## Table of Contents

*Language Value* November 2015, Volume 7, Number 1, pp. i-iv
http://www.e-revistes.uji.es/languagevalue ISSN 1989-7103

## From the Editors

# RESEARCH ON NEW METHODS FOR TESTING LANGUAGE SKILLS: A FOCUS ON SPEAKING PERFORMANCE

The current issue has speaking performance as a conducting thread, specifically, the research undertaken on new methods to test it. Testing language skills and defining the constructs under which tests need to be constructed is a complex task. Particularly, assessing speaking performance is a rather intricate process with many agents and variables involved and many factors affecting the final outcome. Subjectivity, fairness, anxiety, cognitive and metacognitive aspects, particular features of speech, individual abilities, and social constraints among others make the mastery of speaking in a foreign language and its subsequent assessment a definitely arduous task. Critical reflection and research on the part of educational institutions, researchers, test developers or any person undertaking assessment is of paramount importance to guarantee the adequacy, reliability and success of the whole testing process and results.

In the first article of the issue, entitled "New and not so new methods for assessing oral communication", **Gary J. Ockey and Zhi Li** note the necessary and logical evolution experienced by oral communication assessment practices over the past decades, namely those related to the testing process, the construct to be measured, the tasks employed and the technology used with this aim in the process. Authors delve into the broadening of the construct to be assessed, including interactional competence and technology and thus playing a determining role in the type of tasks currently in vogue and present. Five proposed task types in order to assess the degree to which they can contribute to effectively measure such construct are presented. These include: oral proficiency interviews, paired/group oral discussion tasks, simulated tasks, integrated oral communication tasks, and elicited imitation tasks. They are evaluated based on current conceptualizations of the construct of oral communication, and results seem to indicate


DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.1

that they do not assess a broad construct of oral communication equally. Consequently, authors advise test developers to use more than a single task type to more effectively ensure construct representativeness, considering the aspects of oral communication that they aim to include or exclude in their assessment when they select one of these task types.

In their "Comparing candidates' beliefs and exam performance in speaking tests", **Pérez-Guillot** and **Zabala-Delgado** analyse students' beliefs about their performance in the speaking section of a language proficiency exam and compare them with their actual results in the exam. In this way, the authors intend to determine whether students' beliefs were based on their actual level of competence or if they were based on other factors arising from the particular characteristics of this section of the exam, mainly anxiety or stress. The paper suggests that determining the basis for students' beliefs - either self-perceived or factual- and thus signaling the aspects to be modified would allow us to improve the reliability and quality of the exam. The authors claim that when developing a language exam factors outside the content of the exam, related to administration and organisation, as well as those connected to candidate's individual features should be considered in order to shed some light on the differences detected between perceptions and actual results.

**Beltrán-Palanques,** explores two different elicitation techniques among those that may be employed to test pragmatic competence, more specifically, discourse completion tasks/tests (DCTs) and role-play tasks (RPTs). As also reviewed by **Gary J. Ockey and Zhi Li** (this volume), RPTs and DCTs might be regarded as simulated tasks that can be used to assess pragmatic competence. In his study entitled "Revisiting pragmatic tests in the FL context: Towards interactive tests to examine speech act performance", Beltrán-Palanques examines the task effect of the two aforementioned elicitation techniques, which have been designed following an interactive perspective. More specifically, the DCTs, traditionally designed to allow participants to take only one turn, have been elaborated interactively to allow participants to freely interact in the written mode, thereby resulting in interactive DCTs (IDCTs). The RPTs used in his study were open,


DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.1

which also allow interaction between participants. The speech act chosen for the purposes of this study was the illocutionary act of apologies, as an example of an interactive communicative act. His study points out effects across the two elicitation techniques in relation to the speech act outcomes as regards length, amount and typology of apology strategies performed.

The volume also includes a final article entitled "Speaking an additional language: Can study abroad do the trick?" where **María Juan-Garau**, provides an overall picture of the acquisition of speaking abilities in a given second language acquisition (SLA) learning context. Specifically, she tries to ascertain whether studying abroad can indeed 'do the trick' when it comes to effectively and "quickly" learning a second or foreign language, and, if so, under which conditions. The author deals with the features of the study abroad (SA) context, analysing the way it may affect oral performance and thus the acquisition of oral competence. She proceeds focusing on the specific linguistic benefits in the oral domain that can accrue in this learning context. With this aim, an overview of empirical research projects findings is provided. At the same time that an array of individual and external variables that may impinge on successful language acquisition abroad are discussed to show that the SA context is definitely advantageous for the development of language learners' speaking skills as long as the relevant conditions outlined in the study are met.

Two books reviews have also been included in this issue. The first one, authored by **Elena Martín Monje**, reviews the volume entitled *Changing Methodologies in TESOL*, whose author, Jane Spiro, presents a work aimed at the student, student teacher and practising teacher of TESOL who might be interested in the teaching and learning process of English language nowadays as well as in the way it is experienced worldwide. The reviewer describes the work as "a valuable resource for pre-service TESOL courses" which may also be used as a reference book for teacher trainers and individual teachers to be informed of the latest developments in the field.


DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.1

Finally, **Raquel Lázaro Gutiérrez** reviews the volume *An Introduction to Interaction: Understanding Talk in Formal and Informal Settings*, authored by Ángela Cora García. The reviewer presents this volume as "a comprehensive guide for (mainly, but not only) linguistics, sociology, communication and even business students on the theories and research methodologies of conversation analysis", and concedes that when students reach its last chapter they are expected to be able to apply the basics of conversation analytical research to the transcripts provided or to any conversation.

***Mª Carmen Campoy-Cubillo***
***Nuria Edo-Marzá***

**Editors**
**Universitat Jaume I, Spain**


DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.1

*Language Value* November 2015, Volume 7, Number 1 pp. 1-21
http://www.e-revistes.uji.es/languagevalue ISSN 1989-7103

# New and not so new methods for assessing oral communication

**Gary J. Ockey**
gockey@iastate.edu
**Zhi Li**
zlisu2010@gmail.com
**Iowa State University, USA**

**ABSTRACT**

The assessment of oral communication has continued to evolve over the past few decades. The construct being assessed has broadened to include interactional competence, and technology has played a role in the types of tasks that are currently popular. In this paper, we discuss the factors that affect the process of oral communication assessment, current conceptualizations of the construct to be assessed, and five tasks that are used to assess this construct. These tasks include oral proficiency interviews, paired/group oral discussion tasks, simulated tasks, integrated oral communication tasks, and elicited imitation tasks. We evaluate these tasks based on current conceptualizations of the construct of oral communication, and conclude that they do not assess a broad construct of oral communication equally. Based on our evaluation, we advise test developers to consider the aspects of oral communication that they aim to include or exclude in their assessment when they select one of these task types.

**Keywords**: *oral communication, speaking, assessment, methodology*

## Introduction

Practice and research in assessing oral communication is regarded as the "youngest sub-field in language testing" (Fulcher 2003, p. 13). The testing process, the construct to be measured, the tasks used to measure the construct, and the technology used to aid in the process all continue to evolve as the field matures. These developments have helped to minimize some of the challenges that are faced in the assessment of oral communication. In this paper, we discuss the current state of the assessment of second language oral communication in the light of some of these developments. We begin by briefly outlining the oral communication assessment process. Our aim with this section is to provide an indication of some of the factors to be considered when assessing oral communication. Next, we provide an oral communication construct, which is in line with current conceptions in the field. The greater part of our paper is provided in the next section, which describes the tasks that are currently being used to aid in assessing



DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.2

oral communication. Along with the description of these tasks, we analyze them based on the degree to which they can be used to effectively measure our construct of oral communication, given the factors presented in an oral communication assessment process.

## I. THE ORAL COMMUNICATION ASSESSMENT PROCESS

A number of factors contribute to the score that a test taker receives on a test designed to assess his or her ability to communicate orally. Figure 1 provides a graphic display of how some of these factors affect scores.

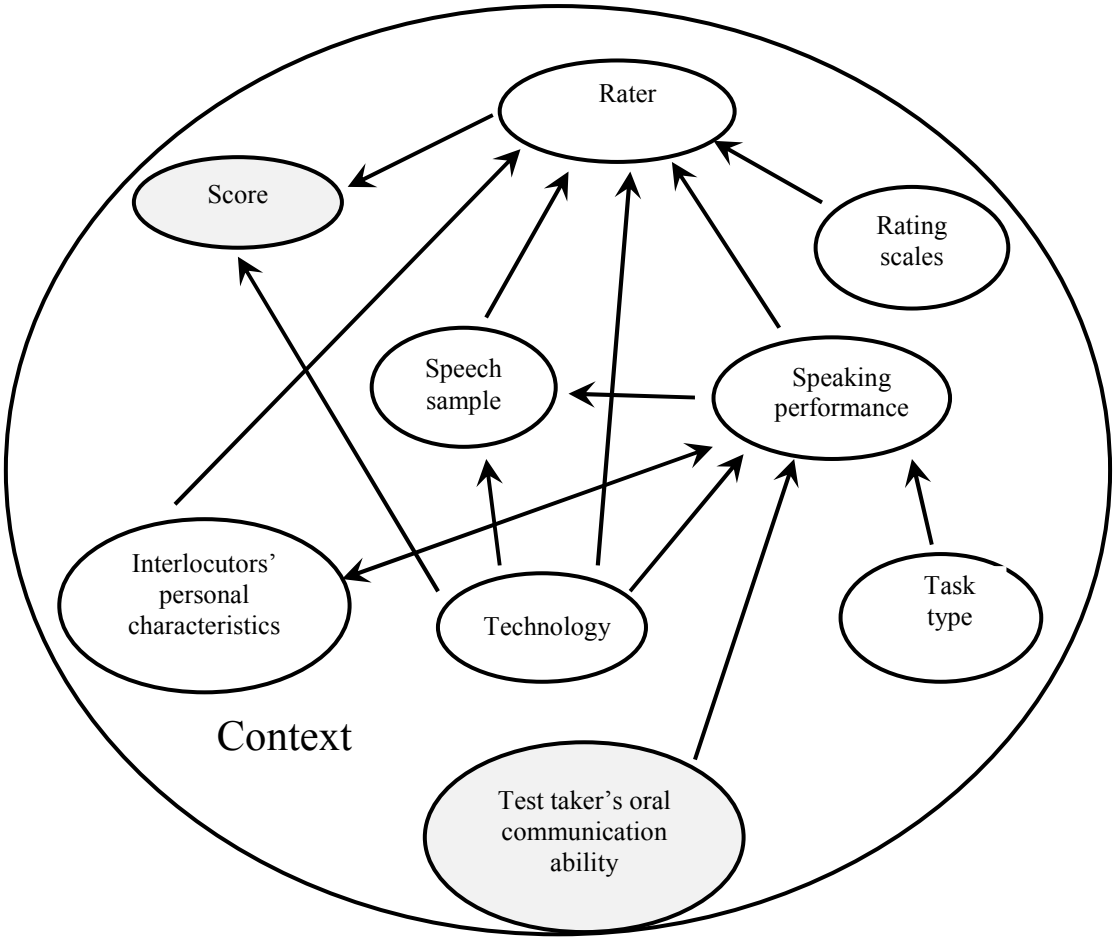


Figure 1 Model of assessment of oral communication

The conceptualization builds on the earlier models of Kenyon (1992), McNamara (1996), Skehan (1998), Bachman (2001), and Ockey (2009). The model focuses on factors that have an impact during the administration of the test. Other factors, such as the impact of the score on instruction, are not explicitly identified in the model, but are considered to be part of the testing context. In the figure, the test taker's oral communication ability is depicted by an oval at the bottom of the large circle. The aim of the assessment is to measure this ability. The score that is assigned to the test taker based on the assessment is shown in the upper left part of the large circle. This score is used to indicate the test taker's oral communication ability. As can be seen in the figure, task type, other interlocutors' personal characteristics, technology used for the assessment, the actual speaking performance and resulting speech sample, rating scales, and raters can all have an impact on scores during a test administration. These factors, coupled with the context (e.g., stakes, consequences, sociopolitical situation, and cultural expectations of stakeholders) of the assessment, may all be sources of construct-irrelevant variance in an oral communication assessment.

A number of factors have influence on test scores. Task types (as well as the specific prompts used for a particular task type) can affect a test taker's speaking performance. Familiarity with a task may be an advantage for test takers. For instance, some test takers may do better with a group or paired discussion because they are used to talking to others in a group setting. The personal characteristics of the other interlocutors involved in the assessment can influence speaking performance and can have an effect on how raters evaluate a test taker's ability. For example, in a one-on-one interview, an interviewer can affect the scores by being more or less supportive during the interview, and in a group/paired oral test, the members of the test taker's group might be very assertive, thus having an impact on the test taker's speaking performance. The rater may also compare the performance of the group members in a group/paired discussion, thereby making it possible for the abilities of the other members of one's group to have a direct effect on a test taker's score through the rater. Technology can also affect a test taker's speaking performance. For instance, a test delivered over the telephone or the internet may be interrupted by a slow or unclear connection, making it difficult for the test taker to understand the prompt. Technology can also affect the speech sample if, for example, the recording device does not function effectively. Technology might also

influence scores through the rater, who, for instance, may not be able to use the data entry procedures effectively. Finally, technology can affect a test score directly if the scoring system does not work properly.

Rating scales can also have an influence on scores. Since they are designed to measure the construct that the tester aims to assess, they therefore play a crucial role in linking the scores to the construct that the test is designed to measure. To be effective, rating scales must clearly reflect the construct and be easily interpretable. Raters can be human, computer (automated scoring), or a combination of both. Human raters and/or computer automated scoring engines play a key role in the oral communication assessment process, and can affect scores in several ways, depending on how they are trained or programmed to interpret the rating scales and evaluate elicited speech samples. All of these factors work in a given assessment context, which also influences the scores assigned to test takers. There are many contextual factors, including physical features of the setting, such as the temperature of the room and level of external noise, and psychological factors, such as the test taker's anxiety and motivation. Given all of these factors in an oral communication assessment process, it is crucial that test designers define the oral communication ability that they aim to assess as clearly as possible, and then consider each of these factors to best ensure a valid assessment process.

## II. THE CONSTRUCT OF ORAL COMMUNICATION

The definition of the construct spells out the key components or essential aspects of the ability test developers wish to measure. In the context of assessing speaking, Fulcher (2003: 23) defined speaking ability as "the verbal use of language to communicate with others". Fulcher's (2003) definition is in line with other more recent definitions, such as that of Jamieson, Eignor, Grabe, and Kunnan (2008: 74), who defined speaking ability as "the use of oral language to interact directly and immediately with others… with the purpose of engaging in, acquiring, transmitting, and demonstrating knowledge". These broad definitions of oral proficiency suggest that this ability includes: 1) interactional competence; 2) appropriate use of phonology; 3) appropriate and accurate use of vocabulary and grammar; and 4) appropriate fluency.

Interactional competence can be viewed as an individual's underlying ability to actively structure appropriate speech in response to incoming stimuli, such as information from another speaker, in real time. That is, interactional competence can be considered as the individual attributes that test takers need to engage in real-time interactive communication, which may not be necessary to engage in non-interactive communication. More specifically, interactional competence entails the ability to take turns, open and close gambits, respond to others, and negotiate and develop topics with appropriate pragmatic use in a given context. Research suggests that interactional competence is not adequately assessed with description or prepared oral presentation tasks (Ockey, Koyama, Setoguchi, and Sun 2015). Other research indicates that test takers prefer real-time tasks in which they actively co-construct meaning with other interlocutors. They also feel that such tasks are better indicators of their oral communication ability in the second language (Brooks and Swain 2015).

Appropriate use of phonology relates to the effective use of both segmental and prosodic aspects of language. At the segmental level, pronunciation refers to the ability to articulate words and create the physical sounds that endow a word with a meaning. Prosodic aspects of phonology include stress, increased volume on a syllable, intonation, voice movement, and pitch (Fulcher 2003). A major conundrum in assessing second language oral communication relates to how to assess accent, an important aspect of phonology. Strength of accent has been defined as, "the degree to which (the accent) is judged to be different than the local variety, and how it is perceived to impact the comprehension of users of the local variety." This definition emerged partly as a result of research which has shown that high-stakes assessments that are rated by local raters who are familiar with the speakers' first language can assign much more lenient ratings than raters who are not familiar with the local first language (Carey, Mannel and Dunn 2011). While some argue for more acceptance of various accents when assessing oral communication (Abeywickrama 2013; Smith and Bisazza 1982), others note the importance of accent in oral communication, and argue that to be fair to test takers, oral communication assessments should carefully consider the accent of the input (Elder and Harding 2008; Ockey and French 2014) and judge the strength of the test takers' accent as a part of their oral communication ability.

Appropriate and accurate use of vocabulary and grammar refer to vocabulary breadth, how many words are known; vocabulary depth, how well and effectively the words are known and can be used (Nation 1990); grammatical breadth, how many grammar structures are known and can be used; and grammatical depth, how accurately and effectively these grammatical structures can be used. Grammar and vocabulary have been treated as separate constructs, but research suggests that human raters do not assign distinct scores for vocabulary and grammar in oral communication assessments (Batty 2006; Hunston, Francis and Manning 1997). Given the strong relationship between scores on vocabulary and grammar, it can be argued that they should be treated as one sub-ability of oral communication.

As one of the four components in the construct of oral communication, fluency, which refers to naturalness of rate of speech, pausing, and repetition has attracted a lot of attention. The temporal aspects of fluency include various measures of quantity, rate, pausing, and language repair (Bosker, Pinget, Guene, Sanders and de Jong 2012; Ginther, Dimova and Yang 2010). Research indicates that the temporal measures of fluency are important components of oral communication. Sato (2014) labeled fluency within interactions as interactional oral fluency and argued that fluency is a 'perceived phenomenon', in which temporal aspects of fluency are interwoven with interactional features.

Having laid out the process of assessing oral communication and the factors that affect it, along with the construct of this ability, we now turn to the types of task that have been used to assess oral communication. We describe each task and provide an analysis of the extent to which it assesses the construct of oral communication that we have provided, given the process and accompanying factors of assessing this ability.

## III. TASK FORMATS IN TESTING ORAL COMMUNICATION SKILLS

Many of the types of tasks that are currently popular in testing oral communication skills are described by Harris (1969), who introduced three main types of "oral production tests" used in the 1960s, namely: 1) scored interviews, 2) highly structured speech samples, and 3) paper-and-pencil tests of pronunciation. The first type of task requires one or more trained interviewers/assessors to engage in conversations with test

takers and rate their performance based on established scales. Scored interviews and their variants are still widely used as one of the dominant tasks in assessing oral communication ability. The second type of tasks relies on pre-set stimuli and does not involve interlocutors. Typical examples in this category include sentence repetition and reading a passage aloud. The tasks in the family of highly structured speech samples have gone through ups and downs in the past decades but have seen a certain degree of revival in recent years, thanks at least in part, to the emergence of automated speech rating systems. The last type of tasks described in Harris (1969) requires written responses about finding rhyme words, identifying word stress and phrase stress. This type of paper-and-pencil tests of oral production has mostly disappeared from language testing, probably as a result of the popularity of the communicative language teaching paradigm.

We now discuss five popular types of task used to assess oral communication. The task types are oral proficiency interviews, paired/group oral discussion tasks, simulated tasks, integrated oral communication tasks, and elicited imitation tasks. The first three task types are variants of scored interviews, while the last two are variants of the highly structured samples discussed in Harris (1969).

### III.1 Oral proficiency interviews

Oral proficiency interviews are one of the most commonly used task formats for assessing oral communication. A typical oral proficiency interview task requires a test taker to respond to questions on different topics posed by an interviewer, who usually chooses the topics, initiates the conversations, and sometimes rates the speech samples elicited from the test taker in the test. In these tasks, test takers are expected to respond to questions but they usually have limited opportunities to demonstrate their ability to negotiate meaning, open and close gambits, or elicit opinions from the interviewer.

One example of this task format is the American Council on the Teaching of Foreign Languages (ACTFL) Oral Proficiency Interview (OPI), which was developed in the 1980s and has been widely used as a "standardized procedure for the global assessment of functional speaking ability" in a number of foreign languages (ACTFL 2012; Liskin-Gasparro 2003). The task format used in the ACTFL OPI is a face-to-face or telephonic

interaction between a test taker and a certified interviewer on a series of personalized questions. In the OPI tests, an interviewer initiates and leads the conversations by choosing topics of a variety of natures, including personal, general, and abstract. In this testing context, the interviewer, who is not familiar with the test taker, controls the topics, asks the questions, and generally directs the course of the conversation. The ACTFL OPI test includes four mandatory phases, namely, warm-up, level checks, probes, and wind-down. In the warm-up step, the interviewer asks the test taker simple questions and establishes rapport. At the level checks step, the interviewer asks the test taker a number of questions with the aim of deciding the proficiency floor, or the proficiency level that the test taker can successfully demonstrate. In the probes step, the interviewer asks questions with a higher level of proficiency than the level expected by the test taker in order to determine the proficiency ceiling, or the highest possible proficiency level that the test taker can sustain.

There are two technology-mediated variants of the ACTFL OPI: telephonic OPI and the Internet-delivered version of OPI or OPIc. In the telephonic OPI, test takers call the testing center and take the test via phone, rather than face-to-face as in regular OPI test. Instead of involving a human interlocutor, the OPIc uses an avatar, through which the one-on-one interview model is simulated.

The speech samples elicited from the real time (face-to-face and telephonic) OPI-styled tasks make it possible to assess a number of aspects of oral communication, such as global tasks and functions, context and content, grammatical accuracy, and text type. Specifically, the ACTFL rating rubric contains detailed descriptions regarding test taker's performance in terms of fluency, pronunciation, grammar, and vocabulary. These aspects of oral communication closely resemble three of the four aspects of oral communication discussed in the construct section of the paper. However, OPI tasks are likely to fall short in assessing many aspects of interactional competence (Johnson and Tyler 1998; van Lier 1989). One of the major concerns is that the speech samples elicited from the OPI tasks do not exhibit key features observed in natural conversations, such as “reactive and mutual contingency” (van Lier 1989: 501), which refers to the spontaneous and interactive sequence of speech between two speakers. This may be related to the unequal power relationship between test takers and the interviewer (Johnson and Tyler 1998). The interviewer asks the questions, while the test takers can

only respond to what is asked. In addition, the interviewer's behaviors and degree of involvement can contribute to the "asymmetrical and pseudosocial" nature of the OPI interactions. That is, the discourse that the test taker produces can be affected by the personal characteristics of the interviewer. For example, the interviewer may be much less friendly than other interviewers, which can lead to an unrepresentative sample of discourse from the test taker (Brown 2003). The technology-mediated variants of the OPI-type tasks have similar limitations, but they probably assess even less of the interactional competence aspect of the oral communication construct, given that test takers cannot ask for clarification of a question or interact with the interviewer at all.

### III.2 Paired and group oral discussions

Paired and group oral discussions can address some of the limitations of the OPI-type tasks. In this format, pairs or small groups of students have a discussion with each other. A trained interlocutor acts as a moderator and may or may not participate in certain aspects of the task. Test takers can be paired or grouped as equal status speakers based on different criteria, for example, proficiency level or interpersonal relationship. One example task is the group discussion in the College English Test – Spoken English Test (CET-SET), which uses a computer program to group three to four test takers and requires them to sustain a 4.5-minute face-to-face discussion on a given topic (He and Dai 2006).

The potential of group oral tasks in assessing oral communication was recognized in the 1980s and the last few decades have witnessed more implementations of this type of tasks in both high-stakes contexts such as the Cambridge Main Suite Examinations in the UK, including the First Certificate in English (FCE) and the Cambridge Certificate in Advanced English (CAE), CET-SET in China, and the speaking section of a provincial exit exam in Canada (Turner 2009), as well as various local English placement tests, such as the Kanda Assessment of Communicative English in Japan (**Ockey, Koyama, Setoguchi and Sun** 2015) and a placement test at Michigan State University in the USA (Winke 2013).

Currently, paired and group oral discussion tasks are mainly carried out in a face-to-face manner, thus requiring the physical presence of each participating test taker. However,

such tasks could be completed via synchronous voice-based computer-mediated communication (CMC), as has been done to aid in English teaching (Alastuey 2011; Lin 2014). For example, video-conference techniques, such as Adobe Connect and Skype, could be used as a testing platform to connect test takers who are not in a face-to-face context. In addition, computer technology could be used to group test takers based on pre-established criteria such as English proficiency level, personality traits, and topic familiarity.

Since paired and group oral tasks are designed to elicit interaction among test takers, accordingly, the rating rubric for paired and group oral tasks generally includes the sub-construct of interactional competence and can therefore accommodate a broader coverage of the oral communication construct than OPI-type tasks. For example, in the group oral placement test described in Bonk and Ockey (2003), test takers' performances are rated on pronunciation, fluency, grammar, vocabulary/content, and communicative skills/strategies, the latter being essentially another name for interactional competence. In the group discussion task of the CET-SET, the evaluative criteria include: 1) accuracy in pronunciation, stress/intonation, and use of grammar and vocabulary, 2) range of vocabulary and grammatical structures, 3) size (percentage) of contribution to group discussion, 4) discourse management, 5) flexibility in dealing with different situations and topics, and 6) appropriateness in the use of linguistic resources (Zheng and Cheng, 2008). These aspects of oral communication fit quite closely with the four aspects of oral communication described in the construct section above, thus suggesting that group and paired tasks aim to assess all four components of the construct. Empirically, it has been found that peer-to-peer discussion provides test takers with a better opportunity to demonstrate their ability to engage in complex interaction, compared with test taker-to-interviewer interaction, as is the case with OPIs (Brooks 2009). In this sense, paired and group oral tasks can tap into a fuller range of oral communication abilities than OPI-type tasks.

Given the complex interaction patterns exhibited in paired and group oral tasks, this task type has attracted much attention. With regard to the effects of interlocutor traits, research suggests that the test takers' familiarity with other test takers (O'Sullivan, 2002), as well as personality (level of extraversion), English proficiency level, and the number of participants, may influence test takers' performance in group discussion

tasks, as shown in a study on Japanese secondary school students conducted by Nakatsuhara (2011). On the other hand, Ockey, Koyama, and Setoguchi (2013) investigated the effect of interlocutor familiarity on test takers' performance in a Japanese university. A comparison of the scores of the two groups of test takers, namely a class-familiar group and a class-unfamiliar group, showed that interlocutor familiarity did not exert a significant influence on four rating categories (pronunciation, fluency, lexis and grammar, and communication skills), suggesting that at least for some contexts, interlocutor familiarity may not have a significant impact on scores elicited from the group oral. The effects of prompts in oral discussion tasks are reported by Leaper and Riazi (2014), who compared the turn-taking features, syntactical complexity features, accuracy, and fluency in the test taker's discourse elicited with four prompts. It was found that the prompts that allowed for an account or extension of personal experiences tended to elicit longer and more complex turns, whereas the prompts with factual content yielded shorter and less complex turns.

To sum up, the group and paired oral tasks have the advantage of providing test takers with the opportunity to demonstrate their interactional competence. This opportunity seems to stem from the rather loose controls placed on the task. That is, test takers seem to be able to demonstrate their interactional competence because the task affords them a fair number of opportunities to collaborate with others of equal status. On the other hand, because of this loose control, the task is susceptible to a number of factors, such as the personalities of other test takers with whom they are grouped, which can affect their test scores.

### III.3 Simulated tasks

Simulated tasks are commonly used to assess oral communication in the context of English for specific purposes (ESP). An example of this type of task is role-play tasks, which require a test taker to assume a particular role in a simulated task context, for example, a meeting with a professor during office hours. Another example of a simulated task is the teaching tasks used in assessing the oral communication ability of prospective international teaching assistants (ITAs) in English-speaking universities. The Taped Evaluation of Assistants' Classroom Handling (TEACH), originally

developed at Iowa State University in 1985, is a performance test with simulated tasks for ITAs (Papajohn 1999; Plakans and Abraham, 1990). The TEACH test consists of three phases including a 2-minute preparation, a 5-minute lecture, and a 3-minute question-answering activity. As introduced in Papajohn (1999), some undergraduates are invited to the testing room to form a 'mock class', and ask questions to the ITAs in the TEACH test. In these simulated tasks, test takers select topics in their own field as the teaching content and present the lecture to mock students as well as the assessors.

Another variant of simulated tasks attempts to assess pragmatic competence through computerized discourse completion tasks (DCT) with a video prompt (Sydorenko, Maynard and Guntly 2014). In this task, test takers are presented with video prompts which describe situations requiring them to make appropriate requests. The test takers respond to the prompts orally and then computer technology is used in an attempt to follow up with rejoinders. The aim is to produce multiple conversation turns. **Sydorenko, Maynard and Guntly** (2014) suggested that the computer-delivered DCT is superior to the traditional paper-based DCT in assessing pragmatic competence in that the former elicits simulated and extended discourse in a more authentic way.

The importance of simulated tasks can be more salient in occupation-related English language tests or English for special purposes (ESP) tests. One example is an oral communication test in aviation English for air traffic controllers developed by Park (2015). The test simulates a control tower as a virtual assessment environment in Second Life, an online 3D virtual world. In this role-play task, test takers act as air traffic controllers and give oral directives based on incoming aural information. While Park's tasks rely on input that has been recorded, that is, the task is asynchronous, it is feasible to enable multi-user voice communication in a virtual environment like Second Life. In that situation, test takers' interactional competence could also be elicited and assessed through technology-mediated communication.

As can be seen, there are numerous variants of simulated tasks. Some assess all four of the constructs of oral communication more effectively than others. Of particular note is that a major aim of these tasks is to assess interactional competence, but in some cases it is not clear to what extent they can actually be used to measure this ability.

### III.4 Integrated tasks

Integrated tasks aim to measure more than one subskill. Examples are listen-speak or read-speak tasks. Developers of these tasks recognize that oral communication rarely involves one-way speech, such as a monologic oral presentation with no question and answer session. These tasks normally include extended written or oral stimuli after which the test taker is expected to provide an extended response. We note that many of the task types that we have discussed require both speaking and listening (which is the major reason we use the term 'oral communication' as opposed to 'speaking' throughout the paper). Integrated tasks can be thought of as an extension of the task type of highly structured speech samples in Harris (1969)'s classification. In this paper, to avoid terminology confusion, we limit the term to the tasks that require test takers to produce speech samples based on given input materials without any synchronous interactions.

Integrated tasks have attracted a great deal of attention from researchers and test developers partly due to the influence of the TOEFL iBT which uses this type of test task to assess speaking ability. In an integrated oral communication task, test takers are required to either listen to a short audio clip or read a short passage, and then summarize the input for a hypothetical audience who does not have access to the same input. Since no interlocutor is needed in the testing process, integrated tasks can be computerized, as exemplified in the TOEFL iBT speaking test. In the integrated tasks of the TOEFL iBT speaking test, computers are used to deliver aural and textual input materials and to record a test taker's speech sample responses. These summary-type tasks have gained some popularity in recent years, in part because of their potential to be rated by automated scoring systems. An example of an automated scoring system is SpeechRater$^{SM}$, which is currently used to score the speaking section of the TOEFL Practice Online (TPO).

The speech samples elicited from integrated tasks, such as the read-listen-speak task used in the TOEFL iBT, can be rated for phonology, fluency, and grammar and vocabulary. However, this type of task does not directly measure interactional

competence. In addition, using aural or textual input in integrated tasks can complicate the test-taking performance and may make it difficult to determine what the task is measuring. For instance, in such tasks it is not clear to what extent reading comprehension is assessed, and how much working memory capacity affects a test taker's oral performance. The test taker's strategy use in integrated tasks may also be different from other task types (Barkaoui, Brooks, Swain, and Lapkin 2012). The questions about sub-constructs measured with integrated tasks could be more noteworthy when automated scoring tools are used. In the latest version of SpeechRater, the features used for scoring include speech articulation rate, average length of speech chunks, unique words normalized by speech duration, Acoustic Model scores, and Language Model scores. Considering the limitations in construct representation and model prediction accuracy, Xi, Higgins, Zechner, and Williamson (2012) only endorse applications of SpeechRater in low-stakes contexts.

### III.5 Elicited imitations

Elicited imitation tasks require a participant to listen to a sentence and then repeat the stimulus material (a word, phrase, or sentence) as closely as possible. This task type was commonly used decades ago but, probably because it is not in line with communicative language teaching principles, fell out of popular use until recently. The revival of elicited imitation tasks in language testing is likely attributable to the ease of delivering and scoring these tasks with automated speech scoring systems. An example of one of these systems is Duolingo's, which is an online language learning website and mobile app. The system uses elicited imitation tasks for its English Test (Ye 2014). These tasks types can be scored using automated speech scoring systems which extract multiple acoustic and prosodic features from test takers' speech samples (Bernstein 2013).

Elicited imitation tasks may provide good estimates of a test taker's fluency and pronunciation through use of automated speech recognition (ASR) technology, but they have limited potential for assessing vocabulary and grammar and little or no potential for assessing interactional competence. While elicited imitation tasks can be reliably scored, and with ASR technology are quite practical, they have been criticized for not having the potential for assessing a broad construct of oral communication (Chun 2006;

O'Sullivan 2013). The task formats and the expected responses in elicited imitation tasks do not involve any interactional aspects of real-life oral communication. Moreover, it can be argued that these task types may be poor indicators of phonology and fluency, since it may be possible for a test taker to simply imitate the phrases with no understanding or ability to segment the speech stream into meaningful parts. In short, imitation tasks, such as sentence repetition tasks, are generally believed to have little potential to assess a broad construct of oral communication. It should also be noted that these tasks could result in negative washback on instruction, since to prepare for such tasks, test takers may spend their time repeating sentences rather than using their time to engage in meaningful discussions with other language users.

## IV. CONCLUSION

Assessing oral communication is a rather complicated process, as shown in Figure 1 at the beginning of the paper. A review of the popularly used tasks for assessing oral communication suggests that a number of factors should be considered in determining which task types to include in a speaking test (see Table 1 for a summary of the oral communication testing tasks) for a particular context. Firstly, a clear construct definition should be elaborated. In other words, it is necessary to spell out what should be counted as oral communication in a particular context. We propose that at least four key aspects of oral communication should be assessed, namely, interactional competence, grammar/vocabulary, phonology, and fluency. Secondly, testing tasks and the corresponding scoring rubric should be reviewed with reference to the constructs. The task types listed in Table 1 have been briefly reviewed in this paper and summarized in Table 1. A check mark indicates that this task has good potential for assessing the ability, a question mark indicates that it has limited potential to assess the ability or potential to assess only certain aspects of the ability, and an X indicates that the task has little or no potential to assess the ability. Since each task type has its own merits and drawbacks, our general suggestion is that, after considering the tasks that might be most appropriate for a particular context based on the extent to which they assess all aspects of the construct and their feasibility, test developers should use more than one task type to best ensure construct representativeness.

Table 1. Summary of the characteristics of oral communication testing tasks

| **Tasks** | **Example of task format** | **Constructs measured** | |
|---|---|---|---|
| **Oral proficiency interviews** | face-to-face interviews or phone interview with an examiner in ACTFL OPI | Interactional competence | ? |
| | | Fluency | ☑ |
| | | Grammar & vocabulary | ☑ |
| | | Phonology | ☑ |
| **Paired or group oral discussions** | unstructured discussion among peers in CET-SET | Interactional competence | ☑ |
| | | Fluency | ☑ |
| | | Grammar & vocabulary | ☑ |
| | | Phonology | ☑ |
| **Simulated tasks** | mini-lecture presentation and question answering in the TEACH test for international teaching assistants | Interactional competence | ? |
| | | Fluency | ☑ |
| | | Grammar & vocabulary | ☑ |
| | | Phonology | ☑ |
| **Integrated tasks** | summarization after listening to or reading input materials in TOEFL iBT Speaking test | Interactional competence | ? |
| | | Fluency | ☑ |
| | | Grammar & vocabulary | ☑ |
| | | Phonology | ☑ |
| **Elicited imitations** | sentence repetition in Duolingo English Test | Interactional competence | ☒ |
| | | Fluency | ☑ |
| | | Grammar & vocabulary | ? |
| | | Phonology | ? |

## REFERENCES

**Abeywickrama, P.** 2013. "Why not non-native varieties of English as listening comprehension test input?" *RELC Journal 44*(1), 59-74.

**ACTFL.** 2012. *ACTFL Proficiency Guidelines 2012 – Speaking*. Retrieved from http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/english/speaking

**Alastuey, M. C. B.** 2011. "Perceived benefits and drawbacks of synchronous voice-based computer-mediated communication in the foreign language classroom". *Computer Assisted Language Learning*, *24*(5), 419–432.

**Bachman, L. F.** 2001. *Speaking as a realization of communicative competence*. Paper presented at the meeting of the American Association of Applied Linguistics – International Language Testing Association (AAAL-ILTA) Symposium. St. Louis, Missouri.

**Barkaoui, K., Brooks, L., Swain, M. and Lapkin, S.** 2012. "Test-takers' strategic behaviors in independent and integrated speaking tasks". *Applied Linguistics*, *34*(3), 304–324.

**Batty, A.** 2006. "An analysis of the relationships between vocabulary learning strategies, A Word Associates Test, and the KEPT". *Studies in Linguistics and Language Education: Research Institute of Language Studies and Language Education, 17*, 1-22.

**Bernstein, J. C.** 2013. "Computer scoring of spoken responses". In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell.

**Bonk, W. and Ockey, G. J.** 2003. "A many-facet Rasch analysis of the second language group oral discussion task". *Language Testing*, *20*(1), 89–110.

**Bosker, H. R., Pinget, A.-F., Quene, H., Sanders, T. and de Jong, N. H.** 2012. "What makes speech sound fluent? The contributions of pauses, speed and repairs". *Language Testing*, *30*(2), 159–175.

**Brooks, L.** 2009. "Interacting in pairs in a test of oral proficiency: Co-constructing a better performance". *Language Testing*, *26*(3), 341–366.

**Brooks, L. and Swain, M.** 2015. "Students' voices: The challenge of measuring speaking for academic contexts". In B. Spolsky, O. Inbar, and M. Tannenbaum (Eds.), *Challenges for language education and policy: Making space for people* (pp. 65–80). New York: Routledge.

**Brown, A.** 2003. "Interviewer variation and the co-construction of speaking proficiency". *Language Testing*, *20*(1), 1–25.

**Carey, M. D., Mannel, R. H. and Dunn, P. K.** 2011. "Does a rater's familiarity with a candidate's pronunciation affect the rating in oral proficiency interviews?" *Language Testing*, *28*(2), 201–219.

**Chun, C. W.** 2006. "Commentary: An analysis of a language test for employment: The authenticity of the PhonePass Test". *Language Assessment Quarterly*, *3*(3), 295–306.

**Elder, C.** and **L. Harding.** 2008. "Language testing and English and an international language: Constraints and contributions" in Sharifian, F. and M. Clyne (eds.): *Australian Review of Applied Linguistics (special forum issue) 31*(3), 34.1–34.11.

**Fulcher, G.** 2003. *Testing second language speaking*. London and New York: Longman.

**Ginther, A., Dimova, S. and Yang, R.** 2010. "Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring". *Language Testing*, *27*(3), 379–399.

**Harris, D. P.** 1969. *Testing English as a Second Language*. New York: McGraw-Hill Book Company.

**He, L. and Dai, Y.** 2006. "A corpus-based investigation into the validity of the CET-SET group discussion". *Language Testing*, *23*(3), 370–401.

**Hunston, S., Francis, G. and Manning, E.** 1997. "Grammar and vocabulary: Showing the connections". *ELT Journal, 51*(3), 208-216.

**Jamieson, J., Eignor, D. R., Grabe, W. and Kunnan, A. J.** 2008. "Frameworks for a new TOEFL". In C. A. Chapelle, J. Jamieson, and M. K. Enright (Eds.), *Building a validity argument for the Test of English as a Foreign Language* (pp. 55–95). New York: Routledge.

**Johnson, M. and Tyler, A.** 1998. "Re-analyzing the OPI: How much does it look like natural conversation?". In R. F. Young and A. W. He (Eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency* (pp. 27–52). Amsterdam and New York: Benjamins.

**Kenyon, D. M.** 1992. *Development and use of rating scales in language testing.* Paper presented at the annual meeting of the Language Testing Research Colloquium. Vancouver, Canada.

**Leaper, D. A. and Riazi, M.** 2014. "The influence of prompt on group oral tests". *Language Testing*, *31*(2), 177–204.

**Lin, H.** 2014. "Computer-mediated communication (CMC) in L2 oral proficiency development: A meta-analysis". *ReCALL*, 1–27.

**Liskin-Gasparro, J. E.** 2003. "The ACTFL Proficiency Guidelines and the Oral Proficiency Interview: A brief history and analysis of their survival". *Foreign Language Annals*, *36*(4), 483–490.

**McNamara, T.** 1996. *Measuring second language performance*. London: Longman.

**Nakatsuhara, F.** 2011. "Effects of test taker characteristics and the number of participants in group oral tests". *Language Testing*, *28*(4), 483–508.

**Nation, I.S.P.** 1990. *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle.

**O'Sullivan, B.** 2002. "Learner acquaintanceship and oral proficiency test pair-task performance". *Language Testing 19*(3)*,* 277-295.

**O'Sullivan, B.** 2013. "Assessing speaking". In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 156–171). Wiley.

**Ockey, G. J.** 2009. The effects of group members' personalities on a test taker's L2 group oral discussion test scores. *Language Testing*, *26*(2), 161–186.

**Ockey, G. J., & French, R.** 2014. From one to multiple accents on a test of L2 listening comprehension. *Applied Linguistics*. doi: 10.1093/applin/amu060.

**Ockey, G. J., Koyama, D., Setoguchi, E., and Sun, A.** 2015. The extent to which TOEFL iBT speaking scores are associated with performance on oral language tasks and oral ability components for Japanese university students. *Language Testing*, *32*(1), 39–62.

**Ockey, G. J., Koyama, D., & Setoguchi, E.** 2013. Stakeholder input and test design: A case study on changing the interlocutor familiarity facet of the group oral discussion test. *Language Assessment Quarterly, 10*(1), 1–17.

**Papajohn, D.** 1999. The effect of topic variation in performance testing: The case of the chemistry TEACH test for international teaching assistants. *Language Testing*, *16*(1), 52–81.

**Park, M.** 2015. *Task selection and simulation for authentic aviation English assessment*. Paper to be presented at the invited colloquium "Task-based language assessment in practice: Justification and realizations" at the 6th International Conference on Task-Based Language Teaching. Leuven, Belgium. September 16-18, 2015

**Plakans, B. S. and Abraham, R. G.** 1990. The testing and evaluation of international teaching assistants. In D. Douglas (Ed.), *English language testing in U.S. colleges and universities* (pp. 68–81). Washington, DC: NAFSA.

**Sato, M.** 2014. Exploring the construct of interactional oral fluency: Second Language Acquisition and Language Testing approaches. *System*, *45*, 79–91.

**Skehan, P.** (1998). Processing perspectives to second language development, instruction, performance, and assessment. *Thames Valley Working papers in Applied Linguistics, 4*, 70-88.

**Smith, L. and Bisazza, J.** 1982. The comprehensibility of three varieties of English for college students in seven countries. *Language Learning, 32(2),* 259-269. **Swender, E.** 2003. Oral Proficiency Testing in the Real World: Answers to Frequently Asked Questions. *Foreign Language Annals*, *36*(4), 520–526.

**Sydorenko, T., Maynard, C. and Guntly, E.** 2014. Rater behaviour when judging language learners' pragmatic appropriateness in extended discourse. *TESL Canada Journal*, *32*(1), 19–41.

**Turner, C. 2009. Examining washback in second language education contexts: A high stakes provincial exam and the teacher factor in classroom practice in Quebec secondary schools. *International Journal of Pedagogies and Learning, 5*, 103-123.**

**Van Lier, L.** 1989. Reeling, writhing, drawling, stretching, and fainting in coils: oral proficiency interviews as conversation. *TESOL Quarterly*, *23*, 489–508.

**Winke, P.** 2013. The effectiveness of interactive group orals for placement testing. In K. McDonough and A. Mackey (Eds.), *Second Language Interaction in Diverse Educational Contexts* (pp. 247–268). Amsterdam/Philadelphia: John Benjamin Publishing Co.

**Wu, S.-L. and Ortega, L.** 2013. Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Language Annals*, *46*(4), 680–704.

**Xi, X., Higgins, D., Zechner, K. and Williamson, D.** 2012. A comparison of two scoring methods for an automated speech scoring system. *Language Testing*, *29*(3), 371–394.

**Ye, F.** 2014. *Validity, reliability, and concordance of the Duolingo English Test*. Pittsburgh, PA. Retrieved from https://s3.amazonaws.com/duolingo-certifications-data/CorrelationStudy.pdf

**Zheng, Y. and Cheng, L.** 2008. Test review: College English Test (CET) in China. *Language Testing*, *25*, 408–417.

*Received: 3 June 2015*

*Accepted: 5 August 2015*

*Cite this article as:*

**Ockey, G.J. & Zhi, L.** 2015. "New and not so new methods for assessing oral communication". *Language Value* 7 (1) 1-21. Jaume I University ePress: Castelló, Spain. http://www.e-revistes.uji.es/languagevalue. DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.2

**ISSN 1989-7103**

*Language Value* November 2015, Volume 7, Number 1 pp. 22-45
http://www.e-revistes.uji.es/languagevalue ISSN 1989-7103

# Comparing candidates' beliefs and exam performance in speaking tests

**Cristina Pérez-Guillot**
mcperez@upv.es
**Julia Zabala-Delgado**
juzadel@upv.es
**Universitat Politècnica de València, Spain**

**ABSTRACT**

The development of a language exam is not a linear process but rather a round cycle in which, by using the test, we obtain information that will in turn be applied to improve each of the steps in the cycle. The goal of our study was to analyse students' beliefs about their performance in the speaking section of a language proficiency exam and compare them with their actual results in the exam, in order to determine whether their beliefs were based on their actual level of competence or if they were based on other factors, such as anxiety or stress due to the particular characteristics of this section of the exam.

**Key words**: *perceptions, competence, candidates, proficiency exam, oral skills, exam development*

## I. INTRODUCTION

The development of a language test is a process that involves several stages, from designing the test and endowing it with adequate contents, to administering the test and analysing the results obtained. However, developing a test is not a linear process but instead a round cycle in which, by using the test, we obtain information that will in turn be applied to improve each of the steps in the cycle.

The goal of our study was to analyse students' beliefs about their performance in the speaking section of a language proficiency exam in relation to other skills, and compare them with their actual results in the exam in order to determine whether their beliefs were based on their actual level of competence or if they were based on other factors, such as anxiety or stress due to the particular characteristics of this section of the exam. Since *the examination of reliability depends upon our ability to distinguish the effects (on test scores) of the abilities we want to measure from the effects of other factors*


DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.3

(Bachman 1990, p.163), being able to differentiate external factors from the actual level of competence of the candidate would necessarily improve test reliability.

Consequently, determining the basis for their beliefs, either factual or self-perceived, would allow us to determine which aspects of the process could be modified to improve the reliability and thus the quality of our exam.

## II. STATE OF THE ART

The process of foreign language acquisition has been examined from different points of view – cognitive, psychological, linguistic, pragmatic and cultural, to mention just a few – and the exact nature of the process is still unknown.

Traditional language learning theories focus mainly on the study of what is learned and what is not learned in a language, explaining both processes by means of the strategies used to acquire knowledge and the reasons for success or failure in acquisition. This approach focuses on learning itself, on specific objectives and on the means used to achieve them and the results obtained. However, it fails to pay attention to the factors surrounding this process, the factors that add complexity to the process and that include not only objective components, but also subjective or external components, which will largely contribute to the final outcome. The starting point is therefore to consider language learning as a broad field in which external factors play an important role and, amongst them, those characteristics that are individual to each student and make their learning unique.

Examining the process of second language acquisition from this broad point of view, there are a number of subjective variables that belong to the students' individual field and that have a significant effect on their learning. This explains the different degrees of success in learning a language achieved by different subjects who follow the same programme and have a comparable intellectual ability. However, it is worth mentioning here that, although social and affective strategies are mentioned in the literature (Dörney 1994; Gardner and Lambert 1959; Hardison, 2014; Horwiz 1995; Jee 2014; Sparks et al. 2011), many of the authors reviewed (Bachman 1990; Bachman and Palmer 1996; Cohen 2003; O'Malley and Chamot 1990) focus primarily on cognitive and

metacognitive variables and the relationships between them, considering social and affective factors as crucial but difficult to quantify.

However, some studies (Baddeley 2007; Carroll and Sapon 1959; Conway et al. 2007; Pimsleur 1966) attribute this difference in the degree of success to the students' ability, on the cognitive aspect, leaving aside the emotional aspect, i.e. their attitude, motivation and beliefs about their own learning process. However, "if we were to devise theories of second language acquisition or teaching methodologies that were based only on cognitive considerations, we would be omitting the most fundamental side of human behaviour" (Brown 2000: 142). The affective domain is difficult to describe scientifically since it refers to emotion or feeling, yet the emotional side of human behaviour is intrinsically related to the cognitive side and needs to be taken into consideration.

Sustained by developments in the field of foreign language teaching towards student-centred learning, the study of these factors has become increasingly important together with, more specifically, the study of how students' perceptions and beliefs are a fundamental aspect on their path to learning a language. In fact, Foreign Language Anxiety (FLA) has been defined as a particular type of anxiety occurring specifically in foreign language learning situations,

> "a distinct complex construct of self-perceptions, beliefs, feelings and behaviours related to classroom language learning arising from the uniqueness of the language learning process", "a phenomenon related to but distinguishable from other specific anxieties" (Horwitz and Cope 1986: 128-129)

and it is made up of three principal components: (a) communication apprehension, (b) test anxiety, and (c) fear of negative evaluation.

Motivation has also been considered influential in the degree of success of foreign language students, and this includes both instrumental motivation – the desire to obtain something from studying a second language – and integrative motivation – the desire to integrate into the culture of the second language (Gardner and Lambert 1959). In fact, and although both types of motivation contribute to second language learning success, students who are the most successful are those who are interested in the culture of origin and native speakers and have a desire to integrate into the society in which the language is used (Falk 1978).

Bachman and Palmer (1996) observed two types of variability in students' performance in language tests: (1) variability due to differences between individuals in terms of the language skills, strategies and processes used, as well as personal characteristics such as cultural and emotional differences, etc., and (2) variability due to the different characteristics of the method or tasks used in the test, such as the assessment modes or types of tasks used. According to Dornyei (2009) individual differences should be considered as higher level amalgams or constellations of cognition, affect and motivation that act as "wholes".

As a consequence of this approach, the subjective variables of the acquisition process mentioned above also play an important role in language testing, since the design of a test needs to take into consideration not only the characteristics of the tasks – test format, input provided, time allotted – but also the individual characteristics of the users – the positive or negative emotions or feelings they may have about their learning process, the examiner, the subject or context, the presence or absence of excessive anxiety when faced with the task, their motivation, etc. Accordingly, two aspects need to be taken into consideration simultaneously: (1) the characteristics of the task, which need to reflect the construct of the test and mirror target language use, and (2) the individual characteristics of the learner, which will affect their learning process and therefore their performance in a test situation.

As can be seen from the aforementioned arguments, although it would be desirable that the primary factor in the outcome of a language test were the ability of the test-taker or the adequacy of the test construct and structure of the tasks, in actual fact there are many other variables coming into play, ranging from the context to the individual characteristics of each test-taker.

Such factors become even more relevant in assessing speaking, since speaking in a foreign language is perhaps the most difficult skill to master as it involves a complex process of constructing meaning (Celce-Murcia and Olshtain 2000) which is performed at the same time as the act of speaking and therefore requires the planning and simultaneous monitoring of utterances. In fact, the ability to express oneself orally in a foreign language is a fundamental part of mastering language use and, as mentioned by

Luoma (2004), reflects not only our personality but also our self-image and ability to reason:

> Speaking is also the most difficult language skill to assess reliably. A person's speaking ability is usually judged during a face-to-face interaction, in real time, between an interlocutor and a candidate. The assessor has to make instantaneous judgements about a range of aspects of what is being said, as it is being said. This means that the assessment might depend not only upon which particular features of speech (e.g. pronunciation, accuracy, fluency) the interlocutor pays attention to at any point in time, but upon a host of other factors such as the language level, gender, and status of the interlocutor and the personal characteristics of the interlocutor and candidate. (Luoma 2004: ix).

Furthermore, speaking a language is especially difficult for foreign language learners because effective oral communication requires the ability to use the language appropriately in social interactions (Fulcher 2003).

Traditionally, most students sitting official exams show high levels of stress when dealing with the speaking section of the test and explain their reaction by expressing their doubts about their own speaking ability (Phillips 1992; Stephenson and Hewitt 2001). However, and in light of the above, we believe that the fact that the speaking section of the test causes more stress in students is, in many cases, not because of their ability or lack of it, but because of the construct of speaking mentioned. As Bandura (1997: 37) states: "perceived self-efficacy is not a measure of the skills that one possesses, rather it is a belief about what one can do in the future, and under different conditions, with the skills that one has". Consequently, perceived self-efficacy, the extent of one's belief in one's own ability to reach goals, will probably influence the way people will react in the face of difficulties and, therefore, a more positive perception of one's skills will influence performance in the real world. Real performance in the real world is in turn what performance in an exam situation should be expected to mirror and what exam tasks and context should be expected to elicit.

## III. METHODOLOGY

### III.1. Participants

The participants in our survey[1] were the candidates sitting the CertACLES exam at the Universitat Politecnica de Valencia (UPV) in June 2014. There were a total of 324 candidates for the three examinations carried out in June 2014: 101 for the B1 examination, 186 for the B2 examination, and 37 for the C1 examination. Our survey obtained 201 answers, that is, 62% of the candidates voluntarily took part in the study. It is interesting to note that the higher the level of language exam, the higher the participation of candidates: 55% of B1 candidates participated in the study, compared to 58% of B2 candidates and 62% of C1 candidates.

CertACLES exams are proficiency exams developed by the Language Centre of the UPV in accordance with the model developed by the Spanish Association of Language Centres in Higher Education (ACLES 2011a, b). ACLES introduced a model for a language examination – the CertACLES model (ACLES 2011b) – that would be followed by all higher education institutions belonging to the organisation and that was intended to allow for the assessment of communicative competence with a standard and comparable framework which all member institutions needed to adhere to. This framework is solid enough to provide for a standard tool for measuring language ability while allowing each individual university to adjust their exam to meet the needs of their environment. Each university is therefore in charge of designing its own individual exams, which have to comply with the framework but have to take into consideration each particular context, not only in terms of test construct and specifications, but also in terms of administration dates and frequencies. CertACLES exams measure the four skills – reading, writing, listening and speaking – and give equal weight to each section. They were officially recognised by the Spanish Conference of Rectors in 2011 (CRUE, 2011) and by the Regional Government in Valencia in 2013 (DOGV, 2013).

The profile of the candidates was expected to consist mainly of a student population, although students and staff from other universities in the area who do not offer their own language proficiency tests were also expected to take part. To further specify the

profile of the candidates, information was requested as part of our survey and the results were as follows[1]:

### *III.1.1. Age range*

Fifty-seven per cent of participants were in the 18 to 30 age range, which would correspond to an examination developed by a higher education institution. Interestingly enough, there is a high percentage of participants in the study who were over 30 years of age, which indicated that there was a high number of participants who were either university staff, alumni or external candidates that wanted to sit an official language exam.

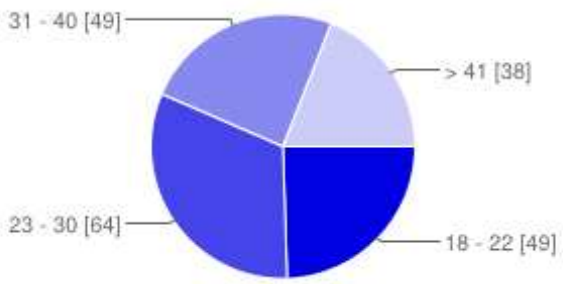


Figure 1. Age range

Table 1. Age range

| **Age** | **No. of candidates within the age range** | **Percentage of candidates within the age range** |
|---|---|---|
| 18 - 22 | 49 | 25% |
| 23 - 30 | 64 | 32% |
| 31 - 40 | 49 | 25% |
| > 41 | 38 | 18% |

### *III.1.2. Education*

The results were as expected given the type of examination and the examining body (a higher education institution), the large majority of participants hold a university degree, and 15% of them have doctoral degrees.

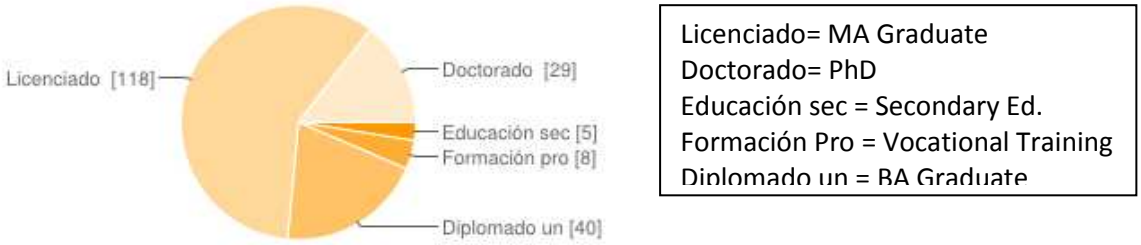


Figure 2. Education

Table 2. Education

| **Education** | **No. of candidates** | **Percentage of candidates** |
|---|---|---|
| Secondary education | 5 | 3% |
| Vocational training | 8 | 4% |
| University degree (BA) | 40 | 20% |
| University degree (MA) | 118 | 59% |
| PhD | 29 | 15% |

### *III.1.3. Motivation for taking a language exam*

Since we were analysing students' perceptions and FLA when confronting a speaking test, we were interested in knowing their reasons for taking the exam. Their reasons would indicate the orientation of their motivation, either integrative or instrumental

(Gardner and Lambert, 1959), and would thus help predict their degree of success. As we can see from the results, illustrated in Table 3, in 75% of the cases the motivation for taking the exam was instrumental. Only 25% of the candidates showed an integrative motivation and stated that the reason for taking the exam was personal satisfaction, which was assumed to mean travelling to other countries and meeting native-speaking people as well as learning the culture of native-speaking countries.

Table 3. Motivation

| **Reasons for taking an exam** | **No. of candidates** | **Percentage of candidates** |
|---|---|---|
| Personal satisfaction | 50 | 25% |
| Mobility grant | 18 | 9% |
| Graduation requirement | 38 | 19% |
| Professional projection | 130 | 65% |
| Other | 10 | 5% |

### III.2. Materials

The main goal of our study, as stated in our introduction, was to analyse students' beliefs about their performance in the speaking section of a language proficiency exam and compare them with their actual results in the exam, in order to determine whether their beliefs were based on their actual level of competence or if they were based on other factors, such as anxiety or stress due to the particular characteristics of this section of the exam. In order to do this, we needed to examine, on the one hand, their feelings with respect to the different sections of the exam in terms of perceived difficulty and candidate anxiety and, on the other hand, the results obtained by the candidates in the actual examination. By looking at the results obtained in the speaking section of the

examination and comparing these results with those obtained by the same candidate in the other sections of the exam, we would be able to compare candidates' self-perception with their actual performance, and thus devise strategies that could be introduced into the design of the examination to reduce the negative influence of factors external to the candidates' language competence.

Accordingly, our study would initially be divided into two different steps: (1) analysing students' perceptions of their performance in the exam, and (2) analysing students' actual results in the exam.

### *III.2.1. Analysing students' perceptions of their performance in the exam:*

For the sake of practicality, we decided to use the free tool for generating surveys provided by Google, *Google Forms*, to create our survey. This tool allowed us to design a relatively simple survey with automatic data processing and charting, but with a compatible table in Excel format to allow further modification or alternative processing of the data obtained. Likewise, the system also allowed the creation of a link to the survey that could be sent to the students' email addresses from the Language Centre's email account. The fact that the tool involved no additional costs and that it was user-friendly, only requiring a few minutes to be able to start using it, was also a key factor in our decision. Google Forms requires the individual who is designing and administering the survey to have a gmail account. This email account does not need to be the one used to send the survey to the participants – which was a question of concern for us since we did not want to use an account not belonging to the university – but it will be visible in the link sent and therefore needed to have some appearance of reputability. To achieve this, we set up a gmail account for the language centre in which not only the name of the language centre was specified, but also the initials of the university, to make the sender easily identifiable.

Before designing our survey, we had to take into consideration the characteristics of the tool we were going to use. Google Forms provides different layouts, allowing us to send respondents in different directions depending on the answer and allowing different types of question formats. The question formats provided are as follows: *Text* – open questions with short answers; *Paragraph Text* – open questions with longer answers;

*Multiple Choice* – controlled answers where one option is chosen from among several; *Checkboxes* – controlled answers where users select as many options as they like; *Choose from a list* – controlled answers in which users select one option from a dropdown menu; *Scale* – controlled answer in which users rank something on a scale of numbers; *Grid* – controlled answer in which users select a point from a two-dimensional grid; *Date* – controlled answer in which users pick a date on a calendar; *Time* – controlled answer in which users select a time of day or a length of time. Our initial intention was to use either the *Text* format or the *Paragraph Text* format, preferring the short-answer questions for the sake of conciseness. However, we also wanted to favour easy processing of the information and we realised that using this type of format would not allow the data to be processed automatically. In the end and after much consideration, we decided to use a Multiple Choice format since it limited the respondents' production and allowed for easier processing by automatically generating charts and summaries of results. In fact, Google Forms can be connected to spreadsheets in Google Sheets, and if a spreadsheet is linked to the form, responses will automatically be sent to the spreadsheet from where information is taken and automatically summarised and presented in a summary of results. For those questions in which we intended to measure a level (level of difficulty, anxiety, etc.), we used the *Scale* format, since the processing of results was similar to that of the Multiple choice format.

### *III.2.2. Analysing students' actual results in the exam:*

The candidate's marks that were analysed belonged to the speaking exam of the CertACLES Certification paper administered in July 2014. This exam aims to evaluate the communicative competence of the candidates and the contents and construct of the exam and the marking criteria are based on the CEFR descriptors. To that end, the exam evaluates the four main communicative macro-skills, i.e. speaking, listening, writing and reading, each with a specific weight of 25% of the total score of the exam.

A candidate is considered to have reached the corresponding language level if the final mark is equal to or higher than 60% of the total possible points, provided that a

minimum of 50% of the possible mark has been attained in each skill. The marks are awarded on a scale of 0 to 10 points (100%) expressed to one decimal place:

- Between 6.0 and 6.9 points (60%-69% of total marks possible) = PASS
- Between 7.0 and 8.9 points (70%-89% of total marks possible) = MERIT
- Between 9.0 and 10 points (90%-100% of total marks possible) = DISTINCTION

The speaking test is conducted by two oral examiners, an interlocutor and an assessor, with paired candidates. The reason for choosing this task format had to do with our goal of mirroring real-life communication, while minimising anxiety and tension in the candidate. As Heaton (1988) states, interviews are adequate attempts to assess oral skills but students are not placed in "natural" speech situations and they are therefore subject to psychological tensions which will necessarily affect their performances. CertACLES exams attempt to minimise this effect by having an interlocutor and an examiner present in the interview to allow the interlocutor to focus on candidates while they speak and avoid interruptions that would occur while the interlocutor takes notes. In this way the interlocutor is responsible for conducting the interview and for giving a global impression of the overall communicative ability of the candidate, but it is the assessor who is responsible for providing an analytical assessment of each candidate's performance. The assessor does not take part in the interaction with the candidates and is thus able to apply a detailed analytical scale with four criteria: (1) grammar, which refers to appropriate use of grammatical forms; (2) vocabulary, which measures the accuracy and the use of lexical forms; (3) Discourse management, which focuses on relevant discourse and coherence; and (4) pronunciation and interactive communication, where the focus is not only on the ability to be understandable but also the candidate's ability to take an active part in the development of the discourse. Moreover, having two candidates taking the exam together allows for equal interaction where there is no power relationship (interlocutor/candidate), but instead a conversation between two members of the same peer group.

Since the exam aims to obtain different types of oral production in a single interview, the interview is divided into three parts:

**Part One**. Conversation between the interlocutor and each candidate. There is a set of standard questions on personal details and preferences grouped by topic (country of

origin, studies, hobbies and interests, education, travel, technology, etc.). The interlocutor can decide which questions to ask based on the responses obtained in order to elicit enough performance by the candidate for assessment within the time frame available.

**Part two**. Simple standardised rubric with minimal language input. The candidates are each given one or two photographs (B1 candidates have one picture to describe and B2 and C1 candidates are given two pictures to allow them to use more complex vocabulary for comparison and contrast). The objective is therefore to compare and contrast during an individual long turn**.** After each candidate has spoken, their partner is asked one question related to the topic.

**Part Three**. Conversation between candidates. The interlocutor gives some pictures to the candidates. They are asked to speak for a set amount of time and justify their opinions, speculate, express preferences and draw conclusions within the target language use defined for each level of examination. At the end of the interaction the interlocutor may ask the candidates further questions on the topic.

### III.3. Procedure

#### *III.3.1. Analysing students' perceptions of their performance in the exam:*

A survey was designed with multiple choice questions on the candidates' profile and their opinion on the difficulty of the different sections (from 0 to 5, 0 being the easiest and 5 being the most difficult). Once the survey had been designed and implemented in Google Forms, we generated the link and sent it out to all the candidates participating in the June 2014 exam sessions (B1, B2 and C1 candidates). The email was sent after they had taken the examination so that their opinions were based on the same exam from which their marks were going to be analysed in our second step. Moreover, and to avoid bias in their responses, the link was sent before results were published and a deadline was established for the collection of responses, no responses being accepted if received after the publication of exam results.

### *III.3.2. Analysing students' actual results in the exam:*

An excel spreadsheet was designed to introduce the candidates' results for the different parts of the exam, that is, listening, reading, writing and speaking. This would facilitate the analysis of the results, and allow for an analysis of the weaknesses and strengths of the different candidates.

A spreadsheet was designed for each of the examinations (B1, B2, C1) and the structure was as follows:

**Candidate number** is the number assigned to each candidate for easier identification; **ID, Name, Surname**, are fields needed to issue the official accreditation certificate; **listening mark, reading mark, writing mark, speaking mark** are individual marks per skill, and **overall mark** is the mark obtained from the weighting of the different skills. Finally, a **register number** is provided for the certificate issued.

## IV. RESULTS

After collecting the data from the survey and analysing the results obtained by the students in the different parts of the exam that they had rated as regards difficulty, the results were as follows.

As we can see in Figure 4, for the speaking section, most of the candidates gave a rank of 3 or higher, indicating higher difficulty; in fact, 45% of the candidates ranked the level of difficulty of the speaking section as 4 or 5.

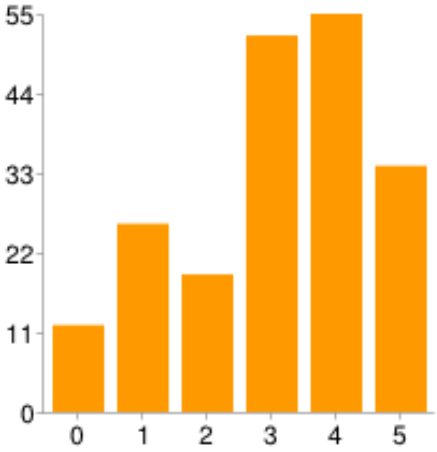


Figure 3. Perceived difficulty of the speaking section

Figure 5 shows the level of difficulty of the listening section, which was ranked higher, although only slightly so; in fact, 53% of the candidates ranked it as having a level of difficulty of 4 or 5.

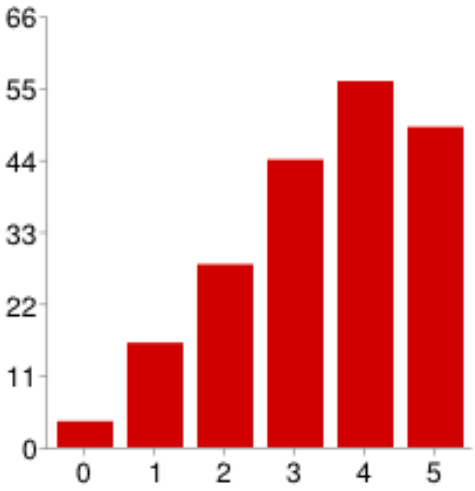


Figure 4. Perceived difficulty of the listening section

Reading was ranked the easiest, as we can see in Figure 6. 72% of the candidates ranked it between 0 and 3 on the scale of difficulty and 13% of those indicated the level of difficulty as non-existent.

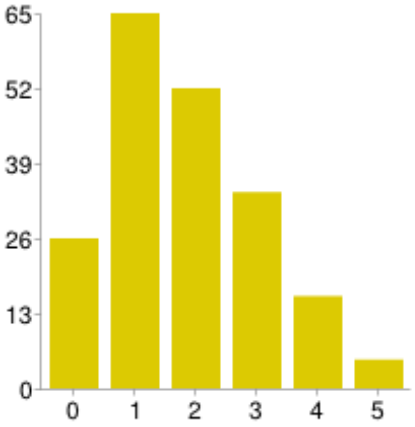


Figure 5. Perceived difficulty of the reading section

As for writing, as shown in Figure 7, it was considered a medium-difficulty section, with only 30% of the candidates ranking the difficulty of the exam above 3.

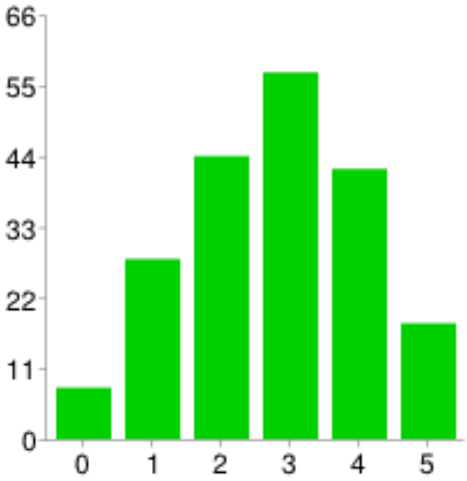


Figure 6. Perceived difficulty of the writing section

In light of these results, candidates considered the listening section to be the most difficult, closely followed by the speaking section, and by the writing and reading sections, which were far behind in terms of perceived difficulty. This contradicted our initial beliefs, since with their own reactions in the classroom and their reluctance to complete speaking and writing tasks, our students usually express more anxiety towards productive skills in the classroom and there is a higher demand for writing and speaking preparation courses, leaving reading and listening as areas that are not specifically prepared by students but are learnt or practised in general English courses.

As for the candidates' marks in the examination, which are translated in the table below, the results obtained were as follows:

As we can see in Figure 8, the vast majority of the candidates (71%) obtained a mark that allowed them to pass the exam.

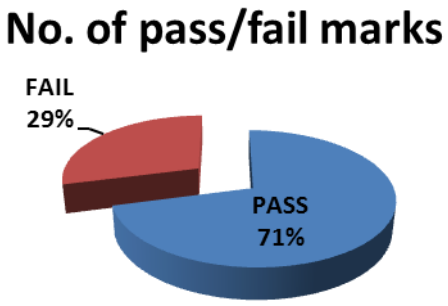


Figure 7. No. of pass/fail marks

It is also important to highlight that, as we can see in Figure 9, 210 candidates passed the examination, 122 of them obtaining either a merit or a distinction.

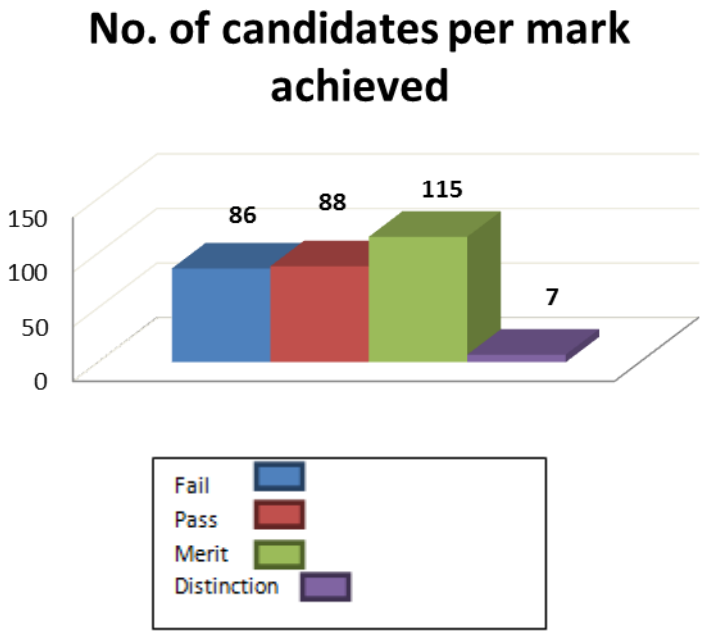


Figure 8. No. of candidates per mark achieved

Figure 10 shows the number of candidates who failed the examination per skill and level of examination. The figure shows three clusters of results indicating the candidates at levels B1, B2 and C1. As we can see, at level B1 a higher percentage of candidates failed the speaking section. In contrast, B2 candidates have higher failing rates in the

writing section, while speaking has the second best results after reading. For the C1 examination, the results are similar to B2, speaking having the second best results of all the sections of the exam.

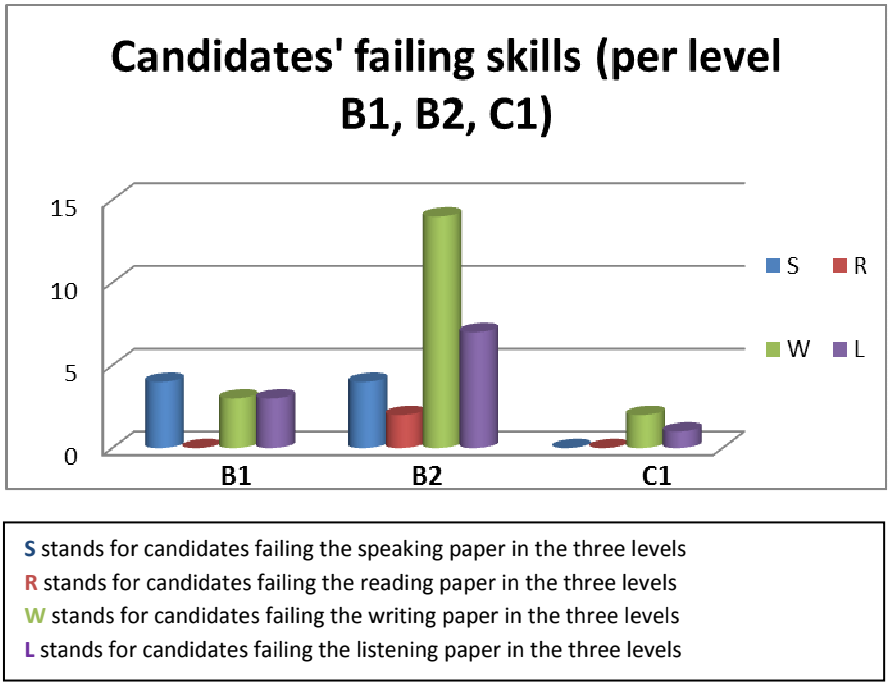


Figure 9. Candidates with only one failed skill

To further illustrate Table 10, in Table 5 we can see the number of candidates with one or more failed skills and a specification of the skills failed. It can also be observed how the results further indicate that candidates predominantly fail because of the writing paper, particularly at higher levels. The lower figures at C1 are due to the small number of candidates who failed (only 5 candidates failed the C1 examination).

Table 4. Illustrative table of candidates failed per level

| | B1 | B2 | C1 | Total number of candidates |
|---|---|---|---|---|
| **S** | **4** | **4** | **0** | **8** |
| **S-W** | 1 | 8 | 0 | 9 |
| **S-R** | 1 | 1 | 0 | 2 |
| **S-L** | 1 | 2 | 1 | 4 |
| **R** | **0** | **2** | **0** | **2** |
| **R-L** | 2 | 1 | 0 | 3 |
| **R-W** | 0 | 0 | 0 | 0 |
| **W** | **3** | **14** | **2** | **19** |
| **L-W** | 2 | 3 | 1 | 6 |
| **L** | **3** | **7** | **1** | **11** |
| **W-R-L** | 1 | 2 | 0 | 3 |
| **S-W-L** | 1 | 2 | 1 | 4 |
| **S-W-R** | 1 | 1 | 0 | 2 |
| **S-R-L** | 2 | 0 | 0 | 2 |
| **ALL** | 2 | 0 | 0 | 2 |

## V. CONCLUSIONS

The results of our study show a mismatch between self-perceived efficacy and actual performance, particularly for higher levels of proficiency. In fact, although according to our survey most candidates ranked speaking as the second most difficult skill, the results of the exam show that the number of candidates who fail the exam because of the speaking section is comparatively lower. This is not the case, however, for B1 candidates, who showed a more accurate level of self-perceived efficacy, as seen in the results in Table 10. This is in line with Bandura's (1997) statement about the perception of self-efficacy, in that candidates produce less accurate assessments as they progress through higher levels of language study.

However, as stated in our results, candidates' perceived efficacy is accurate in the case of reading and listening, since they are both the easiest and the most difficult sections and the candidates' mark reflects this as being so. The greatest mismatch is therefore in the productive skills, since writing is considered a medium-difficulty skill and it is in fact the skill in which the candidates' performance is ranked lower. Speaking is perceived as a high-difficulty skill but this difficulty is not reflected in the candidates' results, as few of them fail because of the speaking section. Therefore, candidates' perception of their efficacy in the speaking section does not seem to correspond to their actual ability, which would indicate that their perceptions are indeed influenced by factors that are external to their actual performance. FLA comes into play and, consequently, modifications in the test process should be arranged to reduce anxiety for candidates. Some of the modifications suggested would be the following:

- Organising exams with paired interviews whenever possible in order to avoid relationships of power with the examiner and thus reduce stress.

- Facilitate the presence of an assessor whenever possible in order to allow the examiner to act only as the interlocutor.

- Individual arrangements for candidates to facilitate schedules and allow them to choose the time of day at which they would feel more comfortable taking the test.

- Flexible examination dates, to eliminate stress in candidates who have conflicting commitments (academic, professional, family-related, etc.).

- Preparation time and warm-up questions to allow them to feel more at ease with the topic, as well as get to know both the interviewer and the other candidate (in cases when the interview is paired).

- Additional prompts to facilitate discussion topics during the exam and prevent candidates from relying on their resourcefulness or imagination.

- Start and finish the interview on a positive note to improve confidence and self-image, which could then be mirrored in real-life performance.

We consider that the results of our study call for further research on factors outside the content of the exam, factors related to administration and organisation, as well as those related to individual characteristics of the candidate (personality, background, etc.), which will undoubtedly explain the difference between perceptions and actual results.

## Notes

[1] The survey was carried out in Spanish to allow all participants to fully understand the questions; the titles and legends in the graphs are therefore in Spanish. Under each graph there is a representation of the information in table format and translated into English.

## REFERENCES

**ACLES** 2011a. Asociación de Centro de Lenguas de Educación Superior. 20 April 2014 <http://www.acles.es/>

**ACLES** 2011b. Modelo de acreditación de exámenes de acles. 20 April 2014 <http://www.acles.es/multimedia/enlaces/9/files/fichero_29.pdf>

**Bachman, L.F. and Palmer, A.S.** 1996 *Language testing in Practice*. Oxford: Oxford University Press.

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

**Baddeley, A.** 2007. *Working memory thought, and action*. Oxford: Oxford University Press.

**Bandura, A**. 1997. *Self-Efficacy: The Exercise of Control*. New York: Freeman.

**Brown, D.H.** 2000. *Principles of language learning & teaching*. New York: Longman.

**Carroll, J.B., and S.M. Sapon.** 1959. *Modern language aptitude test*. New York: The Psychological Corporation/Harcourt Brace Jovanovich.

**Celce-Murcia, M., and E. Olshtain**, 2000. *Discourse and context in language teaching: A guide for language teachers*. Cambridge University Press.

**Cohen** 2003. "The learner's side of foreign language learning: Where do styles, strategies, and tasks meet?" *International Review of Applied Linguistics in Language Teaching,* 41(4), 279-291.

**Conferencia De Rectores De Las Universidades Españolas (CRUE)**. 2011 (8 de septiembre). Propuestas sobre la acreditación de idiomas. 25 April 2014 <http://www.acreditacion.crue.org/export/sites/Acreditacion/doc/Propuestas_acreditacion_de_Idiomas-AG8-09-2011.pdf>

**Conway, A., C. Jarrold, M. Kane, A. Miyake, and J. Towse**, (Eds.). 2007. *Variation in working memory*. Oxford: Oxford University Press

**Diari Oficial De La Comunitat Valenciana (DOCV)**. 2013. Orden 93/2013, de 11 de noviembre, de la Consellería de Educación, Cultura y Deporte, por la que se modifica el anexo único denominado certificados y diplomas que acreditan la competencia en lenguas extranjeras del Decreto 61/2013, de 17 de mayo, por el que se establece un sistema de reconocimiento de la competencia en lenguas extranjeras en la Comunitat Valenciana y se crea la Comisión de Acreditación de Niveles de Competencia en Lenguas Extranjeras. 17 May 2014 <http://www.docv.gva.es/datos/2013/11/13/pdf/2013_10793.pdf >

**Dörnyei, Zoltán**. 2009. "Individual differences: Interplay of learner characteristics and learning environment". *Language Learning,* 59.s1, 230-248.

**Fulcher, G.** 2003. *Testing second language speaking*. New York: Pearson Longman.

**Gardner, R.C., and Lambert, W.E.** 1959. "Motivational variables in second-language acquisition". *Canadian Journal of Psychology/Revue canadienne de psychologie*, *13* (4), 266-272.

**Hardison, D.M.** 2014. Changes in Second-Language Learners' Oral Skills and Socio-Affective Profiles Following Study Abroad: A Mixed-Methods Approach. *Canadian Modern Language Review*, 70 (4), 415-444.

**Heaton, J.B.** 1988. *Writing English language tests* (2nd ed.). Harlow: Longman.

**Horwitz, E.K., Horwitz, M.B., and Cope, J.** 1986. "Foreign language classroom anxiety". *The Modern Language Journal*, 70 (2), 125-132.

http://dx.doi.org/10.1017/CBO9780511733017

**Jee, M.J.** 2014. "Affective factors in Korean as a Foreign Language: anxiety and beliefs". *Language, Culture and Curriculum*, 27 (2), 182-195.

**Luoma, S.** 2004. *Assessing speaking*. *Cambridge University Press.*

***O'Malley, J.M. and Chamot, A.U.*** *1990. Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.

**Pimsleur, P.** 1966. *Pimsleur language aptitude battery (PLAB).* New York: The Psychological Corporation.

**Phillips, E.M.** 1992. "The effects on language anxiety on students' oral test performance and attitudes". *The Modern Language Journal*. 76 (1), 14-26.

**Sparks, R.L., Humbach, N., Patton, J.O.N., and Ganschow, L.** 2011. "Subcomponents of second language aptitude and second language proficiency". *The Modern Language Journal*, 95 (2), 253-273.

**Stephenson, J., and Hewitt, E.** 2011. Los desafíos de realizar la investigación en el aula: el caso de un estudio réplica sobre la ansiedad en el aprendizaje del inglés. *Dedica. Revista de educação e humanidades*, 1, 365-382

*Received: 3 December 2014*

*Accepted: 22 July 2015*

*Cite this article as:*

**Pérez-Guillot, C. & Zabala-Delgado, J.** 2015. "Comparing candidates' beliefs and exam performance in speaking tests". *Language Value* 7, 22-44. Jaume I University ePress: Castelló, Spain. http://www.e-revistes.uji.es/languagevalue.

DOI**:** http://dx.doi.org/10.6035/LanguageV.2015.7.3

**ISSN 1989-7103**

*Language Value* November 2015, Volume 7, Number 1 pp. 46-61
http://www.e-revistes.uji.es/languagevalue ISSN 1989-7103

# Speaking an additional language: Can study abroad do the trick?

**María Juan-Garau**
maria.juan@uib.eu
**Universitat de les Illes Balears, Spain**

**ABSTRACT**

The present article sets out to provide an overall picture of the acquisition of speaking abilities in a given second language acquisition (SLA) learning context, namely study abroad (SA), so as to ascertain whether SA can indeed 'do the trick', as it is popularly assumed, and, if so, under which conditions. Section 1 will characterise the SA context and look at how it can affect oral performance, bearing in mind the opportunities for target language contact and practice it offers learners. Section 2 will focus on the specific linguistic benefits in the oral domain that can accrue in this learning context. In this regard, we will present an overview of empirical research findings with special attention to the SALA (Study Abroad and Language Acquisition) and COLE (Context, Contact and Competence Level) research projects. Section 3 will concentrate on individual variables that may affect oral development abroad in combination with external variables such as SA programme conditions. Finally, Section 4 will provide a summary of the main ideas presented and draw some conclusions.

**Keywords**: *study abroad, speaking skills, oral competence, EFL, SLA, learning context*

## I. STUDY ABROAD AS A LANGUAGE LEARNING CONTEXT

Research on bilingualism and multilingualism has paid increasing attention to the learning contexts in which languages are acquired, whether naturalistic or instructed. Such research aims at finding empirical evidence regarding the potential efficiency of a given learning context in promoting students' ability to communicate more fluently, accurately and with higher degrees of complexity in an additional language. In the present article we will focus on the study abroad (henceforth SA) context, situated at the naturalistic end of the learning context continuum (see Juan-Garau 2012, Pérez-Vidal 2011).

As a natural learning setting, SA allows for the testing of learner hypotheses by paying attention to relevant input, for the practice of common speech acts embedded in daily routines, and for the contextualisation of learning in a myriad of authentic situations, enabling better memorisation and retrieval, in contrast to the relative dreariness of


DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.4

explicit learning within the confines of a classroom (DeKeyser 1991, Juan-Garau and Pérez-Vidal 2007). SA conditions, thus, allow learners to pay attention to form incidentally while focusing on meaning (Ortega 2005).

SA research has proliferated over the last decades, basically as a result of the increasing number of SA programmes that secondary schools and particularly higher education institutions offer their students as part of their educational experience (Wang 2010), giving rise to a massive-scale student mobility operation in Europe and worldwide. Such research can shed new light on both the positive aspects and the challenges of SA. Most studies of language learners abroad have focused on the acquisition of oral skills, reflecting higher expectations for gains in this area. Indeed, the SA context affords constant opportunities for learners to be exposed to comprehensible input from a plethora of target-language speakers and to practise speaking by getting a variety of things done while interacting and negotiating meaning in a second (L2) or additional language. Learners, however, need to seize the contact opportunities the SA context affords in order to enhance their speaking abilities.

Barbara Freed's 1995 edited volume, entitled *Second Language Acquisition in a Study Abroad Context*, constituted the first important landmark in the study of the achievements of language learners abroad and inspired further research in this area, which has grown exponentially since then. In his foreword to the volume, as Kinginger (2013) aptly remarks, Charles A. Ferguson referred to the "myths" that surround SA, including the belief that the only way to achieve "real fluency" in an L2 or foreign language is to travel to a place where the target language is spoken. In the next section we will survey SA research that has focused on learners' speaking abilities. In so doing, we will try to dispel some of the myths surrounding oral development in SA settings and to ascertain whether SA can "do the trick" as far as the acquisition of oral competence in an additional language goes.

## II. BENEFITS IN THE ORAL DOMAIN AFTER SA

Overall oral proficiency, often measured through the ACTFL (American Council on the Teaching of Foreign Languages) Oral Proficiency Interview (OPI) or similar protocols, has been found to register considerable gains abroad (e.g. Isabelli-García 2003, Kang 2014, Lindseth

2010, Segalowitz and Freed 2004). Research, furthermore, indicates that progress following residence abroad mostly accrues in the oral-aural skills, while it is much less apparent in the written and reading skills (e.g. Brecht and Robinson 1995, Dyson 1988, Lapkin et al. 1995, Meara 1994). We will next examine the domains of complexity, accuracy and fluency (CAF) in relation to speaking skills and will finally consider pronunciation as well. In so doing, we will pay special attention to research conducted as part of the SALA (Study Abroad and Language Acquisition) and COLE (Context, Contact and Competence Level) research projects.[1]

Speech complexity does not seem to increase much as a result of SA. Pérez-Vidal and Juan-Garau (2011) examined SALA-COLE project participants' oral productions regarding both syntactic and lexical complexity in the at-home and SA settings. As regards syntactic complexity, clauses per T-unit were seen to decrease slightly at home, while an increase close to significance was registered abroad. Subordination also increased non-significantly in the SA setting, but not at home. Similarly, lexical diversity, as measured by Guiraud's index, registered a non-significant improvement that was more noticeable abroad. Learners significantly increased, nonetheless, their use of formulaic sequences as an effect of the period spent abroad. All in all, these results confirm the findings obtained by Juan-Garau and Pérez-Vidal (2007) with a smaller sample, which showed non-significant increases abroad in the domain of complexity. Similarly, SALA-COLE researchers Mora and Valls-Ferrer (2012) found that the complexity of their participants' oral productions remained largely unmodified. However, Llanes and Muñoz (2013), in a study comparing the effects of SA and at-home learning contexts in children and adult learners, reported significant gains in oral complexity, with the SA setting appearing more beneficial for children's oral skills than for older learners.

As regards accuracy, there is a dearth of SA studies in comparison to the research conducted in relation to fluency. The existing documentation on speaking proficiency reveals that no substantial development tends to take place in the former domain after residence abroad. Thus, Isabelli (2001) examined the null subject parameter and found that L2 Spanish learners abroad benefited from positive evidence, but some problems remained, suggesting restructuring of the parameter but no resetting. Also investigating subject expression in L2 Spanish with oral data, Lopez Ortega (2003) reported that a number of discourse variables affected the presence or absence of subjects. She found signs of progress in her four students after the SA experience, although the quantitative results did not yield conclusive evidence. Based on interviews, picture-sequence tasks and questionnaires, Longcope's (2003) findings indicated that SA had an immediate reflection on learners' fluency, but not necessarily on their grammatical accuracy or syntactic complexity. Similarly, Serrano et al. (2011) reported increased oral fluency and lexical complexity, but not accuracy (or syntactic complexity), for their SA group in comparison with

the domestic semi-intensive programme. For her part, Torres (2003) investigated the acquisition of Spanish clitics to find that the SA context did not appear to have much linguistic benefit over classroom learning for clitic accuracy and use. Finally, Isabelli-García (2010) found no advantage of the SA abroad over the at-home context in the acquisition of Spanish gender agreement for her intermediate level participants over a four-month period.

Other researchers, however, have provided a rosier view on SA accuracy outcomes. Juan-Garau (2014) examined oral accuracy in forty-three SALA-COLE participants, comparing the gains accrued in the SA context to the gains derived from formal instruction at home. Results showed an advantage for the SA context, where learners became more target-like, with long-lasting effects. Participants with lower pre-departure proficiency levels benefited the most from SA. Some contact variables abroad (e.g. practice of listening and writing activities) were associated with the development of oral accuracy. Similarly, Llanes and Muñoz's (2009) participants improved their oral accuracy rates, measured by means of the ratio of error-free clauses and the average number of errors per clause, after just three-four weeks abroad. For his part, Howard's (2001) SA group reached higher levels of accuracy in aspectual marking in French across a wider range of contexts than the AH groups. Howard (2005) also found a more beneficial effect for SA than for classroom instruction in the expression of past time in natural spontaneous speech for his eighteen Irish learners of French. Likewise, using oral interviews, Isabelli and Nishida's (2005) SA learners showed better performance in the use of the subjunctive in Spanish than the AH groups. After a summer abroad, twenty-two out of thirty L2 Spanish learners in Yager's (1998) study were also observed to improve their grammar, along with their pronunciation and overall oral proficiency, on a nativeness scale. Discrepancies in the findings on oral accuracy following a sojourn overseas evidence the complexity of grammatical development patterns as well as the need for more research in this area.

Oral fluency stands out as the clear winner abroad in comparison with the complexity and accuracy domains. Fluency has been examined with respect to various temporal and hesitation characteristics of speech delivery revealing that SA learners increase the length and rate of their fluent speech runs (e.g. Segalowitz and Freed 2004, Towell et al. 1996) while reducing their pauses and dysfluencies (e.g. Freed et al. 2004, Isabelli-García 2003, Segalowitz and Freed 2004). Nevertheless, several studies have reported that not all learners improve their oral fluency abroad (e.g. Segalowitz and Freed 2004), with individual learner differences accounting for the variability observed. Juan-Garau and Pérez-Vidal (2007) presented evidence from the SALA and COLE projects. They considered the oral linguistic development of twelve Catalan-Spanish undergraduate learners of English who spent three months abroad preceded and followed by periods of formal instruction in the at-home setting over the course of three

academic years. The participants' oral abilities, assessed by means of a role-play conducted in dyads, were positively affected by the SA period, although significant gains were obtained only for the fluency measures. Results also indicated that the opportunities for input and interaction that an SA period provides were conducive to linguistic gains in the oral skills, more so when certain sociolinguistic conditions are met (e.g. working in an international setting with target-language speakers or listening to the media), while other conditions may prove detrimental (e.g. living with Catalan/Spanish companions). Other SALA-COLE researchers have also investigated oral fluency with data elicited through an interview task from advanced learners also involved in these projects. Trenchs-Parera (2009) analysed seven dysfluency phenomena comparing learners' performance to native-speaker baseline data. Her findings unveil SA as a context that helps to correct learners' disruptions, producing the impression of more fluent speech. Mora and Valls-Ferrer (2012), for their part, provide evidence of robust gains in oral fluency as a result of SA (captured mostly through time-related aspects of speech production, such as speech rate, mean length of run, pause frequency and duration, and a composite fluency index), and lack thereof during formal instruction at home (see also Valls-Ferrer and Mora 2014 for similar findings).

Turning to pronunciation, several scholars have focused on the development of SA learners' phonetic and phonological abilities, although this remains a largely uncharted area of research. Previous studies investigating speech learning and phonological acquisition abroad have not shown any consistent improvement in speech production and perception in this learning context. For instance, Simões (1996), Stevens (2001) and Díaz-Campos (2004), focusing on Spanish pronunciation by native speakers of English, reported some benefits in phonological ability for SA students, with the last two studies revealing an advantage of SA over at-home groups in the loss of aspiration when producing unvoiced stops. However, findings in these three studies did not always point in the same direction and did not always report any improvement in the areas considered. Individual differences in the participants and the programmes (e.g. pre-departure proficiency level, length of stay, and time spent using the target language) help to account for differences in pronunciation performance (see section III below).

Research conducted within the SALA-COLE projects has also contributed to analyse the impact of SA on pronunciation. Pérez-Vidal et al. (2011) looked at the perception and production of English vowel contrasts. As regards perception, discrimination scores were always higher in the at-home setting rather than abroad, a finding that was confirmed in a recent study by Mora (2014). This was contrary to our expectations but consistent with previous research (Díaz-Campos 2004, Mora 2008). As for production, an analysis of duration and frequency measurements of the vowel contrasts considered revealed that most of the vowels SA

participants produced differed from native speakers' in terms of both their duration and quality (see also Avello and Lara 2014 for comparable results). We concluded that a short-term SA period did not seem to provide enough experience with L2 sounds for changes in the learners' ability to perceive and produce vowel contrasts to develop noticeably. In fact, special focused practice may be needed for those changes to come about (see, for example, Aliaga-García and Mora 2009). Del Río's (2013) study, nevertheless, provides a more positive view of the effects of SA on pronunciation. She examined the development of foreign accent and comprehensibility in the oral productions of 25 adolescent Spanish learners of English before and after a 3-month period of SA. Results showed that SA participants improved significantly between pre-test and post-test in both speech dimensions, thus confirming the beneficial impact of SA context on L2 learners' oral production.

## III. VARIABLES INFLUENCING ORAL GAINS ABROAD

As already hinted at in section I, it has often been assumed by teachers, students, families and society at large that SA is superior to formal instruction AH for language learning in general and fluency in particular. This belief, however, is often unfounded. In fact, DeKeyser (2007) argues that SA does not always bring about sizeable linguistic gains, as this learning context is not always as obviously ideal to practise a foreign language as people tend to assume. He further claims that the SA and FI contexts are not in opposition, as the declarative and procedural knowledge gathered in the AH setting can bear fruit later on abroad, eventually leading to automatic language use. The problem, however, is that learners abroad, at least to begin with, often feel the pressure to communicate orally in real time and to do so in a fluent and comprehensible manner. Given these difficulties, for SA to bring about language development, certain conditions, to which we turn next, must be met (Pérez-Vidal and Juan-Garau 2011).

One major variable affecting SA outcomes is pre-departure proficiency level, often interacting with other individual learner differences. Several studies have indicated that learners may need to have a certain command of target-language lexical and grammatical forms and structures prior to the stay (e.g. DeKeyser 2007, Isabelli and Nishida 2005, Segalowitz and Freed 2004). Thus, the existence of a threshold level for substantial acquisition abroad to take place has been posited. In a recent study by Kang (2014) with Korean university participants learning English, intermediate-level students were the ones who made the most out of their SA period in terms of improving their speaking abilities, while low-level learners' oral skills remained unchanged. These findings appear to lend support to the aforementioned threshold-level tenet that learners

must have a well-developed lexical and grammatical base to really benefit from SA.

Individual variables are also seen to influence SA gains (Lafford 2006). Among them, learners' readiness and ability to benefit from the contact opportunities at hand has been signalled as an essential factor for SA success. In this respect, Segalowitz and Freed (2004) emphasise the importance of delving into learner-context interactions to explain why a given context may be advantageous to some learners and not others. In fact, Freed et al. (2004) conclude that it is not context *per se* that promotes language gain, but rather the quantity and quality of contact within that context. In the same vein, Kinginger (2013) underscores the variability in post-SA language learning outcomes and the need to further investigate student activity abroad to be able to understand this phenomenon. Language learning, according to this author, is thus seen "as a dialogic, situated affair that unfolds in intercultural contexts and includes significant subjective dimensions" (Kinginger 2013: 5). Therefore, it is important to consider the extent to which learners abroad engage with the host community, which in turn may be affected by how they are received by the institution where they are to study or by their host families. Students abroad may be eager to seek social interaction or, on the contrary, they may avoid contact with their interlocutors. In the latter case, as Kinginger (2013: 5-6) remarks, they may "cling to social networks of co-nationals, or remain virtually 'at home' via the Internet", an option not available some decades ago that can now greatly diminish foreign language practice abroad. Pre-departure preparation can play an important role in avoiding such pitfalls and helping students make the most of their sojourn, as Pérez-Vidal (2014) mentions. The development of self-regulatory strategies among students – including motivation maintenance, goal-setting, and language-learning strategies – is also considered crucial to sustain and intensify foreign language learning abroad (Allen 2013).

Learner attitudes, motivation, and beliefs have also been seen to affect how informal contact abroad relates to acquisition. Thus, Brecht and Robinson (1995: 318) remark that: "Understanding student attitudes may contribute substantively to our knowledge of second language acquisition (SLA) as well as to the success of study abroad programs." These scholars note that students' opinions and beliefs may have an effect on their behaviour and, consequently, influence learning outcomes either positively or negatively. As Churchill and DuFon (2006) indicate in their overview of SA research, several studies have investigated both the effect of pre-departure motivation on the sojourn experience and the effect of time abroad on learner motivation. Hernández (2010) reports that there is a positive relationship between participants' integrative motivation and their interaction with the L2 culture, which, in turn, leads to a significant improvement of their speaking skills (see also Isabelli-García 2006). Yashima et al. (2004), for their part, show that pre-departure attitudinal and motivational

variables relate to the willingness to communicate and the communication behaviour exhibited by Japanese learners of English while abroad. These authors find that willingness to communicate results in more frequent communication in the L2 (see also MacIntyre 2007), as the learner seeks further opportunities to interact with host nationals and, in turn, that behaviour invites hosts to communicate with the sojourner more extensively, thus reinforcing language development in the L2. Research investigating how the length of the period abroad can affect learner attitude and motivation has not produced consistent findings so far. Allen (2002) finds that her learners do not develop positive attitudes towards French native speakers after six weeks abroad, which makes her question the usefulness of short stays to promote positive attitudes towards the target community. Other studies, however, have reported increased motivation after short-term immersion programmes (e.g. Simões 1996). Thus, SALA-COLE researchers Trenchs-Parera and Juan-Garau (2014) find that the SA period heightens the development of positive motivational stands and the reduction of anxiety. Longer stays have generally been thought to promote integration in the host community and yet some of the learners in Hoffman-Hicks' (1999) study did not develop positive attitudes towards native speakers of the target language even after a year abroad. As Yager (1998) points out, if students spend less than a semester overseas, their attitudes and motivation while abroad become even more important in order to take full advantage of their shorter stay in the host country. Another factor that might interact with learner motivation and attitudes towards the host context is previous linguistic experience. Allen (2002) concludes that her more proficient learners are better prepared to benefit from the opportunities to interact that an SA context offers. Yashima et al. (2004), however, find that it might well be the students' perception of their abilities to communicate in the foreign language – rather than their proficiency level as such – that interacts with their willingness to communicate.

SA programme characteristics can also impinge on learning outcomes (for a review of programme features, see Paige et al. 2002 and Coleman 2013). One of particular significance to learner progress is length of stay. In this respect, as Churchill and DuFon (2006) point out, existing research suggests that (a) even short stays can produce benefits (see, for example, Juan-Garau and Pérez-Vidal 2007, Llanes 2010), although (b) longer stays tend to benefit learners more, particularly in the domains of pronunciation and fluency, and (c) learner development only approaches native-like norms at best. Lara (2014), in a study conducted within the SALA-COLE framework, explored the linguistic development abroad of learners whose length of stay differed (three vs. six months). She did not find compelling evidence to suggest that a given length of stay was more beneficial than the other in terms of post-SA benefits. Similarly, Llanes (2010) did not find significant differences in the language gains of participants experiencing a

two- or a three-month stay. These results are indicative that a wider difference in length of stay may be needed (e.g. a semester vs. an academic year) for significant changes to appear.

## IV. CONCLUSIONS

Research to date as regards the development of speaking skills in SA settings reveals that this context – providing extensive opportunities for L2 exposure and practice – has indeed the potential to enhance L2 learners' oral abilities, even though findings are more positive and conclusive for fluency than they are for accuracy and complexity (see, for example, DeKeyser 1991, Freed et al. 2004, Isabelli 2001, Juan-Garau and Pérez-Vidal 2007, Pérez-Vidal and Juan-Garau 2011, Segalowitz and Freed 2004, Towell et al. 1996). The latter two domains would possibly show significant gains after longer, or more intensive, periods of immersion abroad than the ones considered herein (Mora and Valls-Ferrer 2012). As for the effects of SA on pronunciation and phonological development, previous research in L2 speech learning has not provided sufficient robust evidence to suggest that an SA setting can enhance L2 speech perception and production. This is an area, however, where more research is clearly needed. It is also important to remark that considerable variation is found both within and across individuals (i.e. intra- and inter-speaker variation) in the literature on language acquisition in SA contexts,[2] which often makes it difficult to draw a linear developmental pattern over time (see, for example, Jensen and Howard 2014).

Considering all of the above, we gather that the SA context can no doubt be advantageous for the development of language learners' speaking skills, but it may end up not being so for all learners (DeKeyser 2007). Consequently, we should bear in mind Freed et al.'s (2004: 298) caveat that: "it is not the learning context per se that promotes various types of learning but rather […] the nature of the interactions, the quality of the experiences, and the efforts made to use the L2 that render one context superior to another with respect to language gain". Thus, in the preceding section (III), we have considered an array of individual variables that impinge on successful language acquisition abroad, including learners' pre-departure proficiency level and preparation, their readiness to benefit from contact opportunities abroad by engaging with the host community, and the development of self-regulatory strategies as well as positive attitudes, motivational stands and beliefs. These individual variables, often interacting with external variables such as programme characteristics (e.g. length of stay), affect learners' socialization and ultimately their language learning success – or lack thereof – abroad.

Future research should bring together the variables summarised in the previous paragraph to

provide a fuller characterisation of learners who succeed in improving their speaking skills abroad. As DeKeyser (2014: 321) points out, more fine-grained studies combining quantitative and qualitative data are needed to provide "a reliable documentation of background, process and outcome variables with the in-depth documentation of students' activities, and the quantity and quality of their interactions, especially from the students' perspective".

Intimately connected with the development of linguistic competence in general and of speaking skills in particular, Coleman's (2013) concentric circles model enables us to understand better the dynamic nature of socialization overseas. According to this model, learners gradually move outwards from the inner circle of co-nationals, through contact with other outsiders (generally other international students), towards the outer circle of locals. This progression, Coleman (2013: 31) argues, "is not universal, automatic or uni-directional, but given motivation, time and effort, alternatively labelled agency, movement tends to be centrifugal". He thus clearly pinpoints some of the ingredients that contribute to learners' successful oral development abroad. In their attempt to maximise their SA experience, however, students should not only invest time and effort in language learning but also they should ideally get institutional and pedagogical support in the form of preparatory pre-departure sessions, monitoring during the actual period abroad, and follow-up activities (Beattie 2014). In this sense, Kinginger (2011: 70) states that: "Every effort should be made to ensure that language learners abroad enjoy access to – and engagement in – the practices of their host communities as well as guidance in their efforts to learn and to interpret their experiences". In the same vein, Davidson (2010: 23) claims that residence abroad "holds enormous potential for meeting the needs of education in the 21st century". To unfold this potential, though, SA needs to be well integrated into the learners' curriculum and well supported by all the stakeholders involved, including of course learners themselves, who should engage actively in their learning process. In sum, as long as the relevant conditions we have outlined are met, SA *can* do the trick in terms of enhancing learners' speaking abilities.

## Notes

[1] These are longitudinal research projects – based at Universitat Pompeu Fabra (Barcelona) in collaboration with the Universitat de les Illes Balears (Palma) – that focus on the acquisition of English as a foreign language by students who partake in SA in addition to formal instruction at home. Learners' language development in these two learning contexts is analysed over a three-year period, contrasted against native-speaker baseline data, and interpreted with the help of qualitative data derived from learner questionnaires and diaries.

[2] In the case of the SALA-COLE studies which we have reported on, there is no inter-group variation in terms of such variables as motivation, proficiency or gender – among others – as the very same learners participated in both the at-home treatment and the SA period.

**Acknowledgements**

This research received financial support through HUM2007-66053-C02-01/02, FFI2010-21483-C02-01/02 and FFI2013-48640-C2-2-P from the Spanish Ministry of Economy and Competitiveness. I would like to express my gratitude to the anonymous reviewers of this paper for their suggestions. Thanks are also due to Dr Carmen Pérez-Vidal and the rest of my colleagues in the SALA and COLE projects.

## REFERENCES

**Aliaga-García, C. and Mora, J.C.** 2009. "Assessing the Effects of Phonetic Training on L2 Sound Perception and Production". In Watkins, M.A., A.S. Rauber and B.O. Baptista (Eds.) *Recent Research in Second Language Phonetics/Phonology*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing, 2-31.

**Allen, H.W.** 2002. "Does Study Abroad Make a Difference? An Investigation into Linguistic and Motivational Outcomes". PhD dissertation. Emory University, USA. In *Dissertation Abstracts International-A*, 63 (4), 1279.

**Allen, H.W.** 2013. "Self-regulatory Strategies of Foreign Language Learners". In Kinginger, C. (Ed.) *Social and Cultural Aspects of Language Learning in Study Abroad.* Amsterdam: John Benjamins, 3-15.

**Avello, P. and Lara, A.R.** 2014. "Phonological Development in L2 Speech Production during Study Abroad Programmes Differing in Length of Stay". In Pérez-Vidal, C. (Ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts.* Amsterdam: John Benjamins, 137-165.

**Beattie, J.** 2014. "The 'ins and outs' of a Study Abroad Programme". In Pérez-Vidal, C. (Ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts.* Amsterdam: John Benjamins, 59-84.

**Brecht, R.D. and Robinson, J.L.** 1995. "On the Value of Formal Instruction in Study Abroad". In Freed, B.F. (Ed.), *Second Language Acquisition in a Study Abroad Context*. Amsterdam: John Benjamins, 317-334.

**Churchill, E. and DuFon, M.A.** 2006. "Evolving Threads in Study Abroad Research". In DuFon, M.A. and E. Churchill. (Eds.) *Language Learners in Study Abroad Contexts*. Clevedon: Multilingual Matters, 1-27.

**Coleman, J.A.** 2013. "Researching Whole People and Whole Lives". In Kinginger, C. (Ed.) *Social and Cultural Aspects of Language Learning in Study Abroad.* Amsterdam: John Benjamins, 17-44.

**Davidson, D.** 2010. "Study abroad: When, how long, and with what results? New data from the Russian front". *Foreign Language Annals*, 43, 6-26.

**DeKeyser, R.M.** 1991. "Foreign Language Development during a Semester Abroad". In Freed, B. F. (Ed.) *Foreign Language Acquisition Research and the Classroom*. Lexington, MA: D. C. Heath & Co., 104-119.

**DeKeyser, R.M.** 2007. "Study Abroad as Foreign Language Practice". In DeKeyser, R.M. (Ed.). *Practice in a Second Language. Perspectives from Applied Linguistics and Cognitive Psychology*. Cambridge: Cambridge University Press, 208-226.

**DeKeyser, R.M.** 2014. "Research on Language Development during Study Abroad". In Pérez-Vidal, C. (Ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts.* Amsterdam: John Benjamins, 313-326.

**Del Río, C.** 2013. *Perceived Foreign Accent and Comprehensibility in the Oral Production of Adolescent Learners of English. Study Abroad vs. At Home Learning Contexts*. PhD dissertation. Universitat Pompeu Fabra.

**Díaz-Campos, M.** 2004. "Context of learning in the acquisition of Spanish second language phonology". *Studies in Second Language Acquisition*, 26 (2), 249-273.

**Dyson, P.** 1988. *The Year Abroad.* Report from the Central Bureau for Educational Visits and Exchanges, Oxford University Language Teaching Centre.

**Freed, B. (Ed.).** 1995. *Second Language Acquisition in a Study Abroad Context.* Amsterdam: John Benjamins.

**Freed, B., Segalowitz, N. and Dewey, D.** 2004. "Context of learning and second language fluency in French: Comparing regular classroom, study abroad and intensive domestic immersion programs". *Studies in Second Language Acquisition,* 26 (2), 275-301.

**Hernández, T.A.** 2010. "The relationship among motivation, interaction and the development of second language oral proficiency in a study-abroad context". *The Modern Language Journal*, 94 (4), 600-617.

**Hoffman-Hicks, S.D.** 1999. *The Longitudinal Development of French Foreign Language Pragmatic Competence: Evidence from Study Abroad Participants*. PhD dissertation. Indiana University.

**Howard, M.** 2001. "The Effects of Study Abroad on L2 Learners' Structural Skills". In Foster-Cohen, S. and A. Nizegorodcew (Eds.) *Eurosla Yearbook* 1. Amsterdam: John Benjamins, 123-141.

**Howard, M.** 2005. "Second Language Acquisition in a Study Abroad Context: A Comparative Investigation of the Effects of Study Abroad and Formal Language Instruction on the L2 Learner's grammatical development". In Housen, A. and M. Pierrard (Eds.) *Investigations in Instructed Second Language Acquisition*. Berlin: Mouton de Gruyter, 495-530.

**Isabelli-García, C.** 2003. "Development of oral communication skills abroad". *Frontiers: The Interdisciplinary Journal of Study Abroad*, 9, 149-73.

**Isabelli-García, C.** 2006. "Study Abroad Social Networks, Motivation and Attitudes: Implications for Second Language Acquisition". In DuFon, M.A. and E. Churchill (Eds.) *Language Learners in Study Abroad Contexts*. Clevedon: Multilingual Matters, 231-258.

**Isabelli-García, C.** 2010. "Acquisition of Spanish gender agreement in two learning contexts: Study abroad and at home". *Foreign Language Annals*, 43 (2), 289-303.

**Isabelli, C.A.** 2001. *The Impact of a Study Abroad Experience on the Acquisition of L2 Spanish Syntax: The Null Subject Parameter*. PhD dissertation. University of Illinois at Urbana-Champaign.

**Isabelli, C.A. and Nishida, C.** 2005. "Development of Spanish Subjunctive in a Nine-month Study-abroad Setting". In Edington, D. (Ed.) *Selected Proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*. Somerville MA: Cascadilla Press, 78-91.

**Jensen, J. and Howard. M.** 2014. "The effects of time in the development of complexity and accuracy during study abroad". *Eurosla Yearbook*, 14, 31-64.

**Juan-Garau, M.** 2012. "Context Matters: Variability across Three SLA Learning Settings". In E. Piechurska-Kuciel and L. Piasecka (Eds.) *Variability and Stability in Foreign and Second Language Learning Contexts*. Cambridge: Cambridge Scholars Publishing, 220-242.

**Juan-Garau, M.** 2014. "Oral Accuracy Growth after Formal Instruction and Study Abroad: Onset Level, Contact Factors and Long-Term Effects". In C. Pérez-Vidal (Ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins, 87-110.

**Juan-Garau, M. and Pérez-Vidal, C.** 2007. "The effect of context and contact on oral performance in students who go on a stay abroad". *Vigo International Journal of Applied Linguistics*, 4, 117-134.

**Kang, D-M.** 2014. "The effects of study-abroad experiences on EFL learners' willingness to communicate, speaking abilities, and participation in classroom interaction". *System*, 42, 319-332.

**Kinginger, C.** 2011. "Enhancing language learning in study abroad". *Annual Review of Applied Linguistics*, 31, 58-73.

**Kinginger, C.** 2013. "Introduction: Social and Cultural Aspects of Language Learning in Study Abroad". In Kinginger, C. (Ed.) *Social and Cultural Aspects of Language Learning in Study Abroad.* Amsterdam: John Benjamins, 3-15.

**Lafford, B.** 2006. "The Effects of Study Abroad vs. Classroom Contexts of Spanish SLA: Old Assumptions, New Insights and Future Research Directions". In Klee, C. and T. Face (Eds.) *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*. Somerville MA: Cascadilla Proceedings Project, 1–25.

**Lapkin, S., Hart, D. and Swain, M.** 1995. "A Canadian Interprovincial Exchange: Evaluating the Linguistic Impact of a Three-month Stay in Quebec". In Freed, B. (Ed.) *Second Language Acquisition in a Study Abroad Context*. Amsterdam: John Benjamins, 67-94.

**Lara, A.R.** 2014. *Complexity, Accuracy and Fluency Development through Study Abroad Programmes Varying in Duration*. PhD dissertation. Universitat Pompeu Fabra.

**Lindseth, M.U.** 2010. "The development of oral proficiency during a semester in Germany". *Foreign Language Annals*, 43 (2), 246-68.

**Llanes, A.** 2010. *Children and Adults Learning English in a Study Abroad Context*. PhD dissertation. Universitat de Barcelona.

**Llanes, A. and Muñoz, C.** 2009. "A short stay abroad: Does it make a difference?" *System*, 37, 353-365.

**Llanes, A. and Muñoz, C.** 2013. "Age effects in a study abroad context: Children and adults studying abroad and at home". *Language Learning*, 63 (1), 63-90.

**Longcope, P.D.** 2003. *What is the Impact of SA on L2 Learning? A Descriptive Study of Contexts, Conditions and Outcomes*. PhD dissertation, University of Pennsylvania.

**Lopez Ortega, N.R.** 2003. *The Development of Discourse Competence in Study Abroad Learners: A Study of Subject Expression in Spanish as a Second Language.* PhD dissertation. Cornell University.

**MacIntyre, P.D.** (2007). "Willingness to communicate in the second language: Understanding the decision to speak as a volitional process". *The Modern Language Journal*, 91 (4), 564-576.

**Meara, P.** 1994. "The year abroad and its effects". *Language Learning Journal*, 10, 32-38.

**Mora, J.C.** 2008. "Learning context effects on the acquisition of a second language phonology". In Pérez-Vidal, C., M. Juan-Garau and A. Bel (Eds.) *A Portrait of the Young in New Multilingual Spain*, 241-263.

**Mora, J.C.** 2014. "The Role of Onset Level on L2 Perceptual Phonological Development after Formal Instruction and Study Abroad". In Pérez-Vidal, C. (Ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts.* Amsterdam: John Benjamins, 167-194.

**Mora, J.C. and Valls-Ferrer, M.** 2012. "Oral fluency, accuracy and complexity in formal instruction and study abroad learning contexts". *TESOL Quarterly*, 46 (4), 610-641.

**Ortega, L.** 2005. "What do Learners Plan? Learner-driven Attention to Form during Pre-task Planning". In Ellis, R. (Ed.) *Planning and Task Performance in a Second Language.* Philadelphia, PA: John Benjamins, 77-110.

**Paige, R.M., Cohen, A.D., Kappler, B., Chi, J.C. and Lassegard, J.P.** 2002. *Maximizing Study Abroad: A Program Professionals' Guide to Strategies for Language and Culture Learning and Use.* Minneapolis, MN: Center for Advanced Research on Language Acquisition, University of Minnesota.

**Pérez-Vidal, C.** 2011. "Language Acquisition in Three Different Contexts of Learning: Formal Instruction, Stay Abroad, and Semi-immersion (CLIL)". In Ruiz de Zarobe, Y., J. M. Sierra and F. Gallardo (Eds.) *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts*. Bern: Peter Lang, 103-127.

**Pérez-Vidal, C.** 2014. "Study Abroad and Formal Instruction Contrasted: The SALA Project". In Pérez-Vidal, C. (Ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts.* Amsterdam: John Benjamins, 17-57.

**Pérez-Vidal, C. and Juan-Garau, M.** 2011. "The Effect of Context and Input Conditions on Oral and Written Development: A Study Abroad Perspective". *International Review of Applied Linguistics in Language Teaching*, 49 (2), 157-185.

**Pérez-Vidal, C., Juan-Garau, M. and Mora, J.C.** 2011. "The Effects of Formal Instruction and Study Abroad Contexts on Foreign Language Development: The SALA Project". In C. Sanz and R.P. Leow (Eds.) *Implicit and Explicit Language Learning. Conditions, Processes and Knowledge in SLA and Bilingualism.* Washington DC: Georgetown University Press, 115-128.

**Segalowitz, N. and Freed, B.** 2004. "Context, contact and cognition in oral fluency acquisition". *Studies in Second Language Acquisition*, 26 (2), 173-199.

**Serrano, R., Llanes, A. and Tragant, E.** 2011. "Analyzing the effect of context of second language learning: Domestic intensive and semi-intensive courses versus study abroad in Europe". *System*, 39 (2), 133-143.

**Simões, A.R.M.** 1996. "Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish". *Hispania*, 79 (1), 87-95.

**Stevens, J.J.** 2001. *The Acquisition of L2 Spanish Pronunciation in a Study Abroad Context*. PhD dissertation. University of Southern California.

**Torres, J.P.** 2003. *A Cognitive Approach to the Acquisition of Clitics in Spanish: Insights from Study Abroad and Classroom Learners*. PhD dissertation. Cornell University.

**Towell, R., Hawkins, R. and Bazergui, N.** 1996. "The development of fluency in advanced learners of French". *Applied Linguistics*, 17 (1), 84-119.

**Trenchs-Parera, M.** 2009. "Effects of formal instruction and a stay abroad on the acquisition of native-like oral fluency". *The Canadian Modern Language Review*, 65 (3), 365–393.

**Trenchs-Parera, M. and Juan-Garau, M.** 2014. "A Longitudinal Study of Learners' Motivation and Beliefs in at Home and Study Abroad Contexts". In C. Pérez Vidal (Ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins, 259-282.

**Valls-Ferrer, M. and Mora, J.C.** 2014. "L2 Fluency Development in Formal Instruction and Study Abroad. The Role of Initial Fluency Level and Language Contact". In Pérez-Vidal, C. (Ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts.* Amsterdam: John Benjamins, 111-136.

**Wang, C.** 2010. "Toward a second language socialization perspective: Issues in study abroad research". *Foreign Language Annals*, 43 (1), 50-63.

**Yager, K.** 1998. "Learning Spanish in Mexico: The effect of informal contact and student attitudes on language gain". *Hispania*, 81, 898-913.

**Yashima, T., Zenuk-Nishide, L. and Shimizu, K.** 2004. "The influence of attitudes and affect on willingness to communicate and second language communication". *Language Learning*, 54 (1), 119-152.

*Received: 8 June 2015*

*Accepted: 11 October 2015*

*Cite this article as:*

**Juan-Garau, M.** 2015. "Speaking an additional language: Can study abroad do the trick?". *Language Value* 7 (1), 45-60. Jaume I University ePress: Castelló, Spain. http://www.e-revistes.uji.es/languagevalue. DOI**:** http://dx.doi.org/10.6035/LanguageV.2015.7.4

**ISSN 1989-7103**

*Language Value* November 2015, Volume 7, Number 1 pp. 62-91
http://www.e-revistes.uji.es/languagevalue ISSN 1989-7103

# Revisiting pragmatic tests in the FL context: Towards interactive tests to examine speech act performance[1]

**Vicent Beltrán Palanques**
vbeltran@uji.es
**Universitat Jaume I, Spain**

**ABSTRACT**

The aim of this study is to explore the task effect of two different research methods, namely those of interactive discourse completion tasks/tests (IDCTs) and role-play tasks (RPTs). The two research methods employed in this study adopt an interactive approach that allows participants to freely interact not only in the oral mode but also in the written mode. This paper compares the apology strategies elicited by means of IDCTs and in RPTs in terms of strategy length, amount of strategies and classification of strategies across the two elicitation techniques. Results from this comparison will be presented and discussed, and pedagogical implications suggested.

**Key words:** research methods, IDCTs, RPTs, apologies.

## I. INTRODUCTION

Testing pragmatics is a relatively young field of research within interlanguage pragmatics (ILP). Although different research methods are used to conduct ILP research (see Kasper and Roever, 2005), discourse completion tests/tasks (DCTs) and role-play tasks (RPTs) are typically used. Over the years, several authors have empirically examined the two aforementioned research methods so as to improve their effectiveness (Houck and Gass 1996; Sasaki 1998; Yuan, 2001; Martínez-Flor 2006, 2013; Duan 2008; Eslami-Rasekh 2012; Beltrán-Palanques 2013). Yet, there is still a need to further explore this particular field of research to shed more light on this specific issue. According to Kasper and Roever (2005), research methods in ILP can be classified as follows: (1) observational data of spoken interaction involving authentic discourse; elicited conversation and RPTs; (2) questionnaires as written DCTs and multiple choice

[1] The research conducted in this article is part of the Education and Innovation research project: *Proyecto de Innovación Educativa Universitat Jaume I 2779/13 Parámetros de aproximación a la evaluación de las destrezas orales en lengua inglesa: tipología, diseño de test y criterios de validación*.



DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.5

questionnaires; (3) rating scales; (4) oral and narrative forms of self-report; (5) diaries; and (6) verbal reports. Two of the most widely employed research methods in ILP are DCTs and RPTs.

### I.1. Discourse completion tasks/tests (DCTs)

DCTs involve a task that contains a description of a given situation (i.e. a prompt), and an empty space, intentionally left, in which the participant has to provide a specific utterance (Kasper and Roever 2005; Roever, 2010). Kasper (2000) distinguishes four main types that derive from the original version of the DCT, i.e. (1) the classic DCT; (2) the dialogue construction DCT; (3) the open item verbal response only; and (4) the open time free response. The classic DCT typically involves a first turn of a dialogue that serves as a stimulus, and a missing gap, which is also known as a rejoinder (Johnston et al. 1998). The second type, the dialogue construction, does not include the response of a hearer, and the gap may or may not be introduced by an interlocutor's turn. The third and fourth types do not require a construction of a dialogue. Specifically, in the third type, the open item verbal response only, involves a written verbal response, and in the fourth type, participants can provide verbal and non-verbal responses or even opt out. Additionally, other types of DCTs can also be found in the literature, such as the content-enriched DCT prompts proposed by Billmyer and Varghese (2000: 543), which might "elicit more robust external modification and elaboration than do the archetypal content-poor prompts which most DCTs studies to date have used". Other DCTs can be for instance, the free discourse completion tasks (Barron 2003), also known as the *dialogue production tasks* in pragmatic variation (Schneider 2008) in which two participants elaborate a dialogue, or the multiple-rejoinder DCT advanced by Cohen and Shively (2003) in which participants provide different responses over a dialogue. Furthermore, other research methods including visual aids have been developed such as the cartoon oral production tasks (Rose 2000) or interactive DCTs (IDCTs) (Martínez-Flor and Usó-Juan 2011; Beltrán-Palanques, 2013) that may also incorporate enhanced photos (Martínez-Flor and Usó-Juan 2011). Finally, multimedia DCTs have also been designed, such as the computer-assisted interactive DCTs (Kuha 1997), the multimedia elicitation task (Schauer 2004, 2009), which is a computer-based system, the

computerized extended DCTs with video prompts (Sydorenko et al. 2014), or the IDCTs performed by means of Google Docs (Beltrán-Palanques in press).

By means of DCTs, researchers may obtain a relatively large amount of data in a short period of time (Roever 2010). Nevertheless, the fact that DCTs can be administered in a short period of time does not necessarily imply that their use is not complex (Kasper and Roever 2005). DCTs are not without criticism; for example, Rose (1994) highlights the artificiality of those tests and Sasaki (1998) indicates that DCTs can be seen as a pen and paper method that resembles a typical written test. Considering these aspects, it could be to some extent questioned whether these tests might represent authentic discourse. Participants may be somehow affected by the nature of the mode in which they are asked to produce their utterances, and consequently this could have an influence on speech act production. In line with this, Golato (2003) argues that DCTs might not always provide an accurate depiction of what they would really say in an authentic oral interaction. Another aspect that may be questioned is the lack of interaction since, in most cases, DCTs only allow participants to take one turn. In an attempt to overcome this problem, some researchers have proposed different typologies of DCTs that include an interactive perspective, (e.g. Martínez-Flor and Usó-Juan, 2011; Beltrán-Palanques 2013, in press). Nevertheless, the design of the research method would depend on the purpose of the study. Despite the fact that some limitations can be identified, DCTs are widely used in the field of ILP in order to gather speech act data since, as pointed out by Kasper and Rose (2002), they show the different forms and strategies that participants employ when confronting a given situation. Furthermore, as pointed out by O'Keeffe et al. (2011: 23 their emphasis), "without this methodology, it would have been difficult if not impossible to conduct such research because some speech acts are very difficult to 'obtain' in any other way". Hence, although some limitations can be found, DCTs are typically employed in the field of ILP to gather speech act data, and their desing would to some extent depend on the purpose of the study.

### I.2. Role-play tasks (RPTs)

RPTs are also widely used to obtain speech act data in the field of ILP, especially in foreign language (FL) contexts where obtaining authentic discourse appears to be rather complex. The RPT method, as indicated by Crookall and Saunders (1989: 15—16, their emphasis) may be seen as "a social or human activity in which participants 'take' on and 'act out' specified 'roles', often within a predefined social framework or situational blueprint (a 'scenario')". In this particular type of research method, participants are encouraged to take part in specific scenarios, perform them orally and say what they would say in such concrete situations and circumstances (Crookall and Saunders 1989; Roever 2010). Moreover, as in the case of DCTs, RPTs may contain contextual information about the context in which the simulated social encounter might occur as well as about the relationship of the participants.

According to Kasper and Dahl (1991), two main types of role-plays depending on the level of interaction can be distinguished, namely those of closed or monologic and open or interactive. Closed RPTs involve the response of one participant to a particular situation without having the response of another interlocutor, whereas open RPTs provide participants with opportunities for interacting and elaborating different turns. Thus, open RPTs can involve as many turns and discourse phases as necessary since interlocutors can interact until the communicative purpose is achieved. It seems that there is a clear advantage of open RPTs over closed RPTs because participants can communicate somehow freely and researchers can then gather further features of spoken production due to its interactive nature. However, as in the case of DCTs, their design would depend on the purposes of the study.

It is also worth mentioning that it remains uncertain whether by means of RPTs an accurate representation of authentic language use in real contexts could be provided (Kasper 2000). It is suggested that data derived from RPTs might somehow differ from that found in natural contexts (Kasper 2000; Golato 2003). Roever (2010) also indicates that natural data may differ from RPTs in the sense that RPTs are simulated scenarios and participants are aware of that fact. Then, participants might not necessarily produce

the same utterances that they would in a real social interaction where their interaction may have an impact on real life (Roever 2010). That is to say, when performing a RPT, participants act out a given role in a simulated context rather than in a natural setting, and consequently it is not clear whether the discourse created might reflect appropriately the linguistic behaviour that might be at work in natural encounters. Another drawback that could be identified is the lack of visual information, as in most cases, researchers use audio data instead of video data. Hence, nonverbal pragmatics is not captured and consequently researchers might not gather all the information that is present in a social interaction. Despite the fact that some limitations can be identified, RPTs are still used in the field of ILP to gather oral speech act data as they allow researchers to obtain specific spoken data about the particular pragmatic aspects investigated.

## II. **WRITTEN AND SPOKEN DATA: AN OVERVIEW**

Over the last decades, different researchers have carried out different studies contrasting and comparing the speech act outcomes of DCTs and RPTs. Houck and Gass (1996) examined the use of refusals as a response to other speech acts (i.e. suggestions, offers, invitations, and requests) by Japanese learners of English as a second language (SL) in both videotaped open RPTs and written DCTs. Results showed that the data obtained in the RPTs was greater since participants employed longer responses and wider use of negotiation segments than those in the written DCTs. Moreover, lower varieties of speech act realisations were found in the written data. Contrarily, Rintell and Mitchell (1989) did not find significant differences in the responses obtained by means of written DCTs and closed RPTs concerning response type. The authors compared the responses obtained from both closed RPTs and written DCTs as regards the speech acts of requests and apologies. In this study, participants were made up of both language learners of English as SL and native speakers. According to Rintell and Mitchell (1989), findings might have also been affected by the fact that closed RPTs, due to its nature, did not allow interaction between participants. Still, the responses given in the oral tasks were longer. Similar results concerning the content of the semantic formulae were found in the study conducted by Eisenstein and Bodman's (1993). The study focused on how

the speech act of gratitude was expressed by both native speakers and non-native speakers by means of four different research methods, i.e. natural observation, oral DCTs, written DCTs, and RPTs. In comparing the four methods, results showed similar responses regarding the content of the semantic formulae. The main difference, however, lied on the level of interaction of the research methods. Also, when comparing the two types of DCTs, the oral DCT version allowed participants to take more turns than the written DCT, and consequently, longer responses were produced.

Sasaki (1998) compared the production of requests and refusals elicited by a group of Japanese students by means of written DCTs and closed RPTs. In contrasting the data obtained from the two research methods, results demonstrated that the responses differed in terms of length and content. Specifically, the responses elicited in the oral method were longer and they contained more and higher variety of semantic formulae than those in the written method. The author argued that the difference found concerning length could be related to the fact that in the oral task participants employed features of spoken language such as repetitions and hesitations. However, the types of central speech act expressions elicited in both research methods were similar. Yuan (2001) examined the production of compliment and compliment responses in various research methods, i.e. written DCTs, oral DCTs, field notes and natural conversations. Results showed that in providing the participants with only one turn in the written DCTs and oral DCTs, interaction might not take place as in the case of RPTs and natural conversations. Results also revealed that the responses in the oral DCT offered more features of natural language than those found in written DCTs. Similar results were found by Martínez-Flor (2006), who focused on the speech act of suggestions. More specifically, the author examined the task effect on two types of production methods, i.e. phone messages and emails. Results indicated that a large amount of semantic formulae were found in written DCTs. In this respect, the author stated that these results might have been affected by the fact that the oral production tasks employed in the study did only allow participants to elicit more than one turn since the type of oral task was similar to closed RPTs. Furthermore, Martínez-Flor (2006) also reported that the responses found in the written task were longer and further elaborated that those appearing in the oral task. Concerning this, the author pointed out that this might have

been related to the fact that the written task was not a conventional pen-and-paper task since this particular task was developed via email format and that they had more time to think about their responses.

In another FL context, Duan (2008) explored the use of refusal strategies by Chinese learners of English as a FL as a response to four different situations, namely those of invitations, suggestions, offers and requests, in two different research methods, written DCTs and oral RPTs. Results revealed that no significant differences were found in terms of strategies when comparing the two research methods. Nevertheless, the author pointed out that the written DCTs appeared to show longer sentences, whereas the oral RPTs produced more natural expressions since further features of authentic spoken language such as pause fillers and broken sentences were identified. More recently, Eslami-Rasekh (2012) examined data taken from written DCTs and closed RPTs. In this particular case, the author focused on the requests strategies produced by a group of Iranian university students in their first language (L1) (i.e. Persian). Findings demonstrated that longer responses were found in the oral data, which was related to the fact that longer and greater number of alters and supportive moves were used in this the RPTs. Concerning the variety of strategies, results indicated that no differences between the two research methods were noted. However, in the written DCTs, more direct realisations were identified. Besides, the modification devices that appeared in the oral method had a softer tone, and concerning the request perspective, findings showed that the data found in the written DCTs were more hearer-oriented, whereas the oral data presented a more impersonal or collective referent. The author, then, concluded that the data obtained by means of RPTs could provide a better representation of natural speech than that gathered through written DCTs. Martínez-Flor (2013), in the Spanish context, conducted a study to investigate the task effect in learners' production of refusal strategies. In this study two different production methods purposefully designed following an interactive perspective were used: interactive written DCTs and open RPTs. The focus of this study was to compare the outcomes of the two research methods to examine the refusal responses in terms of length, amount and typology. Results demonstrated that the amount, the length and the type of refusal strategies employed were to some extent similar in both research methods. In light of such results,

Martínez-Flor (2013) indicated that the adoption of an interactive approach in the design of the written DCTs, thereby similarly to RPTs, seemed to have positively affected participants' responses since they could negotiate meaning and use different turns, which somehow promoted the elicitation of a rich variety of refusal strategies. Finally, Beltrán-Palanques (2013) conducted a study in the Spanish context in which the speech act under investigation was that of apologies. In this particular study, following Martínez-Flor's (2013) study, an interactive approach was followed when designing both the DCTs (i.e. IDCTs) and the open RPTs. Results from this study were in line with the study conducted by Martínez-Flor (2013) since the quantity and quality of the speech act realisation of apologies were similar across the two research methods due to their interactive nature.

Considering the above sketched literature review, the purpose of this study is to examine the effect of two different research methods, IDCTs and open RPTs, on participants' pragmalinguistic apology sequences. Research questions guiding this study are the following:

- Will the data collected by means of open RPTs and IDCTs elicit similar results as regards length?
- Will the data collected by means of open RPTs and IDCTs elicit similar results as regards amount of strategies used?
- Will the data collected by means of open RPTs and IDCTs elicit similar distribution of strategies?

## III. **METHODOLOGY**

### **III.1. Participants**

This study involved 16 female adult learners whose average age was 22.5. All of them were graduate students and they were studying English as a FL, more specifically a B2.1 level course. Despite the fact that they were studying a B2.1 level course of English, the Quick Placement Test (2001) published by Oxford University Press was employed to verify their proficiency level. Results showed that participants were closer

to the B2 level although they had still not achieved this specific level. The background questionnaire, adapted from Beltrán-Palanques (2013), showed that participants were all bilinguals (Catalan and Spanish). They have been to an English speaking country for less than 2 weeks and exclusively for tourism (e.g. holidays, visiting friends). Furthermore, half the participants had been studying English not only at school, higher school, and university, but also in some private schools. Participants were arranged in pairs so as to outperform the written and oral tasks, and then only 8 out of 16 played the role of apologiser while the remaining 8 participants were asked to produce other speech acts so as to interact in the given scenarios. Specifically, they performed mainly complaints and requests. Nevertheless, this is beyond the scope of the present study and therefore no attention will be paid to this particular issue.

### III.2. Pragmatic aspect under investigation

According to Austin's (1962) classification of illocutionary acts, apologies fall into the category of behabitives, and Searle (1979) assigns this particular speech act to the category of expressives. Searle (1979: 15) indicates that apologies "express the psychological state specified in the sincerity condition about a state of affairs specified in the propositional content". Leech (1983), however, classifies this particular speech act within the convivial speech act type since its illocutionary goal coincides with the social goal, specifically, that of maintaining harmony between the speaker and the hearer in which there is some benefit for the hearer and some cost for the speaker. Aijmer (1996) indicates that apologies are strategies that are used to convey a particular communicative goal, which requires an utterance which purpose is to "set things right (Olshtain and Cohen 1983: 20)", and they are used in situations in which a speaker commits an action that damages another person.

Therefore, at least two participants need to take part in an apology sequence, the offender or hearer and the offended or hearer. Apologies are moves that are typically employed to solve a problem between at least two participants, i.e. the speaker and the hearer, and restore harmony between them. Hence, an apology situation involves a participant uttering an apology that is addressed towards the offended participant in

order to restore problems as well as to re-establish harmony between them (Holmes 1995). By apologising, the speaker seems to understand the situation and accepts that an error has been committed. Moreover, it could be suggested that in a real situation, the speaker might apologise and negotiate the apology with the hearer, and the hearer may accept or reject the speaker' apology/apologies. In a situation in which the apology/apologies is/are accepted, participants might restore to some extent the harmony between them. In addition to this, there are also some other factors that could influence offenders' assumption of responsibility (Olshtain and Cohen 1983). For example, the perception of the degree of the severity of the offense may play an important role. The speaker might not necessarily see a violation of a social norm or an inappropriate act in his/her behaviour (Olshtain and Cohen 1983), or perhaps the speaker might choose to emphasise his or her innocence (Trosborg 1987), and then, not take any responsibility. Furthermore, other influential factors can also play a crucial role in the apology sequence such as age, degree of social distance and power between the participants.

It seems therefore that the speech act of apologies is a rather complex speech act that involves different factors such as understanding that an error has been committed. This could happen in real-life interactions, since the speaker might decide not to apologise, as he/she does not perceive that damage has been caused, or simply because the speaker does not want to apologise. In line with this, it could also be argued that personality traits, the real relationship between/among participants, as well as the possible consequences, or at least, the perceived consequences, could also affect the speakers' decision. In a simulated situation, however, participants might be asked to for example apologise in a given situation, and therefore, they would accomplish the task as required, although it is difficult to know whether he/she would act in the same way in a real interaction.

### III.3. Research methods

The research methods used in this study were open RPTs and IDCTs. These two research methods were purposefully designed following an interactive perspective in

order to allow participants to freely interact in the two tasks. For the purposes of the present study, the research methods used (See Appendix A for the open RPTs and Appendix B for the IDCTs) were the same used by Beltrán-Palanques (2013). The two elicitation techniques included a description of the roles and the contexts of each situation in order to help participants understand each scenario appropriately. The same scenarios were used in the RPTs and IDCTs in order to compare the task effect. The scenarios were designed taking into account the target group of the study, participants' sociocultural context, the setting in which they are administered (Beltrán-Palanques 2013), participants' familiarity with the roles (Trosborg 1995) and with the context (Hudson et al. 1995).

The open RPTs (see Appendix A) and the interactive written IDCTs (see Appendix B) used in this study consist of eight situations which were classified as occurring in the following contexts, namely those of university (i.e. situations 2, 5, 6 and 8), bookshop (i.e. situation 1), students' flat (i.e. situation 3), language school (i.e. situation 4), and theatre (i.e. situation 7). Furthermore, the situations were designed taking into account the sociopragmatic features of social status (i.e. hear-dominant and equal), social distance (i.e. acquaintance and stranger) and severity of offence (i.e. high and low) (Brown and Levinson 1987). Finally, it is also important to mention that some of the situations that were adapted from the studies conducted by from Afghari, (2007), Nureddeen (2008) and Flores-Salgado (2011).

### III.4. Procedure and data analysis

Data for this study was collected by means of open RPTs and IDCTs. Due to the interactive nature of the two research methods, participants were randomly assigned in pairs, and each pair performed both tasks, first the RPTs and then the DCTs. Data was collected in two different phases, the first phase was conducted in the second week of the course, and the second phase during the third week of the course. In both cases, the tasks were conducted in the room where the instruction used to take place. Participants were distributed into different time slots, so during the completion of the tasks, the two participants and the teacher - who is the researcher of this study - were alone in the

room. It is also important to note that in order to avoid time constrains, ample time was provided to perform the different tasks. Then, the two tasks were performed at the very beginning of the course, i.e. the second and the third week, and participants did not receive pragmatic instruction regarding the speech act of apologies in the FL course before administrating the tests. After conducting the study, participants were provided with instruction since the results of the study were also used for pedagogical purposes.

In order to classify the different apology strategies, a taxonomy advanced by Beltrán-Palanques (2013) based on previous research on apologies (Olshtain and Cohen 1983; Blum-Kulka, et al. 1989; Trosborg 1995) was employed. As shown in Table 1, the taxonomy is divided into three main parts, namely those of "Illocutionary Force Indicating Device (IFID)", "Assuming Responsibility" and "Compensating the Other".

| **Strategy** | **Type** | **Example** |
|---|---|---|
| **IFID** | | |
| Apologising | Offer an apology | *I apologise* |
| | Feel obliged to apologise | *I should/must apologise* |
| | Regret | *I'm sorry* |
| | Request for forgiveness | *Please, forgive me* |
| **Assuming Responsibility** | | |
| Accepting the situation | Acceptance | *You're (completely) right; I (really have to) accept/ admit it/ that; I (totally/ really) screw it up* |
| Showing empathy | Understanding the error | *I see what you mean...;*<br>*I see/ understand your point of view* |
| | Showing feelings | *I feel bad about what happened; I feel awful/ bad/ guilty* |
| Lack of intention | | *I didn't mean to do that/ hurt you; It wasn't my intention* |
| Justifying the situation | Internal | *I couldn't come earlier.../ I couldn't make it.* |
| | External | *It was raining a lot; my car broke down; there was a problem..; I couldn't call/ text you to tell you that* |
| **Compensating the Other** | | |

| | | |
|---|---|---|
| Promise of forbearance | Forbearance | *It won't happen again* |
| Repair the situation | Offering | *I will/ can/ could do it for you.* |
| | Requesting | *I would like to know how I can compensate you.* |
| | Suggesting | *Why don't we get started?* |

Table 1. Taxonomy on the speech act of apologies (Beltrán-Palanques 2013: 47-48)

This taxonomy is concurred with most apology classifications (Olshtain and Cohen 1983; Blum-Kulka et al. 1989; Trosborg 1995) which also include direct semantic strategies to offer an apology, speaker's assumption of the responsibility, and a set of semantic strategies that might show speaker's concern towards the error and/or the mistake, as well as some strategies that can be employed to compensate the hearer. In short, the present taxonomy includes the basic strategies that may be used when performing an apology and they can be combined to better express the communicative goal. Nevertheless, it is worth mentioning that further categories of strategies could be identified or other taxonomies may classify apology strategies in a different manner.

## IV. RESULTS AND DISCUSSION

This section presents the results and discussion for the three research questions that guided this study. More specifically, the research questions of this study focused on the length, amount and distribution of apology strategies across the two research methods. The results were analysed using the Paired T-test. Table 2 shows the results as regards length and amount of apology strategies.

| *N* 8 | IDCT Length | RPT Length | IDCT Amount | RPT Amount |
|---|---|---|---|---|
| Mean | 855.50 | 873.88 | 48.50 | 51.25 |
| Standard deviation | 27.553 | 29.240 | 5.372 | 4.979 |
| t | 3.973 | | 5.227 | |
| Sig. Bilateral | .005 | | .001 | |

Table 2. Results as regards length and amount across IDCTs and RPTs.

The first research question focused on whether data collected by means of open RPTs and IDCTs would elicit similar results as regards length. Results revealed that the mean of the length in the IDCTs was 855.50 whereas that of RPTs was 873.88. It was found that the standard deviation of the IDCTs was 27.553 and that of RPTs was 29.240. The $t_8$ value found was 3.973 and the Sig. bilateral was 0.005, i.e. p value = 0.005 ≤ α 0.05, revealing that the research method employed the data collection procedure affected the length of the data elicited. This would imply, in this specific study, that despite the fact that the two research methods followed an interactive perspective differences across the two instruments were found, being the responses elicited in the oral mode longer than those found in the written data. The second research question in this study was concerned with whether the data collected by means of open RPTs and IDCTs would elicit similar results concerning the amount of strategies. Results showed that the mean found as regards the amount of strategies elicited in the IDCTs was 48.50 while in the RPTs was 51.25. The standard deviation in the case of the IDCTs was 51.25 and 4.979 in the RPTs. The t8 value obtained was 5.227 and the Sig. bilateral was .001, i.e. i.e. p value = 0.001 ≤ α 0.05. It was therefore identified that the difference was statistically significant. Dissimilar amount of semantic formulae was found when comparing the data derived from the IDCTs and the open RPTs. These results were similar to the findings concerning the first research question. That is to say, regardless of participants' opportunities for interaction, the research method used seemed to have affected the results obtained since there were statistical differences. Results seemed to suggest that that amount of strategies elicited in the RPTs were slightly greater than in the IDCTs.

Finally, the third research question focused on the distribution of the strategies across the two research methods. In order to examine the distribution of the strategies, the taxonomy presented by Beltrán-Palanques (2013) was used. To examine this aspect, a Paired T-test was applied taking into account the three main categories identified in the taxonomy used for the purposes of this study, i.e. IFID, Assuming Responsibility and Compensating the Other. Table 3 shows the results as regards the distribution of the strategies across the two research methods.

| *N*<br>3 | IDCT<br>Distribution | RPT<br>Distribution |
|---|---|---|
| Mean | 89.33 | 84.67 |
| Standard deviation | 11.590 | 7.767 |
| t | 2.000 | |
| Sig. Bilateral | .184 | |

Table 3. Results as regards the ddistribution of apology strategies across the two research methods.

In this case, the purpose was to explore the distribution of the strategies considering the taxonomy advanced, which was divided into three main parts, to whether participants' apologies distribution across the two research methods were consistent. The mean found concerning the amount of strategies elicited in the IDCTs was 89.33 whereas in the RPTs was 84.67. The IDCTs showed a standard deviation of 11.590 and the RPTs of 7.767. The $t_3$ value obtained was 2.000 and the Sig. bilateral was .184, being the p value set at $\alpha$ 0.05. It was therefore concluded that the results was not statistically significant. This finding could be related to the fact that both research methods offered participants with similar opportunities for performing their tasks and that both followed an interactive approach. Interestingly, the results found as regards the third research question revealed that there were no differences as regards the distribution of the strategies, but there were differences concerning the length of strategies and amount of strategies, being greater in both cases in the RPTs. This could be related to the fact that in the oral mode, participants tended to produce longer strategies and more strategies to better convey their communicative purpose and make themselves understood. Also, in the oral mode, participants might, for example, employ features of spoken language such as repetitions and hesitations (Sasaki 1998), however, in this study, oral features such as repetitions (e.g. I, I), or hesitation were not considered.

These findings appear to partially contradict previous research in which interactive research methods were employed (Beltrán-Palanques 2013; Martínez-Flor 2013). Particularly, the studies conducted by Beltrán-Palanques (2013) and Martínez-Flor (2013) focused also on the length, amount and distribution of two different speech acts, apologies and requests, respectively. In those studies statistical differences across the

two research methods were not identified. Moreover, in both studies, the research methods were characterised for following an interactive approach. The differences as regards length and amount could be related to the fact that in the spoken mode participants might have more opportunities to hesitate, repeat themselves (Sasaki, 1998), as well as to use false starts and other characteristics of the spoken discourse which are not present in the written mode. Nevertheless, in this study, those aspects were not taken into account as they could influence the results. Regarding the distribution of strategies, the results of this study were consistent with previous research in which no statistical differences were found (Beltrán-Palanques, 2013; Martínez-Flor, 2013). However, results showed that regardless of having a similar distribution of strategies, the length and the amount of strategies elicited was similar, possibly related to the fact that participants tended to use longer and a greater quantity of strategies to research their communicative purpose in the oral mode.

In light of the results obtained, it could be suggested that the level of interaction of the research methods would not be the only aspect to take into account when designing research methods. Other aspects such as participants' ability to perform each task as well as personality traits and participants' psychological conditions when taking the tests could also play a role in the completion of tasks. Then, this may imply that perhaps there are further aspects that could affect participants' speech act elicitation, that is to say, external factors to the actual research method. Furthermore, despite the effort made to provide an interactive written instrument, differences between acting out in the spoken mode and written mode might appear which in turn could be logical since there are differences between written and spoken discourses. Finally, it is important to state that this study could have been enriched by means of retrospective verbal reports so as to further obtain insights into participants' thoughts, but unfortunately, due to time constraints, they were not included for the purposes of this study (see for a review Félix-Brasdefer, 2010; Beltrán-Palanques, 2014), which in this case would have been beneficial to examine the differences found. The main limitations of this study are the number of participants as well as having only female participants, and thereby these limitations should be taken into account for further research.

## V. CONCLUSION

The purpose of this study was to compare two different research methods, specifically those of RPTs and IDCTs. The two research methods, which were previously used by Beltrán-Palanques (2013), followed an interactive perspective. It was believed that in order to appropriately compare the outcomes of the two research methods, participants should be given with the same opportunities for interaction. In this specific case, the purpose was to explore whether the interactive nature of the traditional DCT could have an effect on participants' elicitation of apologies. Further conditions were taken into account, for example, the same scenarios were used in both research methods and they were completed in pairs; being in both cases the same pair of participants. This study showed that albeit the two research methods were valid to elicit the speech act data of apologies in specific controlled situations from an interactive perspective, significant differences were found as regards the length of utterances and amount of apology strategies. However, no differences were found concerning the amount of strategies as well as the typology.

Finally, it is worth adding that from the study presented above, some future research and pedagogical suggestions can be provided. On the one hand, it seems that it would be necessary to further develop research methods that foster interaction among participants, if the aim is that of capturing interaction, as well as examine data taking into account the whole discourse, that is to say, analysing not only the speech act investigated, but also all the different speech acts appearing the interaction, as well as how participants negotiate communicate, and make themselves understood over the different turns. Ideally, this type of research methods should be used in combination with verbal reports in order to gather further data (Félix-Brasdefer 2010; Beltrán-Palanques 2014). On the other hand, these two research methods can also be used in the language classroom to assess pragmatic competence as part of a process of instruction of speech acts. Moreover, these research methods could also be used as diagnostic tests before teaching specific speech acts as they could provide teachers with insights about learners' pragmatic knowledge. Hence, the pedagogical value of these research methods

should not be ignored as they can serve not only to assess pragmatic knowledge in a final stage of instruction, but also as a diagnostic test.

## REFERENCES

**Afghari, A.** 2007. "A sociopragmatic study of apology speech act realization patterns in Persian". *Speech Communication*, 49, 177-185.

**Aijmer, K.** 1996. *Conversational Routines in English*. London: Longman.

**Austin, J.** 1962. *How to do Things with Words.* Cambridge MA: Harvard University Press.

**Barron, A.** 2003. *Acquisition in Interlanguage Pragmatics: Learning How to Do Things with Words.* Amsterdam: John Benjamins Publishing Company.

**Beltrán-Palanques, V.** (2013). *Exploring Research Methods in Interlanguage Pragmatics: A Study Based on Apologies.* Saarbrücken: Lap Lambert Academic Publishing GmbH KG.

**Beltrán-Palanques, V.** (2014). "Methodological Issues in Interlanguage Pragmatics: Some Food for Thought". In Alcaraz-Mármol, G. and M. M. Jiménez-Cervantes (Eds.) *Studies in Philology: Linguistics, Literature and Cultural Studies in Modern Languages*. Newcastle upon Tyne: Cambridge Scholars Publishing, 239-253.

**Beltrán-Palanques, V.** (in press). "Synchronous Computer-Mediated Communication in ILP Research: A Study Based on the ESP Context". In Martín-Monje, E., I. Elorza and B. García-Riaza (Eds.) *Technology-Enhanced Language Learning for Specialized Domains: Practical Applications and Mobility.* London: Routledge.

**Billmyer , K., and Varghese, M.** 2000. "Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests". *Applied Linguistics*, 21 (4), 517-552.

**Blum-Kulka, S., House, J., and Kasper, G.** 1989. *Cross-Cultural Pragmatics: Requests and Apologies.* Norwood, NJ: Ablex.

**Brown and Levinson.**1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

**Cohen, A., and Shively, R.** 2003. "Measuring speech acts with multiple rejoinder DCTs". *Language Testing Update*, 32, 39-42.

**Crookall, D., and Saunders, D.** 1989. "Towards an integration of communication and simulation". In Crookall, D. and D. Saunders (Eds.) *Communication and*

*Simulation: From Two Fields to One Theme.* Clevedon: Multilingual Matters, 3-29.

**Duan, A. W.** 2008. "The comparison between written DCT and oral role-plays in investigation upon English refusal strategies by Chinese EFL students". *Sino-US English Teaching*, 5 (10), 8-18.

**Eisenstein, M., and Bodman, J.** 1993. "Expressing gratitude in English." In Kasper, G. and S. Blum-Kulka (Eds.) *Interlanguage Pragmatics*.Oxford: Oxford University Press, 64-81.

**Eslami-Rasekh, A.** 2012. "Eliciting Persian requests: DCT and role play data". *World Journal of Education*, 2 (3), 80-86.

**Félix-Brasdefer, J. C.** 2010. "Data collection methods in speech act performance". In Martínez-Flor, A. and E. Usó-Juan (Eds.) *Speech Act Performance. Theoretical, Empirical and Methodological Issues*. Amsterdam: John Benjamins Publishing Company, 41-56.

**Flores-Salgado, E.** 2011. *The Pragmatics of Requests and Apologies. Developmental Patterns of Mexican* Students. John Benjamins Publishing Company.

**Golato, A.** 2003. "Studying compliment responses: A comparison of DCTs and recording of naturally occurring talk". *Applied Linguistics*, 21 (1), 90-121.

**Holmes, J.** 1995. *Women, Men and Politeness.* New York: Longman.

**Houck, N., and Gass, S.** 1996. "Non-native refusal: A methodological perspective". In Gass, S. and J. Neu (Eds.) *Speech Acts Across Cultures*. Berlin: Mouton de Gruyter, 45-64.

**Hudson, T., Detmer, E., and Brown, J.** 1995. *Developing Prototypic Measures of Cross-Cultural Pragmatics (Technical Report, 7).* Honolulu: University of Hawai'i .

**Johnston, B., Kasper, G., and Ross, S.** 1998. "Effect of rejoinders in production questionnaires". *Applied Linguistics*, 19 (2), 157-182.

**Kasper, G**. 2000. "Data collection in pragmatics research". In Spencer-Oatey, H. (Ed.), *Culturally Speaking. Managing Rapport through Talk across Cultures*. London: A&C Black, 316-341.

**Kasper, G. and Dahl, M.** 1991. "Research methods in interlanguage pragmatics". *Studies in Second Language Acquisition*, 13 (2), 215-247.

**Kasper, G. and Roever, C.** 2005. "Pragmatics in Second Language Learning". In Hinkel, E. *Handbook of research in second language learning and teaching*. Mahwah, NJ: Erlbaum, 317-334.

**Kasper, G. and Rose, K.** *Pragmatic Development in a Second Language (Language Learning Monographs Series).* Oxford: Blackwell.

**Kuha, M.** 1997. "The Computer-Assisted Interactive DCT: A Study in Pragmatics Research Methodology". In Bouton, L. F. (Ed.), *Pragmatics and Language Learning (Vol. 8)*. Urbana-Champaign: Division of English as an International Language, University of Illinois, 99-127.

**Leech, G.** 1983. *Principles of Pragmatics.* London: Longman.

**Martínez-Flor, A.** 2006. "Task effects on EFL learners' production of suggestions: A focus on Elicited phone messages and emails". *Miscelánea: A Journal of English and American Studies*, 33, 47-64.

**Martínez-Flor, A.** 2013. "Learners' production of refusals: Interactive written DCT versus oral role-play". In Martí, O. and P. Salazar (Eds.) *Refusals in Instructional Contexts and Beyond*. Amsterdam: Rodopi, 175-211.

**Martínez-Flor, A. and Usó-Juan, E.** 2011. "Research methodologies in pragmatics: Eliciting refusals to requests". *ELIA: Estudios de lingüística inglesa aplicada*, 11, 47-87.

**Nureddeen, F.** 2008. "Cross cultural pragmatics: Apology strategies in Sudanese Arabic". *Journal of Pragmatics*, 40, 279-306.

**O'Keeffe, A., B. Clancy, and S. Adolphs**. 2011. *Introducing Pragmatics in Use.* London: Routledge.

**Olshtain, E. and Cohen, A.** 1983. "Apology: A speech act set". In Wolfson, N. and E. Judd (Eds.) *Sociolinguistics and Second Language Acquisition*. Rowley MA: Newbury House, 18-35.

**Oxford Quick Placement Test**. 2001. Oxford: Oxford University Press.

**Rintell, E. and Mitchell, C.** "Studying requests and apologies: An inquiry into method". In Blum-Kulka, S., J. House, and G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies.* Norwood, NJ, The US: Ablex, 248-272.

**Roever, C.** 2010. "Researching Pragmatics". In Paltridge, B., and A. Phakiti (Eds.) *Continuum Companion to Research Methods in Applied Linguistics*. London: Continuum International Publishing Group, 240-255.

**Rose, K.** 1994. "On the validity of DCTs in non-western contexts". *Applied Linguistics*, 22, 1-14.

**Rose, K.** 2000. "An exploratory cross-sectional study of interlanguage pragmatic development". *Studies in Second Language Acquisition*, 22 (1), 27-67.

**Sasaki, M.** 1998. "Investigating ELF students' production of speech acts: A comparison of production questionnaires and role plays". *Journal of Pragmatics*, 30 (4), 457-484.

**Schauer, G.** 2004. "May you speak louder maybe? Interlanguage pragmatic development in requests". In Foster-Cohen, S., M. Sharwood-Smith, A. Sorace

and M. Ota (Eds.), *EUROSLA Yearbook* Amsterdam: John Benjamins Publishing Company, 253-273.

**Schauer, G.** 2009. *Interlanguage Pragmatic Development: The Study Abroad Context.* London: Continuum.

**Schneider, K.** 2008. "Small Talk in England, Ireland, and the USA". In Schneider, K. and A. Barron (Eds.) *Variational Pragmatics: A Focus on Regional Varieties in Pluricentric Languages*. Amsterdam: John Benjamins Publishing Company, 99-139.

**Searle, J.** 1979. *Expression and Meaning: Studies in the Theory of Speech Acts.* Cambridge : Cambridge University Press.

**Sydorenko, T., Maynard, C. and Guntly, E.** 2014. "Rater behaviour when judging language learners' pragmatic appropriateness in extended discourse". *TESL Canada Journal*, 32 (1), 19-41.

**Trosborg, A.** 1987. "Apology strategies in native/non-native speakers of English". *Journal of Pragmatics,* 11 (1), 147-167.

**Trosborg, A.** 1995. *Interlanguage Pragmatics: Requests, Complaints, and Apologies.* Berlin: Walter de Gruyter.

**Yuan, Y.** 2001. "An inquiry into empirical pragmatics data-gathering methods: Written DCTs, oral DCTs, field notes, and natural conversations". *Journal of Pragmatics*, 33, 271-292.

## Appendix A

**Open role-play tasks**

Read the following communicative situations and perform them.

Scenario 1:

| |
|---|
| Student A:<br>You are a university student. You are in a bookshop in the city centre looking for a book. All of a sudden, another girl/boy, who is more or less your age, walks into you accidently. *What would you say?* |
| Student B:<br>You are university student. You are in a bookshop with a friend. You are having a look at a book when your friend calls you. Then you turn around and run into an unknown girl/boy who is more or less your age. *What would you say?* |

Scenario 2:

| |
|---|
| Student A:<br>You are a professor who has asked his/her students to submit a paper. Although there is a deadline, some students have not delivered their work yet. While you are in your office hours, a student that you know from previous years comes there to talk to you. *What would you say?* |
| Student B:<br>You are student at university. You have been asked to prepare a paper and deliver it on a particular day but you haven't done it. Since you know that professor from previous academic years, you decide to talk to her/him. *What would you say?* |

Scenario 3:

| |
|---|
| Student A:<br>You share a flat with another girl/boy. You know that your flatmate has organised a dinner at your place. However, you will not go because you will be out for the weekend. Once you get back home, you realise that your flatmate hasn't tidied up the living-room. *What would you say?* |
| Student B:<br>You share a flat with another girl/boy. While your flatmate is out for the weekend, you have a dinner at the apartment with some friends. However, you haven't tidied up the living-room yet and it is rather untidy. *What would you say?* |

Scenario 4:

| |
|---|
| Student A:<br>You work in a language school. You have asked one of your employees to bring you a textbook that you would like to use in a course that you will teach in two weeks. However, she/he hasn't done it. *What would you say?* |
| Student B:<br>You are a graduate student who works at a language school. Your boss has asked you to bring her a textbook. You know that she/he will need it for a course that she/he will teach in two weeks, but you know that she wanted to have it today. However, you have forgotten it at home. *What would you say?* |

Scenario 5:

| |
|---|
| Student A:<br>You are a university student. A classmate that you know from previous courses has been ill for some days, and because of this she/he has asked you if you could lend her your notes. You agree on that but you tell her/him that you need them back on Friday morning because you have an exam on Monday. *What would you say?* |
| Student B:<br>You are a university student. You have been ill for some days and one of your classmates lent you her/his notes. You have told her/him that you will bring the notes on Friday morning since you both have to study for the exam you have on Monday. Unfortunately, you leave the notes at home. *What would you say?* |

Scenario 6:

| |
|---|
| Student A:<br>You are the language coordinator at the language centre of a university. You have to interview a girl/boy for a job. However, she/he is late and you have been waiting for her for about 25 minutes. *What would you say?* |
| Student B:<br>You have finished your English Studies degree. You have an interview with the language coordinator of the language centre of a university at 10 a.m. but since you are caught in a traffic jam you arrive around 25 minutes late. *What would you say?* |

Scenario 7:

| |
|---|
| Student A:<br>You are a rather famous actress/actor who performs monologues in theatres. While performing your monologue, the mobile of someone from the audience starts to ring. You don't pay much attention to that fact. *What would you say?* |
| Student B:<br>You and a friend go to the theatre to see a monologue. While the actress/actor is performing the monologue, your mobile phone starts to ring. You can't find it to turn it off. Eventually, you manage to do it. *What would you say?* |

Scenario 8:

| |
|---|
| Student A:<br>You have registered on a language course at the university. You have attended all the sessions of the course, so you have all the notes. One day, a student that you do not know sits next to you. At the end of the session she/he suggests going for a coffee. While you are showing her/him the notes, she/he accidently drops her/his coffee on them. *What would you say?* |
| Student B:<br>You have registered on a language course at the university but you have not attended any sessions so you do not have the notes. The first day you go to class, you sit next to another student. After the session, you suggest going for a coffee. While she/he is showing you her notes, you accidently drop your coffee on them. *What would you say?* |

## Appendix B

**Interactive discourse completion task**

Read the following communicative situations and perform them.

Scenario 1:

| Student A: |
| --- |
| You are a university student. You are in a bookshop in the city centre looking for a book. All of a sudden, another girl/boy, who is more or less your age, runs into you accidently. *What would you say?* |
| Student B: |
| You are university student. You are in a bookshop with a friend. You are having a look at a book when your friend calls you. Then you turn around and run into an unknown girl/boy who is more or less your age. *What would you say?* |

A: _____
B: _____
A: _____
B: _____
A: _____
B: _____
A: _____
B: _____

Scenario 2:

| Student A: |
| --- |
| You are a professor who has asked his/her students to submit a paper. Although there is a deadline, some students have not delivered their work yet. While you are in your office hours, a student that you know from previous years comes there to talk to you. *What would you say?* |
| Student B: |
| You are student at university. You have been asked to prepare a paper and deliver it on a particular day but you haven't done it. Since you know that professor from previous academic years, you decide to talk to her/him. *What would you say?* |

A: _____
B: _____
A: _____
B: _____
A: _____
B: _____
A: _____
B: _____

Scenario 3:

| Student A: |
| --- |
| You share a flat with another girl/boy. You know that your flatmate has organised a dinner at your place. However, you will not go because you will be out for the weekend. Once you get back home, you realise that your flatmate hasn't tidied up the living-room. *What would you say?* |
| Student B: |
| You share a flat with another girl/boy. While your flatmate is out for the weekend, you have a dinner at the apartment with some friends. However, you haven't tidied up the living-room yet and it is rather untidy. *What would you say?* |

A: _____
B: _____
A: _____
B: _____
A: _____
B: _____
A: _____
B: _____

Scenario 4:

| Student A: |
| --- |
| You work in a language school. You have asked one of your employees to bring you a textbook that you would like to use in a course that you will teach in two weeks. However, she/he hasn't done it. *What would you say?* |
| Student B: |
| You are a graduate student who works at a language school. Your boss has asked you to bring her a textbook. You know that she/he will need it for a course that she/he will teach in two weeks, but you know that she wanted to have it today. However, you have forgotten it at home. *What would you say?* |

A: _____
B: _____
A: _____
B: _____
A: _____
B: _____
A: _____
B: _____

Scenario 5:

| Student A: |
| --- |
| You are a university student. A classmate that you know from previous courses has been ill for some days, and because of this she/he has asked you if you could lend her your notes. You agree on that but you tell her/him that you need them back on Friday morning because you have an exam on Monday. *What would you say?* |
| Student B: |
| You are a university student. You have been ill for some days and one of your classmates lent you her/his notes. You have told her/him that you will bring the notes on Friday morning since you both have to study for the exam you have on Monday. Unfortunately, you leave the notes at home. *What would you say?* |

A: _____
B: _____
A: _____
B: _____
A: _____
B: _____
A: _____
B: _____

Scenario 6:

| Student A: |
| --- |
| You are the language coordinator at the language centre of a university. You have to interview a girl/boy for a job. However, she/he is late and you have been waiting for her for about 25 minutes. *What would you say?* |
| Student B: |
| You have finished your English Studies degree. You have an interview with the language coordinator of the language centre of a university at 10 a.m. but since you are caught in a traffic jam you arrive around 25 minutes late. *What would you say?* |

A: _____
B: _____
A: _____
B: _____
A: _____
B: _____
A: _____
B: _____

Scenario 7:

| Student A: |
| --- |
| You are a rather famous actress/actor who performs monologues in theatres. While performing your monologue, the mobile of someone from the audience starts to ring. You don't pay much attention to that fact. *What would you say?* |
| Student B: |
| You and a friend go to the theatre to see a monologue. While the actress/actor is performing the monologue, your mobile phone starts to ring. You can't find it to turn it off. Eventually, you manage to do it. *What would you say?* |

A: _____
B: _____
A: _____
B: _____
A: _____
B: _____
A: _____
B: _____

Scenario 8:

| Student A: |
| --- |
| You have registered on a language course at the university. You have attended all the sessions of the course, so you have all the notes. One day, a student that you do not know sits next to you. At the end of the session she/he suggests going for a coffee. While you are showing her/him the notes, she/he accidently drops her/his coffee on them. *What would you say?* |
| Student B: |
| You have registered on a language course at the university but you have not attended any sessions so you do not have the notes. The first day you go to class, you sit next to another student. After the session, you suggest going for a coffee. While she/he is showing you her notes, you accidently drop your coffee on them. *What would you say?* |

A: _____
B: _____
A: _____
B: _____
A: _____
B: _____
A: _____
B: _____

*Received: 10 January 2014*

*Accepted: 22 September 2015*

*Cite this article as:*

**Beltrán-Palanques, V.** 2015. "Revisiting pragmatic tests in the FL context: towards interactive tests to examine speech act performance". *Language Value* 7, 61-90. Jaume I University ePress: Castelló, Spain. http://www.e-revistes.uji.es/languagevalue.

DOI**:** http://dx.doi.org/10.6035/LanguageV.2015.7.5

**ISSN 1989-7103**

*Language Value* November 2015, Volume 7, Number 1 pp. 92-96
http://www.e-revistes.uji.es/languagevalue ISSN 1989-7103

# BOOK REVIEW

***An Introduction to Interaction: Understanding Talk in Formal and Informal Settings***
**Ángela Cora García**
**Bloomsbury: London, New Delhi, New York, Sydney, 2013. 358 pages. ISBN-** 978-1-4411-2768-6

**Reviewed by Raquel Lázaro Gutiérrez**
raquel.lazaro@uah.es
**Universidad de Alcalá, Spain**

This textbook is a comprehensive guide for (mainly, but not only) linguistics, sociology, communication and even business students on the theories and research methodologies of conversation analysis. Based on ethnomethodology, a theoretical perspective of sociology which appeared in the 1960s and which explored how people create social order, social structure and situated action (Garfinkel 1967) through the direct observation of human behaviour, conversation analysis emerged in the 1980s as an approach to the study of talk in interaction. Sacks (1984), a graduate student who worked with Garfinkel, thought of talk as the ideal source of data to study human action, as it could be tape-recorded and carefully and repeatedly examined. The conversation analysis methodology was soon found useful for the study of a wide range of conversations (formal, informal, institutional, etc.) held for a great deal of different purposes and in diverse contexts (business, education, media, legal settings, etc.).

In the first chapter of this manual the author, Angela Cora Garcia, explains in a motivating way why talk should be studied and the many benefits students would obtain from it, such as strengthening listening skills and communicative competence, learning how to carry out qualitative research and developing critical thinking and analytical skills. Following this textbook, students are expected to understand the procedures used to organise talk in interaction, how talk is organised and what types of problems tend to emerge in interaction. Although due to its clear structure and easy-to-read writing style it could be used autonomously by students, this textbook is also intended to be used in class. An "Instructor's Manual" is offered, which contains material for class


DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.6

presentation, additional examples and an appendix with additional resources, such as access to video recordings together with their transcripts.

The book is divided into seven parts. The first includes four chapters about the theory, methodology and characteristics of data used for conversation analysis. Its first chapter is a convincing introduction about the importance of the study of talk in interaction through conversation analysis. It opens with a catchy exercise consisting of an excerpt from an interaction between four people having a conversation and a few reflective questions for the readers. After having explained the usefulness of carrying out conversation analysis, the readers is introduced to the structure of the manual and explained how it should be used. The following chapters within this part consist of an introduction to the origins of conversation analysis, that is, ethnomethodology (Chapter 2), and its methodological approach or "how to study interaction from a conversation analytic approach" (Chapter 3). The latter chapter explores such methodological issues as the collection of naturally occurring data, the size of the sample (collections of data or case studies), or ethics and the importance of obtaining permissions for tape recording. The author also warns about a common burden researchers have to cope with when observing talk in interaction, i.e. as conversation analysts deal with naturally occurring discourse, they have to be aware of the "observer's paradox" (Labov 1972), or the influence of the researcher on the speakers, who may change their conversational behaviour when they feel they are being observed.

The last chapter of this part deals with a key issue about the preparation of data: transcription practices. Whether data is gathered using audio or video recorders, researchers have to transcribe the recordings including a variety of details. Based on the transcription conventions developed by Jefferson in the 1970s, the author shows how to transcribe hesitations, pronunciation, laughter, breaths or simultaneous speech.

Part II, consisting of nine chapters, tackles the main characteristics of the organisation of talk in interaction using examples from ordinary conversations in everyday life or institutional interactions. Thus, the turn-taking system is explained throughout Chapter 5, where, following Sacks et al. (1974), the turn constructional and the turn allocational components are described. Sequences of turns are examined throughout Chapter 6, where the concept of adjacency pair is developed, and Chapter 7, where the

characteristic sequential organisation of conversations is explored to discover how participants perform exchanges which are longer than two turns. Thus, two-turn action sequences can be expanded by means of side sequences, insertion sequences or pre-sequences up to the point of building series or chains of adjacency pairs.

Chapters 8 and 9 are devoted to openings and closings, respectively. In order to explain these two mechanisms, the author uses examples from studies on telephone conversations, such as the already classical ones conducted by Schegloff (1979, 1986) or Whalen and Zimmerman (1987). The following chapter deals with error avoidance and repair, both of one's own and of others' discourse, whereas Chapter 11 explains how speakers create, close, change or refocus topics. In this way we learn that topical continuity and coherence depend on the range of techniques participants use to maintain them rather than on the subject matter of the conversation.

The research carried out by Goodwin (1984) on story-telling is the tool used by the author in Chapter 12 to explain how aspects of nonverbal language (facial expressions, hand gestures and gaze direction, amongst others) accompany talk to complete communication. Embodied actions give information about the structure (rather than about the content) of the talk, and make it easier for the audience to follow stories. Part II finishes with a chapter about how speakers refer to others. Membership categorisation analysis is applied to three different settings (telephone calls, legal procedures and the workplace) to show how participants categorise themselves and others.

Whereas the chapters in Part II are focused on particular organisational principles of talk in interaction, the thirteen remaining chapters explore conversations in institutional settings. On reading them, the student can grasp the particular nature of these encounters in terms of how orchestrated routines, work-related tasks or institutional roles are performed through talk. Part III comprises three chapters about telephone calls (emergency calls to the Police, emergency service calls and the influence of technological transformation on talk by telephone), and another one about air traffic communication. Thus, Chapters 14 and 15 summarise research by authors such as Zimmerman (1984), Whalen and Zimmerman (1987) and Schegloff (1968, 1979), amongst others, to explain "how service call takers and callers work together to accomplish the work of the call in a timely and efficient manner", whereas Chapter 16

explains how interactions on the phone have changed due to technological innovations such as caller identification.

The following section (Part IV) deals with talk in medical settings and includes one chapter about the delivery of bad news and another one about the Primary Care consultation. This latter chapter examines the structure of the medical consultation, focusing particularly on the interview physicians carry out on patients to elaborate the diagnosis, on the one hand, and the phase of treatment recommendation on the other. Talk in legal settings such as trials, hearings and other proceedings, police interrogation and pre-trial interviews of victims, or mediation sessions is explored in the following section (Part V). Throughout these pages students can notice how interrogation techniques vary when the interviewee does not cooperate with the interviewer, that is, when the participants in the interaction do not pursue the same objective but are instead on opposing sides of a dispute.

The last two sections explore talk in broadcast media (Chapters 23 and 24 deal with television news interviews and call-in talk shows on radio and television, respectively) and talk in business settings (customer service in Chapter 25 and meetings, interviews and performance appraisals in Chapter 26). Issues such as the performance of roles and identities, power asymmetry and the use of negotiation and persuasion are developed throughout these pages.

Each of the chapters of this textbook is written in an understandable way using a clear and simple style. Students are asked questions at different moments so that they can reflect before contents are presented, can focus better on key aspects of the topics at hand or can compare the examples given in the book with conversations they have in real life. The chapters include summaries of key points and exercises for students. The last chapter actually wraps up as it presents the same examples of conversations which appeared in the first chapter, but this time written using detailed transcripts instead of simplified ones. By the time they reach the last chapter of this textbook students are expected to be able to analyse these detailed transcripts and apply the basic findings of conversation analytical research to them and to other conversations.

## REFERENCES

**Garfinkel, H**. 1967. *Studies in Ethnomethodology.* Cambridge: Polity Press.

**Goodwin, C.** 1984. "Notes on story structure and the organization of participation". In Atkinson, J. M. and J. Heritage (Eds.) *Structures of Social Action: Studies in Conversation Analysis.* Cambridge: Cambridge University Press, 225-246.

**Labov, W.** 1972. *Sociolinguistic Patterns.* Oxford: Blackwell.

**Sacks, H.** 1984. "Notes on methodology". In Atkinson J.M. and J. Heritage (Eds.) *Structures of Social Action: Studies in Conversation Analysis.* Cambridge: Cambridge University Press, 21-27.

**Sacks, H., Schegloff, E.A. and Jefferson, G.** 1974. "A simplest systematics for the organization of turn-taking for conversation". *Language*, 50, (4), 696-735.

**Schegloff, E.A.** 1968. "Sequencing in conversational openings". *American Anthropologist*, 70, 346-380.

**Schegloff, E.A.** 1979. "Identification and recognition in telephone conversation openings". In Psathas, G. (Ed.) *Everyday Language: Studies in Ethnomethodology.* New York: Irvington, 23-78.

**Schegloff, E. A**. 1986. "The Routine as Achievement". *Human Studies*, 9, 111-151.

**Whalen, J. and Zimmerman, D.H.** 1987. "Sequential and institutional contexts in calls for help". *Social Psychology Quarterly*, 50, (2), 172-185.

*Received: 1 September 2014*

*Accepted: 15 Febuary 2015*

*Language Value* December 2015, Volume 7, Number 1 pp. 97-99
http://www.e-revistes.uji.es/languagevalue ISSN 1989-7103

# BOOK REVIEW

*Changing Methodologies in TESOL*
**Jane Spiro**
**Edinburgh University Press: Edinburgh, 2013. 224 pages. ISBN:** 9780748677641

**Reviewed by Elena Martín Monje**
emartin@flog.uned.es
**Universidad Nacional de Educación a Distancia, Spain**

This book offers an overview of key concepts related to TESOL (Teaching English to Speakers of Other Languages) and the methodological changes that this field has undergone in recent decades. The author states that the book "is aimed at the student, student teacher and practising teacher of TESOL interested in how the English language is experienced, taught and learnt in the twenty-first century worldwide" (p.1). This wide angle can be perceived in the numerous examples and case studies that populate the volume, making it a pleasant read.

*Changing Methodologies in TESOL* is divided into three key sections: 1) methods and the language learner; 2) the language in methods; and 3) the world in the classroom, plus an introductory chapter to the whole volume. Carefully structured, each chapter includes an introduction, some discussions connecting current debates in TESOL with real teachers in practice, case studies, suggested tasks and further and guided readings. It should also be mentioned that there is additional material available on the publisher's website for each of the topics covered.

Chapter 1 is an introduction entitled "The Meaning of Methods" and focuses on the concepts of 'methods' and 'methodology' and why this theoretical side is important for TESOL. Furthermore, it makes the case for the important connection between methods and methodology and actual teaching practice. After this introduction, the first of the three main sections, "Methods and the Language Learner" (comprising Chapters 2 and 3), deals with the impact of learners' needs, context and culture on language learning and teaching approaches and reflects on how methods take account of the learner.


DOI: http://dx.doi.org/10.6035/LanguageV.2015.7.7

Chapter 2, "Learning Theories and Methods", explores teaching methods as responses to teachers' beliefs and attitudes towards the language learning process and takes an eclectic approach, very much in line with Kumaravadivelu's point of view (2006). Chapter 3, "The Place of the Learner in Methods", looks into differences among learners and how teachers should be aware of them so that they can vary their teaching methodology accordingly. As Spiro puts it, "a method is not in itself successful or effective; it is 'good' or 'bad' as a method only insofar as it 'fits' with the learners themselves" (p. 55).

The second section, "The Language in Methods" (Chapters 4-7), turns to the impact of the knowledge of language in methods, from the word to whole texts, including writing, speaking and also the pedagogic approaches to vocabulary and grammar. The latter is the focus of Chapter 4, "Grammar in Methods". This is one of the few chapters written by another author (Paul Wickens in this case) and it delves into what teachers should know about how learners learn grammar and grammar itself, giving some ideas about how to use corpora as pedagogic resources. Chapter 5 deals with "Vocabulary in Methods" and offers useful strategies for acquiring vocabulary, as well as an overview on the main trends in the teaching of vocabulary today. Chapter 6, "Teacher Knowledge and the Four Language Skills: Understanding Written and Spoken Language in the Twenty-First-Century World", reflects on the changing pedagogies related to spoken and written language and pays attention to modern concepts such as English as Lingua Franca and World Englishes. However, it is surprising to still find the traditional division into four skills, with no mention of the eight language activities provided by the *Common European Framework of Reference for Languages*, which has become an unavoidable benchmark since its publication at the turn of the century: "The language learner/user's communicative language competence is activated in the performance of the various language activities, involving reception, production, interaction or mediation (in particular interpreting or translating). Each of these types of activity is possible in relation to texts in oral or written form, or both" (p. 14). Finally Chapter 7, "Methods and Principles for Integrating the Four Skills: Reading, Writing, Speaking and Listening", complements the previous chapter by showing how to implement the suggested strategies into the aforementioned four skills, proving that there is a necessary interdependence among them.

The third section, "The World in the Classroom", explores competences that have become essential in the 21st century, from intercultural awareness to digital literacy, paying attention to all the capabilities that are developed when learning a language and how the teacher's methodology can take account of them. Chapter 8, "Multiple Literacies: Professional, Academic and Web Literacies in Methods" focuses on English for Specific Purposes and the development of multiple literacies, including digital competence. Chapter 9, "Cultural Competences in Methods", looks at the important role of culture in language learning and provides tips for teachers to enhance their learners' social, pragmatic and cultural knowledge in the language classroom. Lastly, Chapter 10, "Windows into TESOL Classrooms: Where Are We and Where Are We Going?", with classroom examples from John Eyles, showcases a wide range of classroom types, which reflect the diverse and complex settings language teachers can come across these days, from the high-technology class to the low-technology one, large or small classes or even self-access ones.

*Changing Methodologies in TESOL* will be a valuable resource for pre-service TESOL courses. It may also be used in other ways, as a reference book for teacher trainers and even for individual teachers, helping all of them to keep up with the latest developments in the field and those interested in reflecting on their own practice.

## REFERENCES

**Council of Europe** (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

**Kumaravadivelu, B.** (2006). TESOL Methods: changing tracks, challenging trends. *TESOL Quarterly 40*(1), 59-81.

*Received: 28 July 2014*
*Accepted: 10 February 2015*