-member". And House (1996, quoted in Bardovi-Harlig & Félix-Brasdefer, 2016) argues that from a sociolinguistic point of view learning routines are essential at any stage because they represent the social knowledge shared by members of a community. According to Kasper & Schmidt (1996), making use of routine formulas is along with expressing pragmatic intention indirectly, one of the universally available pragmatic strategies. Coulmas (1981) explains that routine is a property of an expression or strategy that is appropriate for a certain type of situation or is relative to certain communicative ends. They are a means to guide the normal participation of a person in social interaction. Routines such as greetings are universal phenomena.

In our study, we preferred to replace this term "routine formula" with "lexicalized trope- inference" (LTI) (Gallardo, 2005) because it is broader and includes terms such as phrasal verbs, idioms, sayings, formulas or routines.

An area of interest is centred around the study of routines (Bardovi-Harlig, 2008, Bouton, 1988, 1994; House, 1996; Nattinger & DeCarrico, 1992; Roever, 2005, 2006; Taguchi, 2008) conceptualized as formulaic expressions subject to specific social environments that fulfill a specific purpose.

All in all, in previous lines we have pointed out the difference between lexicalized trope-inference and the figurative use that is based on the transgression of conversational maxims. Grice (1975) explains that participants in front of the conversational maxims can be divided into three groups: first, the speakers who follow the maxims without violating them; second, those who violate them, but their transgression is explained by the assumption of a collision with another maxim and, finally, the speakers who violate a maxim but they do not stop to cooperate, since in these cases the speaker does it hoping that the listener discovers such violation and initiates an inferential process through which reveals the implicature or, in other words, the implicit information. That is what Grice (1975: 45) calls *exploiting* a maxim, for example, "if the speaker is saying something that is believed to be false; if the speaker does irony or makes ironic and sarcastic statements; if the speaker denies something; if the speaker distorts information".

Some examples in which the maxim of quality is exploited are irony, metaphor, meiosis and hyperbole (Grice, 1975). Tautologies, on the other hand, are examples of transgression of the maxim of quantity, while ambiguity, obscurity and verbosity are examples that occur when the maxim of the relation and /or manner are violated. In general, indirect speech acts are transgressions of the maxim of relevance.

2.4. Pragmatics assessment

Domain tests are designed to assess the competence of students in different areas of knowledge of language and they are not based on a particular program. Some domain tests are intended to clarify whether students have acquired a certain level of language proficiency. Standardized tests give information on the degree of mastery or knowledge in a given task or skill, that is, compare the score of each student with that of the others and the results are usually represented in percentiles or in the Gaussian curve. They are also called tests referred to the norm (Madrid, 1997). The target domain of most pragmatic tests tends to be the social language in use in everyday life settings (small talk, shopping, table talk), although not always explicitly formulated, and, to a lesser extent, academic settings.

Research has focused mainly on the effects of exposure and socialization, on the use of routines as well as their teachability. There are fewer studies on the use of language in extensive social interaction or on the effect that such use has on the interlocutors. The data are collected in these studies through role playing (Félix-Brasdefer, 2007) and occasionally in natural environments.

Politeness theory has also been criticized for not being applicable to non-Western cultures. Furthermore, Kasper (2006) and Schegloff (2007) have questioned the deterministic understanding of the context as something entirely constituted by the situational variables of power, distance, and imposition of Brown and Levinson (1987) for not taking into account the dynamics of discourse and the internal context of the conversation. Roever (2005, 2006) defends measures in real time of the apprentice integration skills so that the extrapolation to the target domain of the social use of the language can be defended. Gumperz (1982) names prosody, facial expression, structure of the argument, choice of content and adherence to or ostentation of social norms as indexes of the speaker's analysis of the social situation and states that they establish a style of speech. But other aspects that appear in Table 2 should also be taken into account:

Monologic: extended monolog	Dialogic: participation in interaction	Routine formulae	Implicature
Production & recognition of: • Speech styles • Contextualization cues • Discourse structure	Production & recognition of: • Speech styles • Contextualization cues • Sequence organization: pre-sequences, core sequences, post-sequences • Openings & closings • Repair • Response to first-pair parts • Effect on interlocutor	Production and recognition of routine formulae	Comprehension of implicature

Table 2. Components of L2 pragmatic ability with sub-constructs (Roever, 2011: 11).

Conversely, the model described by Purpura (2004) explains how the grammatical resources are used to express a variety of literal as well as implicit meanings in the language in use. Purpura's definition of linguistic ability includes grammatical and pragmatic knowledge and distinguishes the grammatical and semantic meaning (e.g. literal meaning and implicit meaning) of an expression from the layers of pragmatic or implicit meaning that are contextually driven, and which often cannot be derived solely from the arrangement of the words in syntax. His model takes into account how pragmatic (implicit) meanings are superimposed on grammatical structures and semantic meanings in the language in use. Pragmatic knowledge, related to the knowledge of implied meanings, includes the ability to understand and express sociolinguistic, sociocultural and psychological meanings. These pragmatic meanings involving propriety, conventionality, naturalness and acceptability, can only be determined in situations with a high degree of context. The more indirect the expression is, the richer the textual characteristics have to be, so that the meanings can be decoded.

> […] pragmatic meaning embodies a host of implied meanings that derive from context relating to the interpersonal relationship of the interlocutors, their emotional or attitudinal stance, their presuppositions about what is known, the sociocultural setting of the interaction and participation of an interlocutor during talk-interaction. (Purpura, 2004: 262)

To evaluate the pragmatic ability, the following test types have primarily been used:

- *Written Discourse Completion Task* (WDCT)
- *Multiple Choice Completion Task* (MDCT)
- *Oral Discourse Completion Task* (ODCT)
- *Discourse Role Play Task* (DRPT)
- *Discourse Self Assessment Task* (DSAT)
- *Role Play Self Assessment* (RPSA)

The multiple choice tests have been proven to be useful as a measure of pragmatic understanding and also to evaluate the domain implicatures (Bouton, 1988, 1994; Roever, 2001) formulas and routines (Read & Nation, 2004; Roever, 2001) and to measure politeness (Tanaka & Kawade, 1982). They are used very often in traditional grammar tests, since they easily examine in a short time, facilitating the analysis. The need for the multiple response format seems to be increasing with the growing number of computer-based and online tests.

To elicit pragmatic perception data (oral and written) rating scales and verbal reports are often used. On the other hand, as the field of pragmatics has evolved, so have the elicitation methods of experimental data (Barron, Gu & Steen, 2017). Currently methods

from psycholinguistics are being applied (e.g. forced choice tasks such as truth value judgments, felicity judgments and picture selection tasks) and also psychological measures (event related potential (ERP), self-paced reading, eye tracking or functional magnetic resonance imaging (FMRI)). Due to the fact that different methods of measurement are used as well as different methods of analysis, we can better understand pragmatic production, its perception and its understanding.

To summarize, in this section we have presented the most relevant research on pragmatics teaching and assessment and also the investigation gaps in the field. Some pragmatic categories (speech act, implicatures) have been more profusely studied than others and usually in an isolated way. We have also reviewed how pragmatic competence has usually been assessed in foreign language settings through tests such as DCTs, role play, etc. Nevertheless, in the field of foreign language testing, the measurement of pragmatic competence is an area still insufficiently explored. Understanding and expressing pragmatic meanings in communication is so important that failing to evaluate them has serious implications. Bearing this in mind, the purpose in the present study is to find out how pragmatics is assessed in the OOPT since we have not found any research on the content validity of this adaptive high-stake test. The analysis carried out examines what type of items are used and what type of categories are more problematic for the participants in this study.

The online adaptive tests such as the OOPT undergo a thorough process of validation which cannot stop once the test has been administered. This process of validation does not finish until the results of the exam takers are obtained and the judgment of the experts are given. We noticed in this study that there is a need to continue the process of validation of these types of tests since they have great effect in the future and professional life of language students.

3. METHOD

The classification of the items of the pragmatic block in the Oxford Online Placement Test (OOPT) has been approached to assess its construct validity through a qualitative, in-depth study. The participants in this study were first-year undergraduate students of EFL who enrolled in the compulsory subject "Llengua Estrangera per a Mestres: Anglés" at the Teacher Training Faculty of the Universitat de València and were being trained to become primary school teachers in English. Each year, the Faculty of Teaching facilitates the administration of the OOPT test to students who do not have official certification that demonstrates the possession of levels B1 and B2 of language proficiency. For the analysis of the results of this test only the samples of students whose resulting levels ranged between A2

and B1 were collected, which turned out to be 34 in total. Samples were selected among these levels because they contained the highest number of failures.

In 2000, Oxford University Press distributed the *Oxford Online Placement Test* (OOPT). The main objective of the OOPT was not only to measure grammatical or lexical competence but to measure the communicative competence of the examinees. As Purpura (2009: 1) explains, "it measures a test taker's ability to understand a range of grammatical forms and the meanings they convey in a wide range of contexts. It also measures the extent to which learners can use these language resources to communicate in English language situations".

In the development of this test, the descriptors of the CEFRL as well as the current research trends regarding the nature of communicative competence and the learning of foreign languages were taken into account. The most common interlanguage errors were also considered, such as the interference of the mother tongue. The OOPT is composed of two main parts, one devoted to reading and writing skills and one to listening skills. The first contains around 30 questions and evaluates vocabulary, grammar and pragmatics. In words of Purpura (2009: 2): "The *Use of English* section is designed to measure how much learners know about grammatical forms and the meanings (e.g. word meanings, phrasal meanings, sentence meanings) that these forms encode [...] designed to measure the students' knowledge of the pragmatic (i.e., implied) meanings encoded in situated interactions".

All the items in the OOPT were *Written Discourse Completion Task* (WDCT) *and Multiple Choice Completion Task* (MDCT). Figure 1 is an example of these MDCT items.
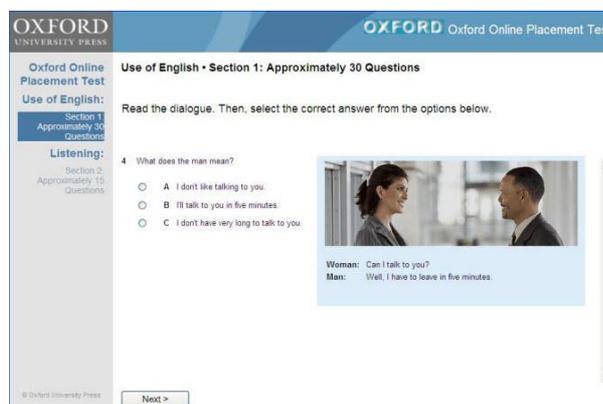
Figure 1. Sample of pragmatic item.

According to Contreras (1990: 170), "in order to carry out a posteriori analysis of a test, there is no possible statistical verification, but rather the logical judgment about what the questions and the objectives are intended to be". Consequently, in our study, we have checked the content validity of some of the pragmatics items of the OOPT, that is to say, if a sufficient and representative sample of the traits which are aimed to be measured is included.

Our purpose was to determine what pragmatic categories were contained in the items analyzed in the previous study by Carrió and Martín (2016).

The authors analyzed those pragmatic items which were failed by a higher number of participants and found out, through an analysis of correspondence, that lack of vocabulary caused the 34 students' failures in the pragmatic part of the test according to their own perception. In this paper, we go deeper and analyze those specific items which were number 12, 14, 15 and 17. There were collected 23, 25, 24 and 27 samples corresponding to number 12,14, 15 and 17, respectively. The total number of samples of erroneous answers was 99, given that one student usually failed several items.

The compilation of data was carried out in six steps. Firstly, an individual recording of the incorrect answers was performed through the summary of each student's error provided online by Oxford University Press. This procedure was carried out due to the fact that the OOPT is adaptive and there may be different items for each student and they can belong to different pragmatic categories. A setback was that we did not have access to the bank of items and being an adaptive test also implied that no traditional item analysis was viable. That is the reason why we could not find the difficulty and discrimination index of each item. The individual recording was followed with items 12, 14, 15 and 17. Thirdly**,** each sample was archived and tagged. After each sample was tagged, it was observed that all of them fell within either of the following three pragmatic categories: direct speech acts (DSA), indirect speech acts (ISA) implicatures (I) and lexicalized trope-inferences (LTI). These coincide with the three components of pragmatic competence chosen by Roever (2005) who worked on the framework for validation proposed by Messick (1989).

Fourthly, a matrix was created for each question in order to classify each sample as direct speech act DSA, ISA, I or as LTI. Fifthly, the distribution of the answers to the four items was found and the most frequent categories determined in each of them. Finally, each item was described, including the percentage of examinees who had failed it and the frequency of pragmatic categories presented. The homogeneity of the frequencies was contrasted through a Chi^2 test.

4. RESULTS

After analyzing 99 samples of incorrect answers given by 34 students to items 12, 14, 15 and 17 in the pragmatic part of the Oxford Online Placement Test (OOPT), we obtained the following results:

The pragmatic categories under which all the wrong answers were classified were mostly lexicalized trope-inferences (LTI), implicatures (I) and indirect speech acts (ISA). Table 3 shows the distribution of these variables. A higher frequency in LTI in items 14 and 15 and implicatures in items 17 and 14 is noted. In item 12, ISA, I and LTI are distributed

homogeneously (30.4% each). In item 15, LTI stands out with 45.8%, followed by ISA (29.2%), DSA and I, both with 12.5%.

Pragmatics categories		Raw no.	%
ITEM 12	**Total**	23	100.0%
	DSA	2	8.7%
	ISA	7	30.4%
	Implicatures	7	30.4%
	LTI	3	30.4%
ITEM 14	**Total**	25	100.0%
	DSA	1	4.0%
	ISA	3	12.0%
	Implicatures	4	40.0%
	LTI	8	44.0%
ITEM 15	**Total**	24	100.0%
	DSA	1	12.5%
	ISA	1	29.2%
	Implicatures	1	12.5%
	LTI	7	45.8%
ITEM 17	**Total**	27	100.0%
	DSA	1	3.7%
	ISA	3	11.1%
	Implicatures	7	51.9%
	LTI	6	33.3%

Table 3. Distribution of pragmatic categories for items 12, 14, 15 and 17.

Different frequencies can be distinguished in the classifications of the items. For example, in item 17, 51.9% of the 27 samples of flawed questions correspond to the category of implicatures, while only 3.7.0% correspond to the category DSA. In item 15, 45.8% correspond to LTI and 29.2% to ISA. On the other hand, in item 14, 84.0% is distributed between implicatures and LTI (40.0% and 44.0%, respectively) versus 4.0% of DSA. The frequency found in ISA and implicatures in item 12 is the same for both, 30.4%, and in DSA, it is 8.7%.

The homogeneity of the frequencies of the answers to items 12, 14, 15 and 17 has been contrasted, concluding that it cannot be admitted that the 4 categories occur in the same

proportion. Significant differences were found in items 14 and 17 through a Chi^2 test of homogeneity, with p-value= 0.008 and p-value= 0.001, respectively, that is, in P14 there are more categories of one type than another. The same can be said for P17.

In summary, according to the results in Carrió and Martín (2016), 48% of the reasons for failure in item 12 raised by students was the incorrect interpretation of the dialogue even knowing the vocabulary. Now, in this study, we have found that in this item the category *ISA* stands out with 30.4% and "anomalous implicatures" with 26.1% of the distribution. The following examples (3) and (4) are samples of the abovementioned pragmatic categories.

> *(3) Indirect speech act (ISA)*
> Man: Let's walk to the beach this afternoon.
> Woman: Isn't it too far in this weather?
> What does the woman mean?
> A I don't know how far it is.
> B I never walk such a long way.
> C I don't want to walk there today

> *(4) Indirect speech act (ISA)*
> Woman: Finished fixing that plug yet?
> Man: It's proving trickier than I thought.
> Woman: It's not exactly rocket science, you know.
> What does the man mean?
> A It was difficult but I managed to fix the plug.
> B I haven't been able to fix the plug yet.
> C Fixing the plug was no problem.

On the other hand, item 14, is the one with the greatest number of lexicalized tropic inferences. Eleven samples have been classified as such, which represent 44.0% and where the main reason for failure is attributed to the lack of understanding of one or more words. Here are two examples in (5) and (6):

> *(5) Lexicalized tropic inference (LTI).*
> Man: I'll tell you what; I can do this job with my eyes closed.
> Woman: So, time to move on then?
> Man: If only!
> What does the man mean by his FIRST comment?
> A I can't understand this job.
> B I have to concentrate to do this job.
> C I no longer feel particularly challenged by this job.

> *(6) Lexicalized tropic inference(LTI) (Phrasal verb).*

Man: Josh gets on my nerves.
Woman: Oh, I really like him!
What does the woman mean?
A Josh scares me.
B Josh annoys me
C Josh worries me.

Also, item 15 coincides with item 14 in that there are a total number of 11 LTI, in this case, 45.8%. This item, together with item 12, is the one with the highest number of ISA, that is, 29.2%. The incomprehension of one or two words is, as in P.14, the most associated reason for failure.

Item 17 is, by far, the one with the greatest number of implicatures, 14 in total, that is, 51.9%. Most students said the reason for failure was "incorrect interpretation even knowing the vocabulary". The second most frequent category in this item is the ITL with a percentage of 33.3. The following samples (7) and (8) are classified as implicatures.

(7) Implicature
Woman: How was your holiday?
Man: I didn't want to come home.
What does the man mean?
A He got lost.
B He went out a lot.
C He had a great time

(8) Implicature
Mother: Have you learnt all your lines for the play tonight?
Son: Yes, but I know I'll forget them as soon as I get on the stage.
Mother: Don't worry, you'll be fine.
What does the son mean?
A I'll be nervous on stage tonight.
B I can't remember my lines now.
C I need to get on the stage as soon as possible.

So far, several examples of the most frequent pragmatic categories have been presented and we can conclude that the pragmatic category that most frequently appears is the lexicalized tropic inference with an average of 38.3%. It is followed by the category "implicature", with 33.7%. On the other hand, while direct speech acts have an average frequency of 7.2%, in indirect speech acts the average is 20.6%.

5. CONCLUSIONS

In this paper we have examined the content validity of the pragmatic items included in the OOPT. As for the first research question "What pragmatic category do some problematic items in the OOPT correspond to?" the results show all samples analyzed corresponding to items 12, 14, 15 and 17 fall into the following categories: lexicalized trope-inferences, implicatures, indirect speech acts and direct speech acts. The frequency of the lexicalized trope-inference category is remarkable among the others. The second category is implicature, followed by indirect speech acts. It should be noted that direct speech acts have hardly any presence among them. Before continuing, we would like to say that although we have focused on the items with higher percentage of failures due to the limited time available and the difficulty of collecting data, it is worth noting that we have observed that the other items share the same categorization. Looking into the non-problematic items could be the goal of future research.

As for the second question: Are the pragmatic categories the same used in other tests of pragmatic competence? We are in a position to affirm that these pragmatic categories match those dealt with in the most relevant research and the OOPT is in the same vein as the tests used up to now. Nevertheless, this does not seem to be the most advisable, since the test offers an insufficient representation of the construct. This is a fact already called into attention by Roever (2011) who reviews existing evidence and advocates a discursive reorientation of tests devoted to foreign language pragmatics.

Furthermore, the fact that these pragmatic categories are observed in the design of the OOPT is not enough, since it would be necessary to obtain a good index of difficulty and discrimination in these items, a task that is very complicated because one of the great problems in the evaluation of the pragmatic ability is the fact of not knowing with certainty what the natural route of acquisition of this ability is. It is not so in the acquisition of morphosyntactic competence, where it is known that, at least in Romance languages, progressive times and indicative mood are acquired before past, participles and subjunctive, etc. Probably, the most critical area of debate in relation to the evaluation of pragmatic competence in a foreign language is to learn what actually is known about the development of pragmatic ability in EFL and identify relatively stable acquisition sequences.

The results of the analysis of the items obtained in this study together with the results of a previous study that analyzed the reasons for pragmatic failure (Martín, 2015; Carrió & Martín, 2016) point to the fact that there is an association between the failure rate, the failure motive and the pragmatic category of each item at issue. For example, items 12, 14, 15 and 17 are items with high failure rates (47%, 50%, 47% and 52%, respectively), as stated by Martín (2015). Between 30.4% and 45% of the samples classified in each of the four items are lexicalized trope- inferences and the main cause of failure pointed out by the students themselves is the incomprehension of one or two words in the dialogues of the pragmatic

items. This indicates that we are primarily facing a problem of receptive knowledge of vocabulary. The pragmatic category with the highest frequency is the lexicalized trope inference with an average of 38.4%. It follows the implicature category, with 33.7%. Speech acts have an average frequency of 7.2% and in indirect speech acts the mean is 20.7%.

These results also clearly show that teachers can barely assess the pragmatic knowledge of students through an adaptive test and that the biggest problem in answering the pragmatic questions of the OOPT for the students participating in this study has been the lack of lexicon, which has hindered the process of pragmatic interpretation by not having acquired the necessary reading comprehension competence. Therefore, the findings clearly suggest that if these students have not passed the threshold level of lexical knowledge, a necessary condition to be able to diagnose the grammatical ability (Alderson, 2007), it is not feasible either to diagnose their pragmatic ability when these circumstances occur.

The underlying philosophy of this type of computerized commercial test clashes with the requirements of any language assessment that one wishes to carry out within a communicative frame of reference. The formal procedure of this test cannot reliably reproduce context-dependent human behaviors. Therefore, we believe that this type of tests (adaptive or not) should be complemented with oral tests that allow interlocutors´ interaction. As stated by Turnbull (2001), the language used in oral and written DCTs is not representative of natural language while in open role-playing games and with experimental techniques, the language is closer to it. Likewise, it is necessary to carefully select items that contain vocabulary known by the learner so that the possible pragmatic deficiency can be isolated more clearly without being confused with lack of vocabulary or scarce reading comprehension given that, as we have seen in the results presented here, what can be actually evaluated is the pragmatic lexical competence and/or reading comprehension. Therefore, we think vocabulary should be taught not so much as words but as a set of words, lexical units and collocations. A lexical pragmatics perspective should be adopted, as indicated by Nattinger and DeCarrico (1992), Lewis (1993, 2000) and the defenders that the language is mainly constituted by grammaticalized lexicon. For this purpose, material and programs that help improve students' skills can be designed and integrated. This would involve the preparation of fragments contextualized within the framework of English for specific purposes and within it, English for academic purposes.

NOTES

1   In our study, we use the term "lexicalized trope- inference" (LTI) (Gallardo, 2005) because it is broader and includes terms such as phrasal verbs, idioms, sayings, formulas or routines.
2   Later, Searle (1969) replaced these concepts with the "propositional act".
3   The notion of generalized implicature has been extensively studied by neo-Griceans (Atlas & Levinson, 1981) as well as the scalar and clausal implicatures that are subclasses of which

Levinson (2000) called Q-, I- and M-Implicatures. Other scholars have also studied the R- and I-Implicatures (Bezuidenhout & Morris, 2004).

REFERENCES

Alderson, J. C. (2007). The challenge of (diagnostic) testing: Do we know what we are measuring? In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner & C. Doe (Eds.), *Language testing reconsidered* (pp. 21-39). Ottawa: University of Ottawa Press.

Atlas, J. D. & Levinson, S. C. (1981). It-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In P. Cole (Ed.), *Radical pragmatics* (pp. 1-62). New York, NY: Academic Press.

Austin, J. L. (1962). *How to Do Things with Words*. Oxford: Calderon Press. Retrieved February 12, 2017 from http://www.dwrl.utexas.edu/~davis/crs/rhe321/Austin-How-To-Do-Things.pdf

Bardovi-Harlig K. (2008). Recognition and production of formulas in L2 pragmatics. In Z.-H. Han (Ed.), *Understanding second language process* (pp. 205–222). Clevedon: Multilingual Matters.

Bardovi-Harlig, K. & Félix-Brasdefer, J. C. (Eds.). (2016). *Pragmatics & Language Learning*, Vol 14. Honololu: Natl Foreign Lg Resource Ctr. University of Hawi´i Press.

Barron A., Gu, Y. & Steen, G (Eds). (2017). *The Routledge Handbook of Pragmatics*. London: Routledge.

Bezuidenhout, A. L. & Morris, E. K. (2004). Implicature, Relevance and Default Pragmatic Inference. In D. Sperber & I. Noveck (Eds.), *Experimental Pragmatics* (pp. 257-282). London: Palgrave Macmillan.

Biesenbach-Lucas, S. (2007). Students writing emails to faculty: An examination of e-politeness among native and non-native speakers of English. *Language Learning & Technology*, *11*(2), 59-81.

Blum-Kulka, S. (1982). Learning to Say What You Mean in a Second Language: A Study of the Speech Act Performance of Learners of Hebrew as a Second Language. *Applied Linguistics*, *III*(1), 29-59. doi: 10.1093/applin/III.1.29

Blum-Kulka, S., House, J. & Kasper, G. (Eds.) (1989). *Cross-cultural pragmatics: Requests and apologies* (*Advances in Discourse Processes, XXXI).* Norwood, NJ: Ablex.

Bouton, L. F. (1988). A cross-cultural study of ability to interpret implicatures in English. *World Englishes*, *17*, 183-196.

Bouton, L. F. (1994). Conversational implicature in a second language: Learned slowly when not deliberately taught. *Journal of Pragmatics, 22*(2), 157-167.

Bouton L. F. (1999). Developing non-native speaker skills in interpreting conversational implicatures in English: Explicit teaching can ease the process. In E. Hinkel (Ed.), *Culture in second language teaching and learning* (pp. 47–70). Cambridge: Cambridge University Press.

Bromberek-Dyzman, K. & Ewert, A. (2010). Figurative competence is better developed in L1 than in L2, or is it? In M. Pütz & L. Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind* (pp. 317–334). Amsterdam: John Benjamins.

Brown, J. D. (2008). Raters, functions, item types and the dependability of L2 pragmatics tests. In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 224-248). Bristol: Multilingual Matters.

Brown, P. & Levinson, S. D. (1987). *Politeness: some universals in language usage*. Cambridge: Cambridge University Press.

Carrió Pastor, M. L. & Martín Marchante, B. (2016). Identificación de errores pragmáticos en un test adaptativo de inglés como lengua extranjera. *ELIA, 16*, 79-102. doi: 10.12795/elia.2016.i16.04

Cohen, A. D. & Shively, R. L. (2007). Acquisition of requests and apologies in Spanish and French: Impact of study abroad and strategy-building intervention. *The Modern Language Journal*, *91*(2), 189-212.

Cohen. D. (1996). Law, social policy, and violence: The impact of regional cultures. *Journal of Personality and Social Psychology, 70,* 961-978.

Contreras, E. (1990). *El profesor universitario y la evaluación de los alumnos*. Madrid: Universidad Politécnica de Madrid.

Coulmas, F. (Ed.) (1981). *Conversational routine*. The Hague: Mouton.

Dulay, H., Burt, M. & Krashen, S. (1982). *Language Two*. Oxford. Oxford University Press.

Félix-Brasdefer, J. C. (2007). Natural speech vs. elicited data: A comparison of natural and role play requests in Mexican Spanish. *Spanish in Context*, *4*(2), 159-185.

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

Gallardo Pauls, B. (2005). El lenguaje y la mente. In A. López & B. Gallardo (Eds.), *Conocimiento y lenguaje*. Valencia: Universitat de Valencia.

Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics, Speech Acts* (pp. 41–58). New York, NY: Academic Press. Retrieved June 10, 2017 from http://www.sil.org/linguistics/bibliographylinguistics/ColePAnd Morgan 1975.htm.

Grice, H. P. (1989). *Studies in the Way of Words*. Harvard, MA: Harvard University Press. Retrieved February 10, 2017 from http://courses.media.mit.edu/2005spring/mas962/Grice.pdf

Gumperz, J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.

Halliday, M. A. K. (1979). *El lenguaje como semiótica social*. México: Fondo de Cultura Económica.

House, J. (1996). Developing pragmatic fluency in English as a foreign language: Routines and metapragmatic awareness. *Studies in second language acquisition, 18*(2), 225-252.

Kasper, G. (2006). Speech acts in interaction: Towards discursive pragmatics. In K. Bardovi-Harlig, J. C. Félix-Brasdefer & A. S. Omar (Eds.), *Pragmatics & Language Learning,* Vol. 11 (pp. 281–314). Honolulu: University of Hawai'i at Manoa.

Kasper, G. & Schmidt, R. (1996). Developmental issues in interlanguage Pragmatics. *Studies in Second Language Acquisition*, *18*, 149–69.

Keenan, E. O. (1976). The universality of conversational postulates. *Language in Society*, *5*(1), 67-79.

Kerbrat-Orecchioni, C. (1996). *La conversation*. Paris: Seuil.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT press.

Lewis, M. (1993). *The lexical approach: the state of ELT and the way forwa*rd. Hove: Language Teaching Publications.

Lewis, M. (2000) Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocation: further developments in the Lexical Approach* (pp. 126-154). Hove: Language teaching publications.

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

Madrid, D. (1997). La evaluación del área curricular de la lengua extranjera. In H. Salmerón (Ed.), *Evaluación Educativa: Teoría, metodología y aplicaciones en áreas de conocimiento* (pp. 252-290). Granada: Grupo Editorial Universitario.

Martín Marchante, B. (2015). *La evaluación de la competencia pragmática a través de una prueba adaptativa.* Unpublished Doctoral Dissertation, Universitat Politècnica de València, Spain.

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

Murray, J. C. (2011). Do Bears Fly? Revisiting Conversational Implicature in Instructional Pragmatics. *The Electronic Journal for English as a Second Language, 15*(2), 1-10.

Murray, N. (2010). Pragmatics, awareness raising and the co-operative principle. *ELT Journal*, *64*(3), 293-301.

Nattinger, J. & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Payrató, LL. (2010). *Pragmàtica, discurs i llenguatge oral. Introducció a l'anàlisi funcional dels textos.* Barcelona: Universitat Oberta de Catalunya.

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

Purpura, J. (2009). The Oxford Online Placement Test: What does it measure and how? Retrieved February 6, 2017 from https://www.oxfordenglishtesting.com/uploadedfiles/6_NewLook_and_Feel/Content/oopt_measure.pdf

Read, J. & Nation, P. (2004). Measurement of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 23-35). New York, NY: John Benjamins.

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

Roever, C. (2001). Web-based language testing: Opportunities and challenges. *Language Learning and Technology*, *5*, 84–94.

Roever, C. (2005). Validation of a web- based test of ESL pragmalinguistics. Retrieved February 22, 2016 from http://ltj.sagepub.com/content/23/2/229.full.pdf

Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing, 23*(2), 229-256.

Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, *28*(4), 463-481.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis.* Cambridge: Cambridge University Press.

Searle, J. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Sperber, D. (1986). *Relevance: Communication and Cognition*. Cambridge, MA: Harvard University Press.

Taguchi, N. (2006) Analysis of appropriateness in a speech act of request in L2 English. *Pragmatics, 16*(4), 513-533.doi: 10.1075/prag.16.4.05ta

Taguchi, N. (2008). Pragmatic comprehension in Japanese as a foreign language. *The Modern Language Journal*, *92*(4), 558-576.

Taguchi, N. (2011). Rater variation in the assessment of speech acts. *Pragmatics, 21*(3), 453-471.

Taguchi, N. (2012). *Context, individual differences, and pragmatic competence*. New York, NY/Bristol: Multilingual Matters.

Taguchi, N. & Roever, C. (2017). *Second Language Pragmatics*. Oxford: Oxford University Press.

Tanaka, S. & Kawade, S. (1982). Politeness strategies and second language acquisition. *Studies in second language acquisition*, *5*(1), 18-33.

Tsutagawa, F. (2013). Pragmatic Knowledge and Ability in the Applied Linguistics and second Language Assessment Literature: A review. *Teachers College Columbia University Working Papers in TESOL & Applied Linguistics*, *13*(2), 1-20.

Turnbull, W. (2001). An appraisal of pragmatic elicitation techniques for the social psycho- logical study of talk: The case of request refusals. *Pragmatics*, *11*, 31-61.

Walters, F. S. (2004). *An application of conversation analysis to the development of a test of second language pragmatic competence*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, IL. Retrieved January 10, 2018 from https://www.ideals.illinois.edu/handle/2142/79792

Walters, F. S. (2007). A conversation-analytic hermeneutic rating protocol to assess L2 oral pragmatic competence. *Language Testing*, *24*(2), 155–183.

Wilson, D. & Carston, R. (2007). Relevance, Inference and Ad Hoc Concepts. *A Unitary Approach to Lexical Prag*matics. Retrieved October 24, 2014 from http://www.phon.ucl.ac.uk/home/robyn/pdf/ Wilson-Carston-Unitary-Approach-2007.pdf

Yamashita, S. (2008). Investigating interlanguage pragmatic ability. In E. Alcón Soler & A. Martinez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 201–223). Bristol: Multilingual Matters.

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