

**Research Article**

# **Unveiling Pragmatic Proficiency: Harnessing Corpus Pragmatics for Effective Implicature Assessment in High School Language Learning**

**Dr. Aglaia Rouki**

National and Kapodistrian University of Athens, Greece

Email: rouki\_aglaia@hotmail.com

**Received:** September 30, 2023

**Accepted:** October 18, 2023

**Published:** October 26, 2023

## **Abstract**

This research delves into the realm of second language acquisition and pragmatic competence, with a specific focus on the intricate domain of conversational implicatures. The aim of this study was to design an effective assessment tool for implicature comprehension by harnessing the power of corpus pragmatics. Utilizing the COCA (Corpus of Contemporary American English) corpus, I crafted a comprehensive test on conversational implicatures. This test was administered to first-grade high school students, with the objective of evaluating both their comprehension and production abilities within pragmatic contexts. The study reveals the synergistic relationship between corpus pragmatics and digital technology, showcasing the seamless integration of authentic language use through online corpora. The results obtained from the test underscore the efficacy of this approach, with overwhelmingly positive outcomes observed among the participants. By illuminating the intersection of corpus pragmatics, technology, and language acquisition, this research not only contributes to the advancement of language pedagogy but also highlights a promising avenue for future language assessment methodologies.

**Keywords:** Pragmatics, Testing Pragmatics, EFL Testing, Corpus Linguistics.

## **1. Introduction**

Within the realm of second language acquisition, pragmatic competence emerges as a cornerstone for effective communication, underpinned by an intricate web of implicatures that bridge language and context. Corpus pragmatics, a pivotal facet of this study, delves into the subtleties of how language is used in authentic communication. This paper embarks on an exploration of the innovative integration of corpus pragmatics and digital technology, specifically online corpora, to construct testing materials for implicature assessment. By uniting comprehension and production tasks, we navigate the nuanced terrain of implicature understanding and expression, while also addressing the challenges learners face in pragmatic contexts. In a landscape where digital literacy converges with language learning, this study ventures into a dynamic intersection, forging new frontiers in language assessment and nurturing pragmatic awareness.

## **2. Literature Review**

### **2.1. A Corpus-Based Approach to L2 Pragmatics**

Pragmatics and corpus linguistics are two domains of research that were initially regarded as mutually exclusive. However, this perception has now changed and common ground has been discovered, leading to the establishment of a new field that is called corpus pragmatics. Corpus pragmatics is defined as “the science that describes language use in real contexts through corpora” (Romero-Trillo, 2017:1). It refers to the study of actual language use that is based on large, computerized collections of language and is regarded as a kind of empirical data based on pragmatics.

Several studies have emphasized the need to raise L2 learners’ pragmatic awareness vis-à-vis the use of naturally-occurring discourse (Schmidt, 1993a,b; Kasper, 1997; Rose, 2000; Eslami-Rasekh, 2005). In Ifantidou (2011a,b; 2013a,b; 2014), pragmatic awareness was defined and tested for the first time in terms of an open-ended array of pragmatically inferred implicatures rather than as a fixed set of routines (Ifantidou, 2011a,b). In this direction, corpora could prove valuable in order to raise pragmatic awareness in EFL learners (Taguchi and Roever, 2017). The global context of sociocultural assumptions, as offered by online corpora, is a facilitating tool because it allows access to real-life settings which trigger more

spontaneous responses (Schauer and Adolphs, 2006; Roever, 2006; Chambers, 2007; Römer, 2009a,b; Ishihara and Cohen, 2010; Taguchi, 2015; Furniss, 2016; Vyatkina, 2016a,b; Bardovi-Harlig, *et al.*, 2017; Vyatkina and Boulton, 2017; Boulton and Cobb, 2017).

## **2.2. Testing Pragmatic Awareness**

Testing pragmatic awareness of second language is a relatively recent enterprise and an underexplored but growing area within second language assessment. The existing literature on tests of pragmatic awareness indicates that the different testing formats vary in terms of their effectiveness and the variables used (Brown, 2001a, b). According to Roever (2011), tests have mainly focused on assessing learners' sociopragmatic and pragmalinguistic abilities. The Speech Act framework for forming tests in interlanguage pragmatics has been criticized for not assessing learners' ability to produce extended monologic and dialogic discourse, thus a re-orientation of pragmatic testing is required. This is the main reason why this framework was not adopted in the present research. Next, I present the main methods used in assessing L2 learners' pragmatic awareness and justify why I chose to incorporate certain of those in my own research.

The main methods of testing pragmatics in an educational context could be divided into five categories. The first one is "Multiple-choice Discourse Completion Tasks (MDCT)", which require the learners to read a situation description and choose how they would continue an utterance. Secondly, "Oral Discourse Completion Tasks (ODCT)" request learners to listen to an orally described situation and record how they would continue it. "Discourse Role-Play Tasks (DRPT)" ask the learners to read the description of a situation and then enact a particular role with the L2 teacher in the situation given. In a similar vein, in "Discourse Self-Assessment Tasks (DSAT)" learners read a written description of a situation and then evaluate their own pragmatic ability to respond correctly to the situation. Finally, in "Role-Play Self-Assessment (RPSA)" learners rate their own performance in the recording of the role play in the DRPT.

Hudson *et al.*, (1992, 1995) were the first to introduce pragmatic tests and distribute them to EFL learners at a US university. The results of those tests were quite promising. Yamashita (1996) applied the Japanese version of the tests and pointed out that out of the 5 tests only the MDCT worked in a satisfactory way for Japanese as a second language. Enochs and Yoshitake (1996) also concluded that the same test types worked well for Japanese university EFL learners. Ahn (2005) applied MDCT to Korean EFL learners and was led to satisfactory results and Liu (2010) found MDCTs useful when having learners generate the speech acts and the situations in which they were used.

For the purposes of my research, I mainly used MDCTs but also other types of tasks that were not mentioned in the list above, such as True/False tasks, open questions or making a judgment-tasks, since these can be easily combined with the material available in corpora. Although MDCTs have been shown to be the most convenient in terms of practicality at the levels of both administration and scoring (Roever, 2011), and are particularly favorable in terms of assessment of pragmatic awareness, I decided to incorporate other kinds of tasks as well, such as open-ended tasks, which reveal how the respondents think about a question; as a result, their responses can be used to expand on and clarify closed responses. This is also the main criticism that MDCTs have received (Brown, 2000), namely that the given options may confuse the respondents, thus not providing information on whether they actually understood the question or simply answered at random.

My aim was to create a test that would use solely authentic material from a variety of contexts and expose L2 learners to a variety of tasks in order to draw conclusions regarding which tasks work most effectively for pragmatic assessment purposes. It is my view that this is an aspect of the present research that may contribute to the field of pragmatics testing and, in particular, the fact that I have attempted to take advantage of all the merits of a wide range of tasks and have also included both closed-ended and open-ended tasks based on real-life instances of language use found in the corpus I employed.

As far as the format of existing tests is concerned, the majority use paper-and-pencil testing formats (Hudson *et al.*, 1992), whereas some others make use of other, less prevalent, types. More specifically, Tada (2005) created computer-delivered tests with video prompts. Roever (2005, 2006) and Itomitsu (2009) developed web-based testing. Rylander *et al.*, (2013) focused on video formats while Timpe (2013) used Skype role-play tasks. I adopted the paper-pencil format, although the video format could also be beneficial, given that the test cannot last for more than one teaching hour and the Greek public high-schools lack adequate technological equipment for all classes in order to make use of other formats that incorporate videos or require internet connection. Based on prior evidence from testing pragmatic awareness of L2 learners, grammatical development does not guarantee a corresponding level of pragmatic development (Bardovi-

Harling and Dornyei, 1998); moreover, even advanced learners often fail to understand and convey the speakers' intentions and politeness values. Therefore, language use is essential in understanding language that is appropriate to situations, users and the message to be conveyed. The responsibility for teaching the pragmatic aspects of language use falls on teachers, who have to face certain challenges, such as lack of sufficient and proper material and training in EFL pragmatics (Eslami-Rasekh, 2005), in how to raise learners' awareness of pragmalinguistic forms and sociocultural norms of interaction and in how to guide learners' observations and discovery of pragmatic rules (Cohen and Ishihara, 2005a, 2005b).

### **2.2.1. Testing Pragmatics with the Use of Corpora**

Except for teaching purposes, corpora have also been used for testing and language assessment. As Park (2014) observed, corpora started being used in language assessment in the 1990s and, since then, test developers have increasingly used them as a source of reference. Various types of corpora, such as large representative corpora, learner corpora or specialized corpora, have been actively used to systematically compare the linguistic features associated with expert users with those encountered in an EFL learner's language. When it comes to EFL pragmatic assessment, the use of corpora is not so wide. A number of representative studies in this field are presented below.

Romero-Trillo (2002) examined the phenomenon of "Pragmatic Fossilization" as one of the major problems that non-native speakers of English face in the learning process. Fossilization refers to the persistence of grammar errors in non-native speakers (Selinker, 1972). Hyland (2002) conducted research on pronoun usage and tested how 40 undergraduate Chinese speakers of English used personal-author pronouns in their academic writings. In his research, he used two corpora, an 'expert' corpus of 240 published journal articles and a 'novice' corpus of 40 project reports written in English by final-year undergraduates in Hong Kong. The results indicated that there were 12 author pronouns (he/she) per text in the 'novice' corpus and 20 in the 'expert' corpus. Also, in the expert corpus there was a significant disciplinary variation with 75% of author pronouns occurring in the social sciences and humanities, whereas sciences and engineering accounted only for 25%. Nevertheless, the 'novice' corpus lacked this variation, since expert writers were three times more likely to use author pronouns in their text than EFL learners. This can be explained by the impersonal portrayal of academic writing in textbooks and style guides. Hyland (2002) advocated that a pragmatic awareness-raising approach where learners will critically evaluate the use of 'I' in their own writing might prove beneficial.

Carrió-Pastor (2016) aimed to identify what aspects of pragmatic knowledge appear at different stages of language learning. For this reason, a corpus comprising of 100 English essays written by Spanish learners of English was created, where 50 essays were at B1 level of proficiency and 50 at B2 level. Focusing specifically on EFL language learners' use of hedges, their aim was to test whether the use of corpora of spontaneously produced written and oral speech could help identify pragmatic knowledge which is associated with different stages of second language learning. The findings indicated that the use of hedges is significantly different depending on the learners' level of proficiency. Thus, the learners' communicative effectiveness was partially associated with their use of hedges and, for this reason, instruction should focus on tasks which raise meta-discursive awareness.

A number of studies on pragmatic testing focused on the use of discourse markers. More specifically, Muller (2005) tested how English and German adult learners use the discourse marker "you know" with the aid of Giessen Long Beach Chaplin Corpus, which consists of recordings of English and German-speaking university learners. Muller identified five functions of these discourse markers-namely "imagine the scene", "see the implication", "reference to shared knowledge", "appeal for understanding" and "acknowledge that the speaker is right"-and found that for two of these functions ("see the implication" and "appeal for understanding") there was no significant difference between German students learning English and native speakers of English. The rest of the functions of "you know" differed considerably.

Huang (2018) conducted a corpus-based study to assess the use of the discourse marker 'well' by Chinese learners of English and compared its frequencies in native speaker data and in Swedish EFL learners. She used the Ubuntu dialogue corpus, a large, publicly-available dialogue-corpus that makes it feasible to build end-to-end deep neural network models directly from the conversation data. The results indicated that while Swedish EFL learners overuse 'well', Chinese-speaking learners, and especially those of upper-intermediate level, significantly underuse it. Huang (2018) concluded that the different L1s influence the use of discourse markers by EFL learners and considered possible pedagogical implications for different first languages and proficiency levels as well as their possible applications to the classroom-instruction of "well".

Overall, although corpora have been used by various researchers to identify stages in language learning and learners' needs and differences in acquisition of language, such as by Granger and Meunier (1994); Granger *et al.*, (2006); Chen (2006); Granger and Vander (2007); Granger (2009); Granger and Paquot (2009, 2011) and Granger and Gilquin (2011), most of the studies have focused on determining learners' proficiency with reference to different genres or to different stages of language acquisition. It seems that little attention has been paid to testing pragmatic awareness by detecting and classifying, for example, errors produced by learners' pragmatic failure (Carrió-Pastor and Mestre-Mestre, 2013a, b). A possible reason for this could be that pragmatic failure is not easily detectable and thus tested. Some researchers even state that learners acquire pragmatic proficiency in their L1 and for this reason this is not of interest to second language teaching (Kasper and Rose, 2002; Dahl, 2004; Björkman, 2011).

### **2.3. Popular Corpora Used for EFL Purposes**

As was suggested in the previous section, several definitions have been offered in the literature regarding what a corpus is. I follow Johansson (1998) according to whom "a corpus is a collection of texts selected and put together in a principled way" (Johansson, 1998:3). In other words, a text corpus is a relatively large collection of texts which have been produced by actual users and can be useful in analyzing how language is really used. A corpus can be categorized according to various criteria, such as source of content, metadata and presence of multimedia or its relation to other corpora (Tognini-Bonelli, 2002). For the purposes of my research, I decided to use a written, monolingual corpus which does not focus on a specific genre.

#### **2.3.1. COCA and Its Merits**

The aim of this section is to present the major advantages of COCA in order to justify its use for the purposes of my research. In order to provide learners with a handy tool for the use of corpora towards raising pragmatic awareness, an appropriate, user-friendly and freely-accessible corpus had to be selected. Having rejected other corpora for reasons that were presented in the previous section, I will next present COCA, which is the largest English corpus, and its benefits over other corpora.

As already stated, COCA is a free, online and easily accessible corpus of 1 billion words which practically means that it provides data of lower-frequency items that cannot be encountered in other corpora, such as the BNC. Furthermore, in terms of collocates, there are 14 times as many in COCA that occur more than 5 times compared to those in the BNC. Another characteristic of COCA that made it the ideal choice is that it is an up-to-date corpus, as, since the early 1990s, 20 million words per year have been added, which is an important indication that it represents contemporary English. As a user, I can easily search both for single words and for collocates within a ten-word window span and compare collocates of two related words (Römer, 2009b).

Another benefit is the fact that COCA includes one billion different texts that come from a variety of sources and genres. This is an important criterion because I aimed to provide my learners with extracts from various types of texts, such as fiction, popular magazines and newspapers (Davies, 2008). Finally, COCA can display example sentences together with frequency searches. These sentences, centered around one key word, serve as an ideal input to observe how words fit and draw conclusions about both their actual and implicated meanings through surrounding words (Scott, 2004). When searching, for example, for the word 'petrol' we are first provided with the number of times this word is encountered in the corpus, followed by the exact contexts where this word is found.

## **3. Methodology**

### **3.1. Setting**

The study took place in the 2<sup>nd</sup> General High School of Piraeus, which is a public school located in the center of Piraeus. It is worth noting that the majority of the learners live in this area. About 225 students attend this school and about 75 of them are in the first grade. I specifically worked with two classes of the first grade consisting of 25 and 23 students, respectively.

### **3.2. Participants**

The participants in the study were twenty 15- to 16-year-old students currently attending the first grade of junior high school with an overall C1 level of English, according to CEFR (Common European Framework of Reference for Languages). The students who participated in my research were selected on the basis of the following two criteria: Firstly, their level had to be at least C1 and, secondly, they had to be willing to participate. The students' level was identified after an English language test had been distributed to two classes of this grade. The sample was equally divided into 10 boys and 10 girls, all of whom had only

attended public schools in their entire school life. All of them had obtained a B2 English Certificate (20% Cambridge First and 80% ECCE) and continued their English lessons either privately or in English language schools (frontistiria) to take the proficiency exam. Three of them (15%) had already obtained the ECPE Certificate of English.

The number of 20 participants for the purposes of the study was deemed to be adequate given that the majority of the tasks of the study tended to have normal distributions. Therefore, although the number of participants in the study was less than 30, which, according to the Probability and Statistical Inference, is the minimum number of participants in order for a study to be statistically significant (Hogg and Tanis, 1997), it did not influence my research hypothesis in a negative way and allowed me to draw safe conclusions regarding the materials I developed. Moreover, the fact that in the methodology I employed I paid more emphasis on the results of the main study, which included 30 participants, and I used the pilot study as an indicator of the efficiency of the teaching and testing material, justifies my choice of having fewer participants in the pilot study.

### **3.3. Research Instrument**

#### **3.3.1. Test on Implicatures**

The first task was an MCQ task consisting of 3 testing items. The total number of points that learners could collect from this task were 6 (2 for every testing item). Each testing item required learners to read a 5- to 6-line context and understand the meaning of a highlighted adjective appearing in it. The learners were provided with 4 alternative answers and they were asked to choose the best option. For example, the first testing item included an irony regarding the use of the word “great” referring to a piece of news. Based on the context provided, the learners had to infer that the use of the word “great” was ironical and that it was actually intended as “really bad”. Some of the chosen adjectives of this task were polysemous and apart from their prototypical meanings also had extended metaphorical meanings. The learners were required to read the context carefully in order to understand which of these meanings was inferred in each case. For instance, the last testing item included the phrase “hot cuisine” and the learners had to infer that out of all the metaphorical meanings of the adjective “hot” the one implied in the relevant task was that of “widely discussed”.

Task 2 was another MCQ task also consisting of 3 testing items. The total number of points the learners could collect from this task were 6 (2 for every testing items). For this task, the learners were required to read 3 different contexts, each consisting of 5-6 lines, and pay attention to a specific highlighted phrase in each case. Based on the contexts, they had to choose 1 out of 4 alternative choices, namely from “a” to “d”, which best illustrated the meaning of the highlighted phrase. For example, the second testing item of this task drew the learners’ attention to the phrase “my soul was bleeding”. They needed to read the context carefully and understand that the meaning of this phrase related to the speaker being extremely sad and, thus, they had to choose “c” as the correct answer.

Task 3 was in the form of dialogues, each consisting of 3 turns. Based on the context provided, the learners had to judge whether the answer provided by one of the two speakers was relevant or irrelevant to the question posed by the other speaker. Subsequently, they had to justify their answer. The task included 3 dialogues and the learners got 1 point for each correct answer and 2 points for each correct justification they provided. Therefore, the total number of points they could gather were 9. For instance, the first dialogue of the task included the question “how old are you?” and the answer “what? I am offended!”. The learners had to assume that the answer was relevant and justify their answer accordingly.

Task 4 was an open-form exercise consisting of 3 testing items. The learners were asked to read the 5-6-line context carefully and understand the intended meaning of certain phrases. They had to explain their view in 1-2 lines. For example, the third testing item included the phrase “I am the bread of life” meaning that “I am a really basic and essential part of your life”. The total number of points the learners could collect for this task was 6 (2 for every correct answer).

For task 5, the learners had to choose between two options and also justify their answer. More specifically, for each testing item they were provided with two 5-6-line contexts with a common highlighted word. In one of the two contexts, the meaning of the word was literal and in the other metaphorical. Based on the context provided, they had to assume which was the literal and which was the metaphorical use and justify their answer. For instance, the first testing item examined the use of the word “doll”. In the first case, the word was used metaphorically, in the sense of “a very beautiful and delicate young lady” and in the second case it

was used literally, namely “toy”. The learners received one point for each correct answer and 2 points for every proper justification (9 points in total for this task).

The last task of the test, task 6, was a True/False task consisting of 4 testing items. For each correct answer, the learners received 1 point (4 points in total for this task). Again, for each item in the task, the learners were provided with a 5- to 6-line context which they had to read carefully and then decide if a sentence about the text, which I provided them with, was true or false. In order to draw learners’ attention especially on the implicatures included in the context, I highlighted certain phrases that constituted examples of irony. For example, in the second testing item of this task, the highlighted phrase was “never been better” and the given assumption was that “something is wrong with the woman but she just doesn’t say”. This assumption was true, since the woman answered ironically to a question regarding how she felt, saying that she was fine, whereas in reality she was not.

### **3.3.2. Rationale behind the Test on Implicature**

The aim of this section is to present the rationale behind the creation of the test on implicature and present how these tests are different from the limited number of tests on implicature created by other researchers.

The seven-page test addressed high-school students aged between 15 and 16 years of age at C1 level of English. Having acquired a level of, at least, C1 was essential in order for the participants to be able to understand the concept of implicature and retrieve it based on the realistic context available, given that the tests were based on original English texts retrieved from the corpus rather than on adapted versions.

For the test, I created six tasks using data that I had retrieved from COCA. The choice of six tasks in the tests was due to time-limitation (45 minutes). I also considered that creating shorter tests would not allow me to draw safe conclusions about the learners’ understanding of implicature. Furthermore, the tasks involved open-ended tasks, MCQ tasks, underline-the-sentence tasks and True/False tasks. The variety of task formats would cater for learners’ learning styles and preferences. On the one hand, closed-ended tasks were chosen because they are easy and quick to answer, and therefore friendlier to the learners, improve the consistency of responses and can also be measured. However, in this type of questions, respondents could always pick one answer at random. For this reason, I decided to also include a number of open-ended tasks, since they allow more in-depth answers which reveal what respondents think with greater accuracy (Farrell, 2015). Every task involved 3 cases of implicature, with the exception of the last task that included 4 items. According to the statistician, this number was considered satisfactory in order to draw conclusions while not being overly tiring or extended for the learners to complete.

Of all the genres in the corpus, I used fiction and articles retrieved from newspapers and articles. Especially fiction was an ideal source of implicature as it includes many instances of implicit use of language. I used ironies, hyperboles, metaphors (equal numbers of conventional and creative metaphors) and indirect answers, since, according to Allott (2018), these are the most common types of implicatures in English texts. These were selected from a variety of contexts, levels of formality and topics, such as formal newspaper articles on politics, environment or technology, restaurant or film reviews, semi-formal opinion articles on social issues, informal dialogues between friends or even ‘slang’ language in every day discussions.

Regarding the length of context provided in every task, the original 5- to 6-line context as occurring in the corpus was preserved for reasons of uniformity across test items and tasks. Limited changes were considered necessary in order to replace certain words, which, based on my teaching experience, would be unknown to C2 learners, an assumption a native speaker of English verified (such as the word “mulct” meaning “to fine somebody”), and to correct certain grammatical structures that were problematic (e.g. “I doesn’t know” instead of “I don’t know”).

Overall, I included a variety of tasks, such as open-ended tasks, MCQ tasks, underline-the-sentence tasks and True/False tasks. The test does not focus on one particular type of implicature, but rather on various types, such as indirect answers, ironies and metaphors retrieved from a variety of genres, such as articles, literature and theatrical scripts. The authenticity of the test items, the length of the context (Bezuidenhout and Cutting, 2002), the variety of tasks included in the test and the wide range of pragmatic phenomena covered allowed learners to gain a more global view of what an implicature is and enabled them to spot it in various contexts. To conclude, I aimed to design testing tools that were ‘global’, ‘all-inclusive’ and realistic in that they covered various types of testing tasks, types of contexts and implicatures.

## 4. Results

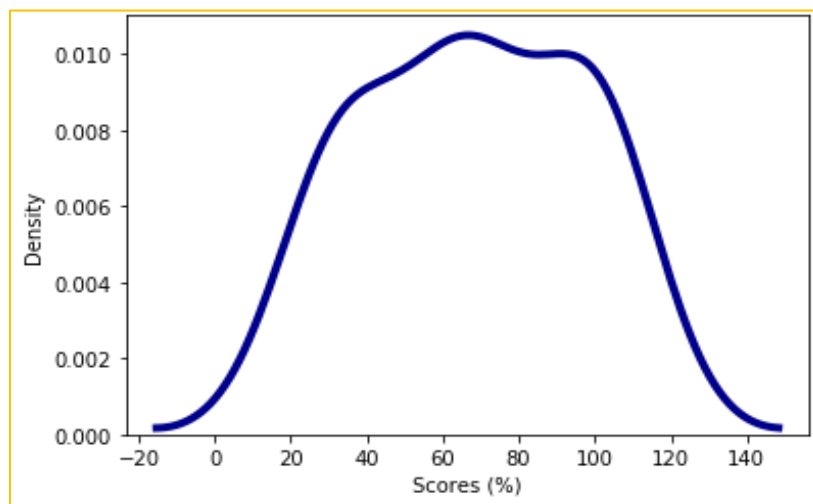
The aim of this section is to present the results of the test administered during the study as well as their qualitative evaluation. For the purposes of the quantitative analysis, *Anaconda* was used. The main reasons I decided to use it are its free and open-source availability and its simplified package management and deployment. Moreover, it provided me with tools to easily collect the data and an environment that was easily manageable for deploying my research.<sup>1</sup>

### 4.1. Quantitative Analysis

In this presentation of test results, I start with an analysis of each task, followed by my conclusions regarding the participants' overall performance. For the purposes of the quantitative analysis of the pilot study tests, the mean ( $\bar{x}$ ) value has been used.<sup>2</sup> The density plots that follow aim at visualizing the distribution of the participants' scores in every task. The peaks of the density plots display where values are concentrated over the interval, while it is worth noticing the skewness of the data distribution, which is a measure of the asymmetry of an ideally symmetric probability distribution.

Skewness is a measure of how much the probability distribution of a random variable deviates from the normal distribution. A positively skewed distribution is a type of distribution in which most values are clustered around the left tail of the distribution while the right tail of the distribution is longer. On the contrary, a negatively skewed distribution is a type of distribution in which more values are concentrated on the right side (tail) of the distribution graph while the left tail of the distribution graph is longer (Hosking, 1992).

Task 1-implicature synonym required participants to understand the meaning of an adjective in a given context and choose a synonym from a set of given options (a to d). The task consisted of 3 items. Seven out of the 20 participants (35%) managed to respond correctly to all three items. Seven out of 20 (35%) responded correctly to 2 out of 3 items and 6 out of 20 (30%) responded correctly to only 1 item (Figure 1).

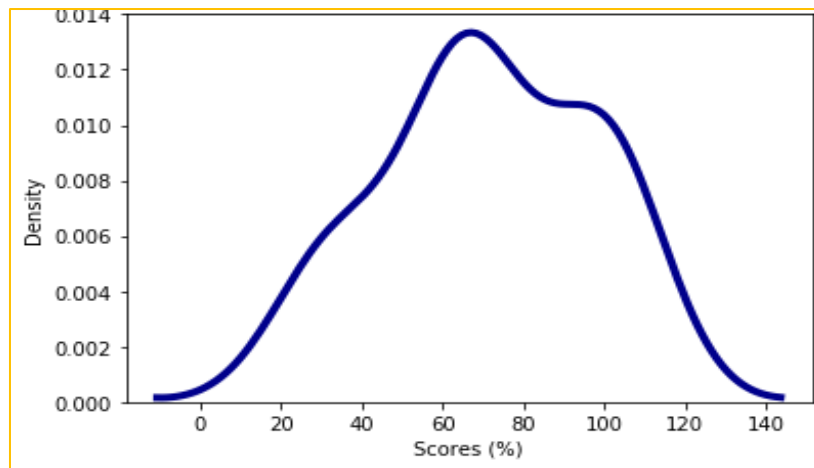


**Figure 1. Participants' scores distribution in Task 1.**

Task 2-Implicature at phrase-level requested participants to choose, from a list of options from "a" to "d", the one that best described the meaning of three implicatures in the form of phrases, namely "walking encyclopedia", "My soul was bleeding" and "a weaponization of the language of diversity", as presented in their context of occurrence. Seven out of the 20 participants (35%) managed to respond correctly to all three items. Nine participants (45%) responded correctly to 2 out 3 items and 4 out of 20 (20%) responded correctly to 1 item (Figure 2).

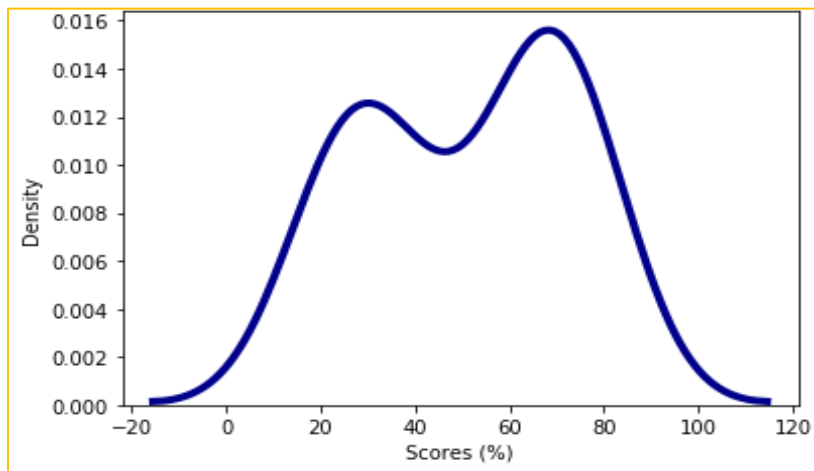
<sup>1</sup> Of the two languages offered, I used Python. I also made use of Anaconda Navigator, which is a desktop graphical user interface (GUI) included in Anaconda distribution. For the creation of the graphs, I used the application Jupyter Notebook, which is an application available by default in Navigator, and, more specifically, the Matplotlib library.

<sup>2</sup> The mean value (or average) is the sum of the values divided by the number of values while the median value is a value separating the higher half from the lower half of a data sample, a population or a probability distribution (Zwillinger, 1995).



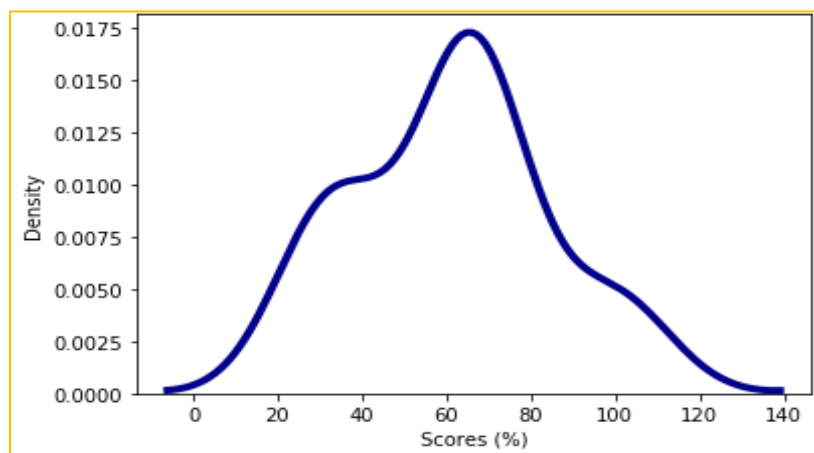
**Figure 2. Participants' scores distribution in Task 2.**

Task 3-Spot the relevant answer required participants to read three dialogues and judge whether an implied answer to a given question in the dialogue was relevant or irrelevant. Apart from answering correctly, the participants also needed to justify their answers. Four participants (20%) scored 77%, 6 participants (30%) scored 66%, 1 participant (5%) scored 55%, 1 more participant (5%) scored 44%, 4 participants (20%) scored 33% and 4 out of the 20 participants (20%) scored 22%, which was also the lowest score (Figure 3).



**Figure 3. Participants' scores distribution in Task 3.**

Task 4-Paraphrase implicature consisted of 3 items and asked participants to express in their own words what the three speakers actually meant by the intended implicatures. Only 3 participants (15%) responded to all the items correctly, 11 participants (55%) managed to respond correctly to 2 out of the 3 items and 6 participants (30%) provided 1 correct answer (Figure 4).



**Figure 4. Participants' scores distribution in Task 4.**



In Task 5-Literal/Metaphorical word, participants were asked to judge whether the use of a word in a given context was intended literally or metaphorically and to justify their answer. Once more, this task comprised 3 items (3 pairs of contexts). The highest score in this task was 88%, which was achieved by only 1 participant (5%) followed by 7 participants (35%) scoring 77%. Nine participants (45%) scored 66%, 1 participant (5%) scored 55% and 2 participants (10%) scored 33% (Figure 5).

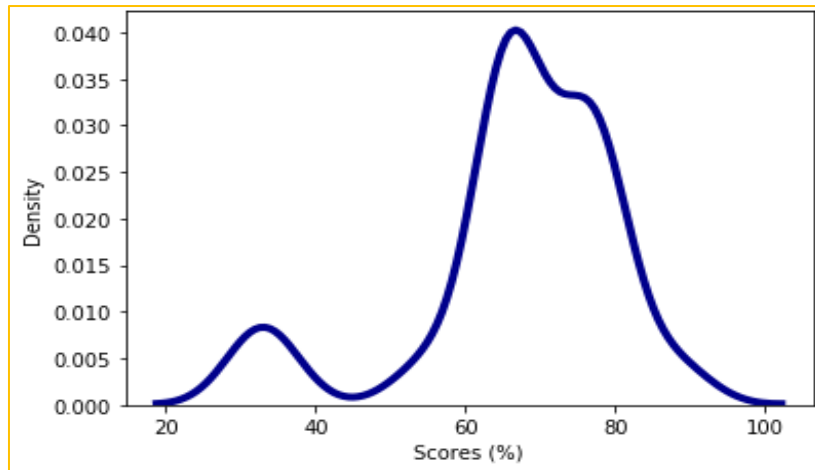


Figure 5. Participants' scores in Task 5.

The last task of the test (Task 6-true/false assumption) consisted of 4 items. Participants were asked to read four contexts and judge whether an assumption provided by me under each context was true or false. Four out of the 20 participants (20%) responded correctly to all items, 5 participants (25%) scored 75%, 6 participants (30%) scored 50% and 5 more participants (25%) scored 25%. The next Figure (Figure 6) summarizes the aforementioned results.

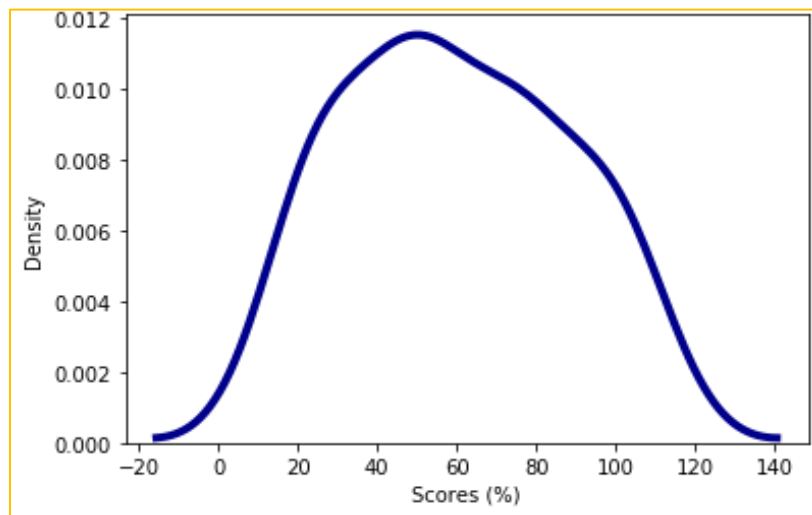
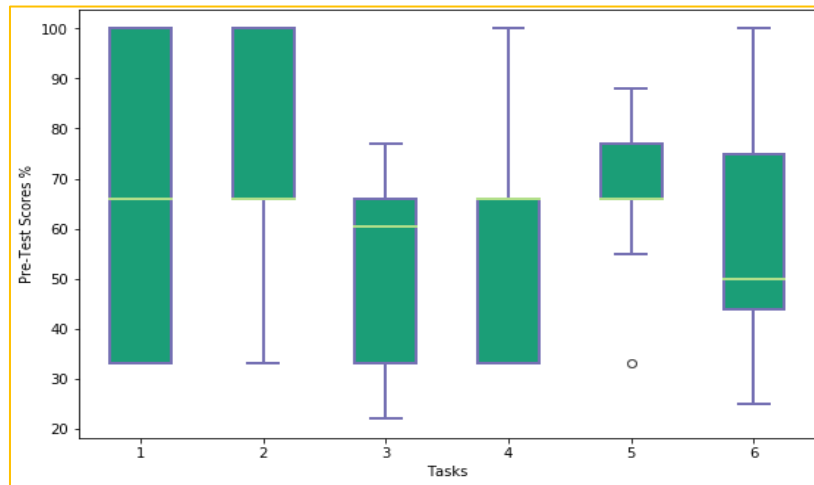


Figure 6. Participants' scores in Task 6.

As depicted in the following boxplot of the data (Figure 7), the task with the most normal distribution is Task 1, whose median value is 66%, its maximum score is 100% and its minimum score is 33%. Furthermore, Tasks 2,4 and 5 share the same median value ( $M=66%$ ), but exhibit different distributions. For Tasks 2 and 4, the maximum score is 100% and the minimum score is 33% while for Task 5 the minimum score is 55% and the maximum score is 90%. It is also worth mentioning that Task 5 includes an outlier below the lower quartile with the score of 33%, which was achieved by 2 participants. This score is different from the majority of the other scores. These three tasks also exhibit the least balanced distribution, since the middle 50% of the scores (Interquartile Range-IQR) are either only above the median value, ranging from 66% to 100% for Task 2 and from 66% to 77% for Task 5, thus making the distribution positively skewed, or only below the median score and range from 33% to 66% for Task 4, thus making the distribution negatively skewed. The task with the greatest distance between the maximum and the minimum score is Task 6, whose minimum score is 25% and maximum score is 100%. Regarding its median value, this is 50% and its

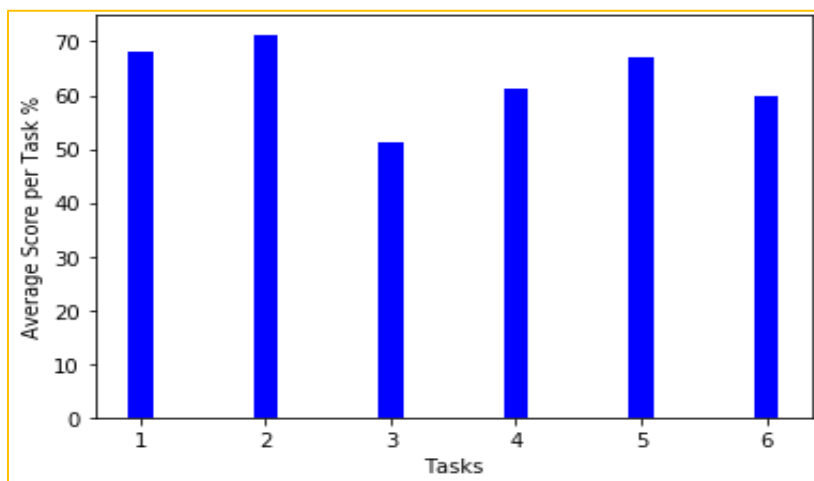
interquartile range (IQR) extends from around 45% to 75%, with the majority of participants scoring greater than the median score.

Finally, the lowest maximum and minimum scores are observed in Task 3, which caused the greatest difficulty to the participants. Its maximum score is 77%, which was achieved by 4 participants and its minimum score is 22%, which was also achieved by 4 participants. Its interquartile range is between 33 and 66% with the lower quartile (Q1) ranging from 33 to 60,5% and the upper quartile (Q3) ranging from 60,5% to 66%.



**Figure 7. Study-box plot of the test per task.**

Overall, based on the boxplot (Figure 7) and the average scores (Figure 8) of the test, it can be concluded that the task which caused the greatest confusion to the participants was Task 3, since their average score in this task was approximately 50% ( $\bar{x} = 51.15\%$ ). Task 4 and Task 6 were found almost equally demanding, as their average scores were approximately 60% (Task 4:  $\bar{x} = 61.2\%$ , Task 6:  $\bar{x} = 60\%$ ). Task 1 and Task 5 could be deemed to be easier, since the average scores were approximately 68% (Task 1:  $\bar{x} = 68\%$ , Task 5:  $\bar{x} = 67.1\%$ ). Finally, the least challenging task was Task 2 as the average score was approximately 72% ( $\bar{x} = 71.3\%$ ). The following bar-chart depicts the mean score of every task (Figure 8).



**Figure 8. Average scores of the test per task.**

In the next section, I offer a qualitative analysis of the data collected through the test in an attempt to complement the quantitative analysis presented above, draw conclusions as to what led participants to provide those specific answers and explain the possible factors that influenced their understanding of the implicatures included in the test.

#### **4.2. Qualitative Analysis**

According to test developed, which included a variety of both comprehension and production tasks in order to take advantage of the merits of both types, it is self-evident that the best way for learners to be able to

speak and write in the L2 is to produce language (Allen *et al.*, 1990). Secondly, the production tasks gave me the opportunity to accurately evaluate the success of my lessons in terms of the learners' active participation (Krashen, 1987). Thirdly, the learners' erroneous responses provided opportunities for corrective feedback, since 'errors' are conceived as problems of production (Swain, 1985, 1993). On the other hand, the comprehension tasks also proved to be very useful, since, as has been claimed, they resulted in increased motivation (Asher, 1977), reduced anxiety and greater likelihood that they would continue practicing (Newmark, 1966, 1971).

As for the types of tasks included, the ones that proved to be more illuminating were the reasoning-gap tasks, in which participants were asked to derive some information from the contexts provided, the opinion-gap tasks, in which participants were asked to convey their own personal views about a particular utterance and the MCQ-tasks, in which learners were asked to select the correct answer from the choices provided (Rabbanifar and Mall-Amiri, 2017).

As stated in the previous section, Task 3-Spot the relevant answer caused the greatest difficulty to the participants. In particular, 12 out of the 20 participants responded correctly without, however, providing a proper justification for their answers. At this point, it needs to be clarified that when I asked the participants to justify their answers, I did not expect them to use any metalanguage, because the current research was purely interested in raising learners' pragmatic rather than metapragmatic awareness and, therefore, it was not concerned with a meta-pragmatic analysis of the link between linguistic and overall relevance of the chosen texts. My intention was to use the justification question as a testing strategy for whether learners were able to infer implicatures and as a means of checking if the participants had answered at random, if the responses provided were relevant to the given questions or if they were able to correctly identify the pragmatically inferred effects from the contexts, which contributed to the creation of a stance towards the topic discussed (Ifantidou, 2014). For this reason, in order for the participants' answers to be regarded as correct no use of metalanguage was required. The only thing they needed to do was to answer correctly that the responses were relevant to the given questions and show that they had understood why they were so. If some participants managed to verbalize the link between the linguistic indexes and the relevant pragmatic effects-as Ifantidou (2014) defined metapragmatic awareness-this was regarded as a correct answer without, however, requiring all the participants to do so. For example, some correct justifications for the first item of the task (Q: Can you give me some simple specifics? A: How old are you?) included meta-language, such as *"this is an irony which intends to show that the speaker is too old to ask this kind of questions"* or *"this ironic answer is used to show that this kind of questions would be expected by a child and not an adult"*) while some other correct responses did not include any meta-language, such as *"this answer is relevant as the speaker wants to show that the question just asked was too childish for an adult"* or *"this is a proper answer which wants to show that such a question is not expected by a mature person"*. Some of the incorrect justifications, a number of which even included meta-language, were *"the answer is irrelevant since it is just an irony and not a real answer"* or *"the answer is not correct because it does not provide enough information for the other person's question"*. Answers such as these might be an indicator of the fact that learners are not adequately familiar with this type of implicature and further practice is required.

Secondly, Task 6-true/false assumption, also exhibited rather low average and median scores. In after-test interviews, participants who answered incorrectly stated that they did not pay adequate attention to the context provided and they just read the phrases in bold instead of the whole texts. They claimed that they had devoted too much time to the previous tasks of the test and they did not have adequate time to examine this last task carefully. Poor time management is considered to be an important factor that may lead to task failure. Learners who do not manage their time effectively and do not use it for the right purposes cannot realistically determine how much time each task requires and, therefore, some of their test questions remain unanswered (Cronk, 1987). Some others reported that they had already felt rather tired from the whole test and preferred to finish it as fast as possible without caring about their answers in the last task. This is probably why they were led to incorrect assumptions about the relevant implicatures. Tiredness is a potential source of bias that influences learners' performance on standardized tests that result from sustained cognitive engagement (Holding, 1983).

Task1-implicature synonym, Task4-Paraphrase implicature and Task 5 Literal/Metaphorical word, which were found relatively easier, focused on the learners' ability to distinguish the literal from the metaphorical use of a word. Given that many of the items provided were used with the same phrasing in their native language (e.g. the metaphor "The fate of your town is in your hands" is also used in Greek "η τύχη της πόλης είναι στα χέρια σου" to refer to someone who has power and whose decisions have a strong impact or the

word “hot” metaphorically used in the sense of “fashionable”), learners did not have difficulty understanding the intended meaning in the given context. A growing body of research in second language acquisition has been conducted on first language transfer. Almost all of the previous researchers believe that first language has interference in second language acquisition and many of them have concluded that the L1 can have a positive effect on L2 understanding and production, as shown in the cases presented above (Nation, 2001). However, as was expected, the more creative implicatures were found more difficult than the less creative ones. For example, the metaphors “the world was his oyster” or “I am the bread of life” received the most incorrect answers compared to the rest.

Finally, the highest average and median scores were obtained in Task 2-Implicature at phrase-level which exhibits one more case of positive L1 transfer (Nemati and Taghizade, 2006). Participants might have already been familiar with the two first phrases in bold (“walking encyclopedia” and “my soul was bleeding”), as there are similar metaphors in Greek (“κινούμενη εγκυκλοπαίδεια”, “η ψυχή μου μάτωνε”), but the more creative metaphor (“weaponization of the language of diversity”) caused some confusion, possibly also due to the relatively more demanding vocabulary.

## **5. Conclusion**

The ensuing discussion centers on the effectiveness of utilizing online corpora to design testing materials for implicature assessment. The study's outcomes are examined through the lens of integrating comprehension and production tasks, addressing challenges encountered by participants, and elucidating the broader implications for language teaching and learning.

The test on implicatures, formulated based on an online corpus, demonstrates the potential of leveraging digital resources to create targeted testing materials. The online corpus provides a vast collection of authentic language use, enabling the construction of tasks that mirror real-world language encounters. This approach, rooted in genuine linguistic contexts, enhances the ecological validity of the assessment and aligns with the principles of authentic language learning (Derakhshan and Eslami, 2015).

The study's test design artfully interwove comprehension and production tasks, resulting in a comprehensive assessment framework. This hybrid approach, drawing from both theoretical foundations (Kasper, 2007) and practical application, enables learners to actively engage in both receptive and productive language skills. The integration of these tasks promotes a holistic understanding of implicatures while fostering language output. Such synergy reflects the essence of communicative language teaching, emphasizing language as a tool for meaningful interaction (Murray, 2010).

The study's findings extend beyond assessment, reverberating in language pedagogy. The utilization of online corpora in test design underscores the symbiotic relationship between technology and language education. Educators can harness the richness of online language data to craft authentic assessment tasks that align with communicative goals. Furthermore, the integration of comprehension and production tasks fosters a holistic language learning experience, cultivating learners' abilities to comprehend and generate meaningful discourse.

The successful incorporation of an online corpus underscores the increasing importance of digital literacy in language education. Engaging with online corpora necessitates a level of technological proficiency, preparing learners to navigate the diverse linguistic landscape of the digital age (Johnson and De Haan, 2013). As language instruction evolves, the ability to access, interpret, and utilize online resources becomes an essential skill set for language learners.

In conclusion, the utilization of online corpora in designing testing materials for implicature assessment represents a pivotal step towards bridging the gap between authentic language use and assessment practices. The study's outcomes underscore the benefits of this approach, emphasizing the synergy of comprehension and production tasks, addressing pragmatic challenges, and highlighting the broader implications for language education in an increasingly digital world. As technology continues to shape language learning, the integration of online corpora paves the way for innovative, contextually grounded assessment methodologies that resonate with the complexities of real-world language interactions.

## **Declarations**

**Acknowledgements:** In the pursuit of my research, I wish to acknowledge the individuals who played a pivotal role in enabling its realization. My foremost gratitude goes to Prof. Elly Ifantidou, for her unwavering

support and dedicated supervision. Her profound academic acumen was instrumental in bringing this work to fruition. I am also indebted to Prof. Vasiliki Nikiforidou for her invaluable feedback, and to Dr. Georgios Mikros for his insightful suggestions during both the study composition and statistical analysis. I extend my heartfelt appreciation to those who aided and encouraged me throughout the study's preparation, including Ioannis Konstas, my friends, colleagues, and students.

**Author Contribution:** The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

**Conflict of Interest:** The author declares no conflict of interest.

**Consent to Publish:** The author agrees to publish the paper in *International Journal of Recent Innovations in Academic Research*.

**Data Availability Statement:** Not applicable.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The proposal for the study was approved by the Institutional Review Board of the National and Kapodistrian University of Athens.

**Informed Consent Statement:** Written informed consent has been obtained from the participant(s) to participate in this study.

**Research Content:** The research content of manuscript is original and has not been published elsewhere.

## References

1. Ahn, R.C. 2005. Five measures of interlanguage pragmatics testing in KFL (Korean as a Foreign Language) learners. Unpublished doctoral dissertation. University of a Hawai, Manoa.
2. Allen, P., Swain, M., Harley, B. and Cummins, J. 1990. Aspects of classroom treatment: toward a more comprehensive view of second language education. In: B. Harley, P. Allen, J. Cummins, and M. Swain (Eds.), *the development of second language proficiency*. Cambridge: Cambridge University Press: 57-82.
3. Allott, N. 2018. Conversational implicature. In: Aronoff, M. (Ed.), *Oxford Research Encyclopedia of Linguistics*, Oxford: Oxford University Press.
4. Asher, J.J. 1977. *Learning another language through actions*. California: Sky Oaks Productions.
5. Bardovi-Harlig, K. and Dörnyei, B. 1998. Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly*, 32(2): 233-259.
6. Bardovi-Harlig, K., Mossman, S. and Su, Y. 2017. The effect of corpus-based instruction on pragmatic routines. *Language Learning and Technology*, 21(3): 76-103.
7. Bezuidenhout, A. and Cutting, J.C. 2002. Literal meaning, minimal propositions, and pragmatic processing. *Journal of Pragmatics*, 34 (4): 433-356.
8. Björkman, B. 2011. Pragmatic strategies in English as an academic lingua franca: ways of achieving communicative effectiveness? *Journal of Pragmatics*, 43(4): 950-964.
9. Boulton, A. and Cobb, T. 2017. Corpus use in language learning: a meta-analysis. *Language Learning*, 67(2): 348-393.
10. Brown, J.D. 2000. *Principles of language learning and teaching*. TESOL Quarterly, 14.
11. Brown, J.D. 2001a. *Using surveys in language programs*. Cambridge: Cambridge University Press.
12. Brown, J.D. 2001b. Pragmatic tests: different purposes, different tests. In: K.R. Rose and G. Kasper (Eds.), *Pragmatics in language teaching*. Cambridge: Cambridge University Press, 310-327.
13. Carrió-Pastor, M.L. 2016. A contrastive study of interactive metadiscourse in academic papers written in English and in Spanish. In: F.A. Almeida, L. Cruz Garcí'a, and V.M. Gonza'lez Ruiz (Eds.), *Corpus-Based Studies on Language Varieties*: 89-114.
14. Carrió-Pastor, M.L. and Mestre-Mestre, E. 2013a. Increasing pragmatic awareness in the L2 classroom. *Studies and global perspectives of second language teaching and learning*. New York: Information Age Publishing.
15. Carrió-Pastor, M.L. and Mestre-Mestre, E. 2013b. A proposal for the tagging of grammatical and pragmatic errors. *Research in Corpus Linguistics*, 1: 7-18.
16. Chambers, A. 2007. Popularising corpus consultation by language learners and teachers. *Language and Computers*, 3-16.

17. Chen, C.W. 2006. The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, 11(1): 113-130.
18. Cohen, A.D. and Ishihara, N. 2005a. A web-based approach to strategic learning of speech acts. Minneapolis, MN: Center for advanced research on language acquisition, University of Minnesota.
19. Cohen, A.D. and Ishihara, N. 2005b. Strategies for learning and performing L2 speech acts. *Intercultural Pragmatics*, 2(3): 275-301.
20. Cronk, J. 1987. Time-management ideas for assistant principals. *NASSP Bulletin*, 71: 55-57.
21. Dahl, T. 2004. Textual metadiscourse in research articles: A marker of national culture or of academic discipline? *Journal of Pragmatics*, 36: 1807-1825.
22. Davies, M. 2008. The corpus of contemporary American English (COCA): 410+ million words, 1990 present. Available online at <http://www.americancorpus.org>.
23. Derakhshan, A. and Eslami, Z. 2015. The effect of consciousness-raising instruction on the pragmatic development of apology and request. *The Electronic Journal of English as a Second Language (TESL-EJ)*, 18(4): 1-24.
24. Enochs, K. and Yoshitake, S. 1996. Self-assessment and role plays for evaluating appropriateness in speech act realizations. *Language Research Bulletin*, 11: 57-76.
25. Eslami-Rasekh, Z. 2005. Raising the pragmatic awareness of language learners. *ELT Journal*, 59(2): 199-208.
26. Farrell, T. 2015. Promoting teacher reflection in second language education: A framework for TESOL professionals. New York: Routledge.
27. Furniss, E.A. 2016. Teaching the pragmatics of Russian conversation using a corpus-referred website. *Language Learning and Technology*, 20: 38-60.
28. Granger, S. 2009. The contribution of learner corpora to second language acquisition and foreign language teaching: a critical evaluation. *Corpora and Language Teaching*, 33: 13-32.
29. Granger, S. and Gilquin, G. 2011. From EFL to ESL: evidence from the international corpus of learner English. *Exploring second language varieties of English and learner Englishes: bridging a paradigm gap*. Amsterdam: John Benjamins, 55-78.
30. Granger, S. and Meunier, F. 1994. Towards a grammar checker for learners of English. *Creating and using English language corpora*. Amsterdam: Rodopi, 79-91.
31. Granger, S. and Paquot, M. 2009. Lexical verbs in academic discourse: a corpus-driven study of learner use. *Academic writing: At the interface of corpus and discourse*. London: Continuum, 193-214.
32. Granger, S. and Paquot, M. 2011. Language for specific purposes learner corpora. In: U. Connor and C.A. Chapelle (Eds.), *Language for specific purposes. The encyclopedia of applied linguistics*, Oxford: Blackwell-Wiley, 1-5.
33. Granger, S. and Vander, V. 2007. Corpus linguistics, language learning and ELT. *Mindbite*, 1: 11-14.
34. Granger, S., Paquot, M., and Rayson, P. 2006. Extraction of multi-word units from EFL and native English corpora. *The phraseology of the verb "make."* *Phraseology in Motion I: Methoden und Kritik. Akten der Internationalen Tagung zur Phraseologie*. Basel: Baltmannsweiler Schneider Verlag Hohengehren, 57-68.
35. Hogg, R.V. and Tanis, E.A. 1997. Probability and statistical inference. Upper Saddle River, NJ: Prentice Hall.
36. Holding, D. 1983. Stress and fatigue in human performance. New York: John Wiley & Sons.
37. Hosking, J.R. 1992. Moments or L moments? An example comparing two measures of distributional shape. *The American Statistician*, 46(3): 186-189.
38. Huang, S. 2018. EFL learners' critical multimodal reflections on the politics of English. *TESOL Journal*, 10(3): e00435.

39. Hudson T.D., Detmer E. and Brown J.D. 1995. Developing prototype measures of cross-cultural pragmatics (Technical Report 7). Honolulu, HI: University of Hawai'i, Second Language Teaching and Curriculum Center.
40. Hudson, T., Brown, J.D. and Detmer, E. 1992. A framework for testing crosscultural pragmatics. Honolulu, HI: Second Language Teaching and Curriculum Center, University of Hawai'i at Manoa.
41. Hyland, K. 2002. Teaching and researching writing. Pearson Education Ltd.
42. Ifantidou, E. 2011a. Genres and pragmatic competence. *Journal of Pragmatics*, 43(1): 327-346.
43. Ifantidou, E. 2011b. Genres and pragmatic understanding. Athens: Patakis Publications.
44. Ifantidou, E. 2013a. Pragmatic awareness: An index of linguistic competence. In *Research Trends in Intercultural Pragmatics*, ed. by Istvan Kecskes, and Jesús Romero-Trillo. Berlin/New York: Mouton de Gruyter, 105-144.
45. Ifantidou, E. 2013b. Pragmatic competence and explicit instruction. In *Pragmatic Development in L1, L2, L3: Its Biological and Cultural Foundations*, ed. by Elly Ifantidou, and Tomoko Matsui. *Journal of Pragmatics Special*, 59: 93-116.
46. Ifantidou, E. 2014. Pragmatic competence and relevance. Amsterdam: John Benjamins Publishing Company.
47. Ishihara, N. and Cohen, A.D. 2010. Teaching and learning pragmatics: Where language and culture meet. Harlow, UK: Pearson Longman.
48. Itomitsu, M. 2009. Developing a test of pragmatics of Japanese as a foreign language. Unpublished doctoral dissertation, Ohio State University, Columbus, OH.
49. Johansson, S. 1998. The approach of the text encoding initiative to the encoding of spoken discourse. *Spoken English on Computer*. Harlow: Longman, 82-98.
50. Johnson, N. and De Haan, J. 2013. Strategic interaction 2.0: Instructed intercultural pragmatics in and EFL context. *International Journal of Strategies Information Technology and Application*, 4: 49-62.
51. Kasper, G. 1997. Can pragmatic competence be taught? Honolulu: University of Hawaii, Second Language Teaching & Curriculum Center. Retrieved from the World Wide Web: <http://www.nflrc.hawaii.edu/NetWorks/NW06/>
52. Kasper, G. 2007. Pragmatics in second language learning: An update language learning roundtable, American Association for Applied Linguistics Annual Conference.
53. Kasper, G. and Rose, K. 2002. Pragmatic development in a second language. Oxford, UK: Blackwell.
54. Krashen, S.D. 1987. Principles and practice in second language acquisition. Prentice-Hall International.
55. Liu, J. 2010. Testing interlanguage pragmatic knowledge. *Pragmatics Across Languages and Cultures*, 7: 467.
56. Muller, S. 2005. Discourse markers in native and non-native English discourse. Amsterdam: John Benjamins Publishing Company.
57. Murray, N. 2010. Pragmatics, awareness raising, and the cooperative principle. *ELT Journal*, 64(3): 293-30.
58. Nation, P. 2001. The role of the first language in foreign language learning. *Asian EFL Journals*, 32(3): 177-175
59. Nemati, M. and Taghizade, M. 2006. Exploring similarities and differences between L1 and L2. *International Research Journal of Applied and Basic Sciences*, 4(9): 2477-2483.
60. Newmark, L. 1966. How not to interfere with language learning? *Language learning: The individual and the process*. *International Journal of American Linguistics*, 40: 77-83.
61. Newmark, L. 1971. A minimal language teaching program. In: P. Pimsleur and T. Quinn (Eds.), *The psychology of second language learning*. Cambridge: University Press.

62. Park, K. 2014. Corpora and language assessment: the state of the art. *Language Assessment Quarterly*, 11(1): 27-44.
63. Rabbanifar, A. and Mall-Amiri. 2017. The comparative effect of opinion gap and reasoning gap tasks on complexity, fluency and accuracy of EFL learners' speaking. *International Journal of Language Learning and Applied Linguistics World*, 16(4): 55-77.
64. Roever, C. 2005. *Testing ESL pragmatics: development and validation of a web-based assessment battery*. Berlin: Peter Lang.
65. Roever, C. 2006. Validation of a web-based test of ESL pragmalinguistics. *Language Testing*, 23: 229-256.
66. Roever, C. 2011. Testing of second language pragmatics: Past and future. *Language Testing*, 28: 463-481.
67. Römer, U. 2009a. Corpus research and practice: What help do teachers need and what can we offer? In: Aijmer, K. (Ed.), *Corpora and language teaching*. Amsterdam, The Netherlands: John Benjamins, 83-98.
68. Römer, U. 2009b. The inseparability of lexis and grammar: Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics*, 7: 140-162
69. Romero-Trillo, J. 2002. A mathematical model for the analysis of variation in discourse. *Journal of Linguistics*, 37: 527-550.
70. Romero-Trillo, J. 2017. Editorial. *Corpus Pragmatics*, 1: 1-2.
71. Rose, K. 2000. An exploratory cross-sectional study of interlanguage pragmatic development. *Studies in Second Language Acquisition*, 22: 27-67.
72. Rylander, J., Clark, P. and Derrah, R. 2013. A video-based method of assessing pragmatic awareness. *Assessing Second Language Pragmatics*, 65-97.
73. Schauer, G. and Adolphs, S. 2006. Expressions of gratitude in corpus and DCT data: Vocabulary, formulaic sequences, and pedagogy. *System*, 34: 119-134.
74. Schmidt, R. 1993a. Foundations of problem-based learning: some explanatory notes. *Medical Education*, 27(5): 422-432.
75. Schmidt, R. 1993b. Consciousness, learning and interlanguage pragmatics. In: G. Kasper and S. Blum-Kulka (Eds.), *Interlanguage pragmatics*. Oxford: Oxford University Press, 21-42.
76. Scott, M. 2004. *WordSmith tools*, 4. Oxford: Oxford University Press.
77. Selinker, L. 1972. *Interlanguage*, IRAL, 10. Mahwah, N.J: Lawrence Erlbaum.
78. Swain, M. 1985. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In: S. Gass, and C. Madden (Eds.), *Input in second language acquisition*. Rowley, MA: Newbury House, 235-253
79. Swain, M. 1993. The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 50: 158-165.
80. Taguchi, N. 2015. Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *Language Teaching*, 48: 1-50.
81. Taguchi, N. and Roever, C. 2017. *Second Language Pragmatics*. Oxford: Oxford Applied Linguistics.
82. Timpe, V. 2013. *Assessing intercultural language learning*. Frankfurt am Main: Lang.
83. Tognini-Bonelli, E. 2002. Corpus Linguistics at work. *Computational Linguistics*, 28: 583-583.
84. Vyatkina, N. 2016a. Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, 28(3): 207-226.
85. Vyatkina, N. 2016b. Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning and Technology*, 20(3): 159-179.
86. Vyatkina, N. and Boulton, A. 2017. Corpora in language learning and teaching. *Language Learning and Technology*, 21(3): 1-8.



87. Yamashita, S.O. 1996. Six measures of JSL pragmatics. Honolulu, HI: Second Language Teaching & Curriculum Center of University of Hawaii at Manoa.
88. Zwillinger, D. (Ed.). 1995. CRC Standard mathematical tables and formulae. Boca Raton, FL: CRC Press, 1995.

**Citation:** Aglaia Rouki. 2023. Unveiling Pragmatic Proficiency: Harnessing Corpus Pragmatics for Effective Implicature Assessment in High School Language Learning. *International Journal of Recent Innovations in Academic Research*, 7(10): 36-52.

**Copyright:** ©2023 Aglaia Rouki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.