



## **KPG oral exams: task design considerations and actual performance**

**[Προφορικές εξετάσεις στο ΚΠΓ: προβληματισμοί σχετικά με το σχεδιασμό και την εκτέλεση των δραστηριοτήτων]**

**Eleftheria Nteliou**

*This paper presents a task analysis model that has been designed for the oral tasks of the Greek State Exams for foreign language certification (known as KPG). The theoretical background of the model is based on the systemic functional approach to language use as well as on the notion of genre as both process and product. The purpose of this paper is to show how the model has been applied to describe and analyze the tasks designed for the KPG oral exams at levels B1 and B2. The analysis of the oral task rubrics shows that the generic process specified by the task is likely to lead test takers to particular lexicogrammatical choices, which, in turn, cause differences in the expected language performance. Implications for future research, language teaching and testing are mentioned.*

∞

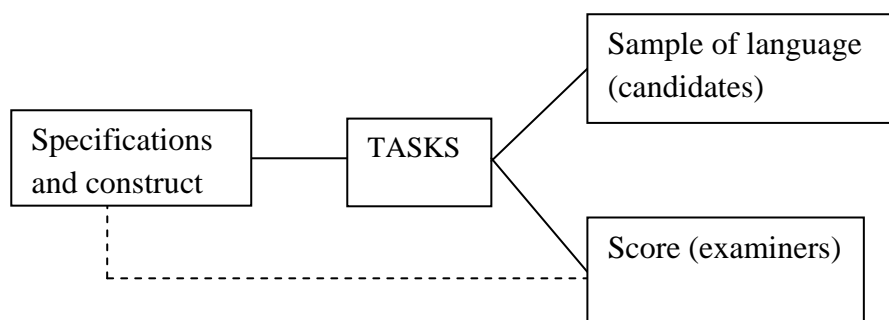
Στο παρόν άρθρο παρουσιάζεται ένα μοντέλο γλωσσολογικής περιγραφής και ανάλυσης των προφορικών δραστηριοτήτων που χρησιμοποιούνται στις κρατικές εξετάσεις πιστοποίησης γλωσσομάθειας (γνωστές, εν συντομία, ως ΚΠΓ). Το θεωρητικό υπόβαθρο του μοντέλου βασίζεται στη συστημική λειτουργική προσέγγιση χρήσης της γλώσσας, καθώς επίσης και στην αντίληψη ότι το κειμενικό είδος δεν είναι μόνο τελικό αποτέλεσμα αλλά και διαδικασία. Ο σκοπός του άρθρου είναι να παρουσιάσει την εφαρμογή του μοντέλου στην περιγραφή και ανάλυση των προφορικών δραστηριοτήτων στις εξετάσεις του ΚΠΓ, στα επίπεδα B1 και B2. Η ανάλυση των εκφωνήσεων των προφορικών δραστηριοτήτων δείχνει ότι η κειμενική διεργασία είναι πιθανό να οδηγήσει τους υποψηφίους σε συγκεκριμένες λεξικογραμματικές επιλογές, οι οποίες, με τη σειρά τους, μπορεί να προκαλέσουν διαφορές στην αναμενόμενη παραγωγή λόγου. Σχολιάζεται, επίσης, η σημασία της παρούσας μελέτης για τον τομέα της διδασκαλίας και εξέτασης των γλωσσών καθώς και για μελλοντική έρευνα.

**Key words:** task demands, generic process, lexicogrammar

## Introduction: oral task characteristics and test performance

In the construct-based approach to assessment, the theoretical description of the language ability to be measured is the first important step to test design, because it normally influences task requirements and the assessment criteria on which actual performance is evaluated (Luoma, 2004, p. 42). In fact, the match between what is intended to be tested (i.e. the construct) and what is actually produced and measured during test performance, constitutes a ‘validity argument’, which needs sufficient evidence to be sustained (Fulcher and Davidson, 2007). Therefore, in test-task validation procedures, tasks play a central role, because the language output they generate constitutes evidence that the test actually measures what it intends to measure.

In oral testing, the central role of ‘task’ and its direct influence on language performance can be seen in Graph 1 which presents a conceptual framework for performance testing. This visual representation was initially proposed by Milanovic and Saville (1996) and later reproduced by O’Sullivan et al (2002, p.35). The framework portrays the various factors (or facets) that should be taken into consideration when designing a test, because they are likely to affect performance and threaten test reliability and validity. According to the framework, the task is directly related to test specifications and the theoretical description of the language intended to be tested (i.e. the construct). Moreover, the task affects the sample of language produced by the candidates, which is evaluated by the examiners and leads to a final score. Based on O’Sullivan et al’s framework, Graph 1 schematically presents how the task relates to the test construct as well as to language performance and assessment.



Graph 1. Tasks and their influence on performance.

This paper focuses on the oral tasks designed for the B1 and B2 level speaking module in the KPG<sup>1</sup> exams in English and presents a model that was used to describe and analyze the speaking tasks in terms of their linguistic demands at the level of oral language production. Apart from their influence on language performance, what is also worth examining is how the task demands relate to the theoretical construct of language ability, as described in the KPG exams specifications (ΥΠΕΠΘ, 2003). The focus of the research is on levels B1 and B2 because exams at these two levels attract a great number of candidates, given that language knowledge certification can also be used as a job qualification, especially at level B2. Moreover, if there is a clear picture of what a B1 and B2 level speaker is expected do with language, we can also make sound inferences about task design and performance at lower and higher levels (i.e. A1-A2 and C1-C2). The next section presents the research background, which relates to tasks and their influence on language production, as well as the theoretical linguistic background on which task design in the KPG exam system is based.

## **Literature review**

The assumption that language output is greatly affected by task design has been widely investigated in the oral testing literature. Studies have shown that different types of tasks (e.g. role play, narration) lead to the elicitation of different language characteristics during oral performance (e.g. Bygate, 1999; Kormos, 1999; Pavlou, 1997; Shohamy, 1988; Young, 1995). O'Sullivan (2008) offers a comprehensive review of the literature on oral tasks and their influence on language production, from a number of perspectives that do not only relate to task types but also to test methods (e.g. 'live' versus 'tape mediated' oral tests) and specific intra-task characteristics, such as planning time and level of difficulty. Because tasks greatly influence language output on which a score is based, Bygate (2009, p. 414) concludes that 'to appraise students' language, we first need to understand the linguistic demands of our tasks'.

The analysis of task demands is likely to lead to expectations or predictions relating to actual performance. O'Sullivan et al (2002) note that there is scarce research on relating test performance to the test designers' predictions or expectations, which are based on their definition of the language construct. For that reason, they first analyzed the construct of spoken ability in the Cambridge ESOL Main Suite speaking tests in terms of a list of language functions. Based on this analysis, they then developed observation checklists, which were used in the *a priori* and *a posteriori* analysis of speaking task output. Using this method, they tried to validate the match between intended and actual test taker language. One of the implications of the study was that the analysis of task content and prediction of the actual performances elicited by test tasks define the construct tested. The prediction of linguistic responses can greatly help in the process of task design, because it can "lead to a greater understanding of how task and task formats can be manipulated to result in specific language use" (O'Sullivan et al., 2002, p. 47).

In order to explore the relationship between expected and actual task performance, we first need to be aware of the construct of language ability, which forms the basis for task design and determines the demands regarding language production. In the case of the KPG exams, the definition of language ability is based on a functional approach to language use, whereby language is viewed as a social phenomenon, taking place within a particular context of communication, which determines the language choices participants make to achieve certain communication purposes (ΥΠΕΠΘ, 2003, p. 31). As Karavas states, the main aim behind test design in the KPG exams is "to evaluate socially purposeful language knowledge and literacy" (2009, p.24).

The functional approach to language use, which is adopted by the KPG exam system, is mainly described in the work of Halliday (Halliday, 2002; Halliday and Hasan, 1976, 1985; Halliday and Mathiessen, 1999, 2004), who has laid the foundation of what is known as 'Systemic Functional Linguistics' (SFL). SFL is a theory which emphasizes the importance of context in making meaningful interactions. In contrast to traditional grammar, according to which language is strictly realized through a set of rules of form and structure, SFL underscores the importance of the context of communication, which should determine language choices. The communication context, which refers to 'the linguistic, physical and psychological dimensions of the situation in which language is used' (Karavas, 2009, p. 25), determines the types of functions that the participants should perform (e.g. to inform, to advise, to persuade, etc.) as well as the production of different text types, which are called 'genres'. In her systemic functional approach to language use, Eggins (1994, p. 30)

distinguishes two types of context, which influence linguistic behaviour: the 'context of situation', which actually determines the register variables of the text (e.g. topic, participant roles, the role of language) and the 'context of culture' that describes the overall purpose of the text with which a genre is associated.

Several SFL applications can be found in the educational context, especially in literacy instruction in English as a second language, most notably in the work of "Sydney School" in Australia (Martin, 1998). However, its relation to assessment procedures has been scarcely investigated. There are some studies that have shown how the SFL principles can effectively be applied in second language learning for classroom assessment procedures (e.g. Huang and Mohan, 2009; Hughes, 2009). Nevertheless, no relation between SFL and the linguistic theory behind high stakes exams has ever been reported. The KPG examination system must be the first and unique to base its description of the language construct and test specifications on the principles of SFL. In particular, this study intends to show how the theoretical approach to SFL can influence task design and lead to specific expectations regarding oral production, focusing on the KPG speaking tests.

### **Research context: The KPG B1 and B2 speaking modules**

The Greek certification system of foreign language proficiency was introduced in April 2003. Tests are designed for A1-A2, B1-B2 and C1 levels, as they are set by the Council of Europe and described by the Common European Framework of Languages (CEFR; Council of Europe, 2001). Regarding the KPG oral exams in English, level B1 and level B2, which both have exactly the same format and structure, were introduced in November 2003 and in May 2007 respectively<sup>2</sup>. Test takers are examined in pairs, but they do not interact with each other. There are two Examiners, one of whom acts as an Interlocutor (i.e. he assigns the tasks to each candidate and marks their performance) and the other is an Examiner-Assessor, who is silent during the oral exam and just marks performance.

Candidates are examined in their ability to respond to personal questions (Activity 1), to talk about something based on one or more visual prompts (Activity 2) and to relay in English information from a Greek text (Activity 3). What differs from level to level are the task demands and the complexity of test input (e.g. the Greek texts used as input in the mediation activities). Moreover, the description of the illustrative descriptors<sup>3</sup> for oral production presents differences in the quality of the expected language output between these two levels. In general, the B1 level tasks and questions are intended to be cognitively and linguistically less demanding than the ones at level B2 and, therefore, the quality of language production at these two levels is likely to differ, at least in terms of complexity.

In particular, KPG oral task designers take into account the expected differences in the B1 and B2 level speakers' speaking skills, which are described in the KPG specifications for these levels (Dendrinos, 2007). The KPG specifications for oral ability at levels B1 and B2 are slightly adapted from the description of oral language ability in the respective levels, provided by the CEFR. Moreover, oral task design incorporates characteristics of the SFL approach to language use. Since, according to SFL, language is a means to make meaning with words within a specific contextual frame, the task designers in the KPG exams should be careful in providing a clear context and purpose of communication as well as specific participant roles. This type of information is usually provided in the input material (e.g. photos, texts) as well as in the task rubrics. In order to investigate the link of task design with task performance in the KPG exams, a tool has been designed, whose purpose is to describe

the linguistic features of expected performance, on the basis of the theoretical considerations that govern task design in the KPG exams.

### **Relating KPG task demands to performance: methodology adopted**

The University of Athens launched a large-scale research project in 2007, whose aim was to linguistically describe and analyse the test tasks and texts that had been used in the KPG exams of all levels and modules in the four languages (i.e. English, French, German, Italian) that were being tested up to that moment<sup>4</sup>. For that purpose, a group of experts specified the linguistic parameters on the basis of which researchers from various areas in the field of Linguistics (e.g. applied and theoretical linguistics, computational linguistics, psycholinguistics, etc.) were asked to linguistically describe the KPG test tasks. The present researcher participated in the group that described and analysed the expectations of the oral tasks designed for the KPG oral exams in English. The project coordinator was responsible for checking all the descriptions and providing feedback. However, the initial task descriptions differed a lot, because the researchers came from very different research and educational backgrounds, which influenced their view of the linguistic terms on which description was based. Consequently, there was a need for clearer description guidelines as well as for a clearly defined theoretical basis on which task description would be conducted.

The theoretical background of the Task Description Model (TDM) that was finally determined mainly adheres to the systemic functional approach to language use, proposed by Halliday (1985, 2002) as well as to a model of language proposed by Knapp and Watkins (1995), which describes grammar in relation to genre and text. Although Knapp and Watkins' model was developed to assist the instruction and assessment of writing skills within the Australian educational context, it was very helpful in the process of associating text purpose with specific lexicogrammatical features in all the tasks used in the KPG exam system. Besides, this model is also based on the principle that language is a socially constructed phenomenon, mainly expressed through the Hallidayan functional approach to language use, which generally maintains that genres relate to specific linguistic structures. However, Knapp and Watkins move one step forward. What makes their model special is the fact that genres are not seen as products or text types, but as sets of generic processes with a specific purpose, which is achieved by the application of "relevant structural and grammatical knowledge to produce appropriate texts" (Knapp and Watkins, 1995, p.26). In particular, in their model, there are five types of generic processes (*description, explanation, instruction giving, argumentation* and *narration*), each of which is realized through specific lexicogrammatical, structural and cohesive elements.

Based on the theoretical frame described above and after several discussions and revisions (see Κονδύλη & Λύκου, 2009), the task description model (TDM) finally included six categories of analysis: *topic, genre, generic process, speaker (or writer) - audience roles* and *lexicogrammar*. These categories were also used in the analysis of the KPG speaking tasks designed for levels B1 and B2, the results of which are presented in this study.

### **The TDM and the B1-B2 KPG speaking tasks: research aims**

The six categories of the TDM were used to analyze the task rubrics designed for Activities 2 and 3, in the B1 and B2 KPG speaking modules and administered over nine consecutive exam administrations from April 2005 to May 2009. *Topic* and *speaker-audience roles* were determined by the information provided in the task rubrics. Regarding *genre*, two types

were analyzed: the ‘monologic talk, based on visual prompts (Activity 2 tasks) and the ‘monologic talk based on Greek text(s)’ (tasks of Activity 3<sup>5</sup>). Concerning ‘generic process’, the features on which description was based came from Knapp and Watkins’ categorization of five types of generic processes, on the basis of their lexicogrammatical, structural and cohesive characteristics:

- Description
- Narration
- Explanation
- Instructions
- Argumentation

Regarding the final category (Lexicogrammar), the description included the following characteristics, most of which were considered within the frame of the Hallidayan approach to functional grammar:

- *vocabulary* (every-day / subject specific)
- *tense* and *aspect*
- *modality* (deontic / epistemic)<sup>6</sup>
- *clause types* (verbal, mental, relational, material, existential)
- *cohesive devices* (additive, temporal, causal, consequential, pronouns etc)
- *impersonal syntax*
- *direct/reported speech*
- *grammatical metaphors* and *idioms*<sup>7</sup>.

An example of B2 task analysis is provided in Table 1.

Activity 2: Having Fun (May 2006) – Task 2 <i>Look at photos 14 and 16 and tell us who you think the people in them are in each case, what their relationship is and what they enjoy doing in their free time.</i>				
Topic:	Genre:	Generic process:	Speaker - Audience roles:	Lexicogrammar
Having fun	Monologic talk (based on visual prompts)	Description (of people, feelings, likes and dislikes)	Test taker - Examiners	Present tenses Epistemic modality (modal verbs and adverbial phrases expressing probability) Types of clauses: material, relational, mental (cognitive and emotive) Type of cohesion: through pronouns

Table 1: Using the TDM categories to analyze an example of B2 level speaking task.

The analysis initially focused on two of the five categories in the TDM: the generic process and the lexicogrammar. Given that Knapp and Watkins associate their five types of generic processes with written production, it would be interesting to find out what types of generic processes are more frequently generated by oral production tasks in the KPG exams. Moreover, the interaction between generic processes and lexicogrammatical choices in spoken performance has been relatively unexplored. Therefore, the aims of the present study were:

- to detect whether there were any types of generic processes or any lexicogrammatical characteristics that tended to appear more frequently than others in the oral task rubrics of Activities 2 and 3, at levels B1 and B2, and
- to examine these characteristics in combination, in order to find out whether there were any systematic patterns of item combinations at each level (e.g. whether a particular type of generic process was expected to systematically lead to specific lexicogrammatical characteristics).

For data analysis purposes, two separate Excel worksheets were created (one for each level of proficiency under examination) in order to count the frequency of the items in each category. The task rubrics of Activities 2 and 3 were also analyzed separately, in order to detect any differences in the frequency of the analytical categories across activities. In this study, the findings focus on the interaction between generic process and its lexicogrammatical realization in terms of modality and cohesion in the oral tasks designed for the B1 and B2 KPG oral exams.

## Findings and Discussion

Based on the analysis of the task rubrics designed for Activity 2 at levels B1 and B2, Table 2 shows in how many tasks test takers are expected to get involved with each generic process. According to Table 2, in most tasks test takers are asked to deal with *Description*. *Explanation* and *Narration* appear to be the next most frequent types of generic processes at levels B2 and B1. However, Table 2 fails to show that in Activity 2, there are some tasks which involve test takers in a combination of generic processes. The most common of these is '*Description and Explanation*'. This combination of generic processes is very frequent, both at level B1 (60%) and at level B2 (94%). For example, the test takers are asked to describe something (e.g. a picture, a situation, their feelings, etc) and explain something, in relation to the situation portrayed in the visual prompts. Furthermore, '*Instructions and Explanation*' as well as '*Narration and Explanation*' are found to be two, less frequent, combinations. The task rubric analysis also revealed that the generic process of *Description* should be more subtly defined because it may include the description of people, events, thoughts, feelings or even the function of providing information about something.

Generic Process	B1 level tasks (total number of tasks: 84)	B2 level tasks (total number of tasks: 190)
Description	59	152
Explanation	18	86
Instructions	7	3
Narration	18	7
Argumentation	2	3

Table 2: Number of Activity 2 tasks with expected types of generic processes

The analysis of the task generic processes could lead to certain assumptions regarding their realization at the lexicogrammatical level. In particular, since test takers at levels B1 and B2 are very likely to be assigned tasks in which they have to describe and explain something, oral text cohesion will probably be achieved through *referential*, *causal* and *temporal* cohesive devices. For example, the coherent description of a situation involves the use of referential pronouns that link the participants together, whereas when an explanation is provided, the use of causal connectives (e.g. because, since, etc.) is necessary. Moreover,

the use of time expressions is necessary in the narration of events (usually at level B1) as well as in some cases when a situation is described or when a prediction is made. Based on these assumptions, Table 3 shows how often each of these types of cohesive devices is expected to appear in the candidates' linguistic performance in the tasks of Activity 2, at levels B1 and B2.

Types of cohesion	Frequency of types of cohesion	
	Level B1 (total number of tasks: 84)	Level B2 (total number of tasks: 190)
Referential (through pronouns)	52%	46%
Causal	29%	47%
Temporal	50%	22%
Additive	20%	15%
Consequential	5%	4%
Concessional	4%	10%

Table 3: Expected frequency of types of cohesion in Activity 2

Another finding relates to the test-takers' degree of familiarization with the situation presented in the task and how this influences their linguistic choices, especially in terms of 'modality', which constitutes another category in the TDM. In particular, at level B2 test takers are frequently presented with a picture, in which the situation or people depicted are unknown to them and the tasks require they say who the people are, what they are doing or how they feel. Consequently, since the candidates do not necessarily know the people in the visual stimuli and have to make inferences about what happens, they need to use specific modal verbs and adverbial phrases expressing probability (e.g. could, may, perhaps, etc.) or mental clauses (e.g. I think / believe that...), all of which constitute ways by which 'epistemic modality' is realized. On the contrary, most of the Activity 2 task rubrics designed for level B1 are phrased in such a way that presupposes that test takers participate in the situation and know the people presented on the prompts. Consequently, no inferences need to be made. In fact, test takers are mostly asked to make descriptions (of people, feelings, the situation), using present and past tenses, without any expressions of modality.

Table 4 shows how many Activity 2 tasks require test takers to express either epistemic or deontic modality or neither of them. After analysing task rubrics and their lexicogrammatical expectations, Table 4 shows that, in fact, more instances of epistemic modality are expected at level B2 than at level B1 (see Table 4). This finding also implies that mental clauses are expected to be more common at level B2 than at level B1, because they are one of the various ways in which epistemic modality can be realized. Regarding deontic modality, it seems less likely to be traced in the tasks of Activity 2, because it relates to the generic process of giving instructions or advice, which has only been encountered in a few B1 tasks. There are also many tasks which do not involve test takers into any kind of modality. These might be the tasks that presuppose that the test takers are familiar with the situation on the visual prompts and they just have to describe or explain it, by using tenses.

Concerning the tasks of Activity 3, the most prevalent type of generic process is *Instructions*, at both levels. However, similarly to Activity 2, oral mediation task rubrics seem to generate combinations of generic processes. In these cases, the most common combination is '*Instructions and Description*', both for level B1 (75%) and level B2 (59%). Moreover,



'Description and Explanation' is another, less frequent, combination, which mostly appears in a few B2 level oral mediation tasks.

Types of modality	B1 level tasks (total number of tasks: 84)	B2 level tasks (total number of tasks: 190)
Epistemic	25	93
Deontic	5	4
Nothing	52	92

Table 4: Number of Activity 2 tasks with expected cases of modality

Since the most common generic process in the tasks of Activity 3 is *Instructions* and test takers are mostly asked to provide instructions or advice, a frequent use of deontic modality expressions should be expected, mainly in the form of modal verbs expressing obligation (e.g. should, can, could, etc.). Interestingly, the analysis of the Greek texts that serve as input in the oral mediation tasks revealed that *Instructions* is also the most common type of generic process in the texts used at both levels. However, what differs in the Greek texts is the way advice and instructions are expressed. For example, information is mostly provided in present tenses and in bullet form. The test takers should use this information to address another interlocutor and advise him/her by using modal verbs (e.g. *should, could*), imperative or other expressions of deontic modality (e.g. *it is necessary that ...*). Moreover, given that in most mediation tasks the test takers are asked to provide advice or instructions and give explanations, then speech cohesion is expected to be achieved by numbering (i.e. instruction steps/advice) and using causal connectives, whereby *additive* and *causal* cohesion are respectively realized.

In summary, these findings have revealed some basic differences between Activity 2 and 3, at both levels of proficiency. Regarding the type of generic process, in the tasks of Activity 2, test takers are usually involved in *Description* and *Explanation*, whereas in Activity 3 *Instructions* is the most frequent generic process. These differences in the generic processes lead to differences in the expected lexicogrammatical characteristics during oral performance in the tasks of Activities 2 and 3. More specifically, in *Description* and *Explanation* the speakers are more likely to use referential pronouns and causal links in order to perform the task. Moreover, if they are presented with an unfamiliar situation they need to resort to epistemic modality, in order to make inferences about the people and what exactly happens. Therefore, since *Description* and *Explanation* are the most frequent generic processes in the tasks of Activity 2, referential, causal cohesion and epistemic modality are expected to be frequently found in the test takers' language output. On the contrary, since *Instructions* is the most common generic process in the tasks of Activity 3, test takers are expected to use additive cohesive devices as well as expressions of deontic modality in order to give instructions or advice.

As far as the differences between levels B1 and B2 are concerned, the analysis of the task rubrics has shown that at level B2 test takers are more frequently asked to imagine that they are part of a fictitious situation. Therefore, more instances of epistemic modality are expected at level B2, whereas at level B1 the use of present tenses to provide simple descriptions is more frequent. Besides, the expression of probability or obligation (i.e. modality) is usually taught after the instruction of tenses, therefore, instances of modality are less expected at B1 level oral production.

## Conclusion, implications and limitations

The description and analysis of the oral tasks used at B1 and B2 level KPG exams revealed systematic interactions between generic process and lexicogrammatical realizations in the oral tasks analyzed in terms of modality and cohesion. In particular, specific generic process combinations were repeatedly observed in the tasks of Activities 2 and 3, which are expected to lead to specific lexicogrammatical characteristics, without any particular deviating cases. This finding offers evidence of task consistency in the oral KPG tasks at levels B1 and B2, which positively adds to the validity of the task design process. Moreover, the tasks designed for Activities 2 and 3 were found to involve the test takers in different generic processes, thus leading the test-takers to different lexicogrammatical choices. Therefore, they are evaluated on different aspects of linguistic ability, something that positively adds to the fairness of their final score.

Task analysis by the means of the TDM resulted in an analytical description of some lexicogrammatical characteristics that should be found in the successful oral performance at levels B1 and B2, in the KPG exams. At the same time, the central role that generic process plays in making predictions about the characteristics of oral performance was highlighted. If the characteristics of expected language production are also observed in real time performance, they could be used to make the illustrative descriptors for oral production at levels B1 and B2 more accurate. Therefore, material writers would be able to base task construction on clearer guidelines. Moreover, the findings of task description could help in describing more accurately the oral assessment criteria that refer to lexicogrammatical choices at these two levels of proficiency. The improvement of the oral assessment scale combined with the oral examiners' training in relating the generic process implied by the task to specific lexicogrammatical characteristics, would hopefully ensure sounder criteria for marking oral performance in the KPG exams.

However, in order to obtain even more accurate descriptions regarding expected oral performance at levels B1 and B2, the TDM should be used to explore the interaction of even more linguistic characteristics. For example, it would be useful to explore how topic relates to lexical choices or to the use of specific tenses. This type of analysis could probably reveal more differences regarding expected oral production at levels B1 and B2. Additionally, the discourse analysis of real test performance could shed some light on how lexicogrammatical choices are actually realized at different levels of proficiency.

Future research should analyse oral task characteristics at levels A1-A2 and C1 to find out how tasks become more demanding from lower to higher levels of proficiency in the KPG exams. Ultimately, all the findings from the application of the TDM to tasks of oral production could be compared with those on written production, thus acquiring a complete picture of the linguistic characteristics of expected language production in the B level KPG exams.

Finally, this study presents some pedagogical implications for teachers, exam providers and material writers. In particular, since task design in the KPG exams is based on the assumption that context affects linguistic choices, test takers need to be taught how to analyze the information provided in the task rubrics. This does not mean that learners should necessarily be exposed to the SFL principles or get familiarized with such terms as *generic process*. Instead, they should be aware of the fact that linguistic choices are neither accidental nor can they be memorized as well as that language is a construct with social and

cultural dimensions. A final implication for language teaching and assessment, which expands the context of the KPG exams, is that language teachers should teach their students how to recognize task goal and text purpose because this ability is likely to help students use grammar and vocabulary appropriately, thus being effectively equipped for success not only in language exams, but also in real life interactions.

**Author's email:** elfie1903@yahoo.com

---

### Notes

1. KPG stands for 'Kratiko Pistopiitiko Glossomathias'.
2. In May 2011 an integrated B level exam was introduced for the first time, which consists of three activities, with an equal number of B1 and B2 questions and tasks. A sample of the new integrated B level test can be found at: <http://rcel.enl.uoa.gr/rcel/news.htm>
3. The illustrative descriptors for language production specify what the candidates are expected to be able to do with language, depending on their level of proficiency. Information about the KPG illustrative descriptors for oral production at levels B1 and B2 can be found at: <http://rcel.enl.uoa.gr/rcel2/texts/B1-Specs.pdf> / <http://rcel.enl.uoa.gr/rcel2/texts/B2-C1-Specs.pdf>
4. The scientific head of this project is Prof. Dendrinou. The project was launched in 2007-2008 by the University of Athens within the frame of a larger research project, entitled "Exam system and language learning certification" (also known by the Greek acronym SAPIG), which was funded by EPEAK, EKT (European Community Fund) and national resources. Since January 2009, the project has evolved as autonomous scientific work, involving experts from various disciplines. More information about the project can be found at: [http://rcel.enl.uoa.gr/sapig/gr\\_ereuna\\_01\\_glosso.htm](http://rcel.enl.uoa.gr/sapig/gr_ereuna_01_glosso.htm).
5. In the one-sided talk (i.e. Activity 2), task response is based on a set of visual prompts that accompany the tasks. In Activity 3, which is also known as 'mediation', the test takers are asked to respond to the task, by relaying relevant information found in a Greek text to an interlocutor who does not understand / speak Greek.
6. According to Eggins (1994, p.179, 187), *modalization* (or else, 'epistemic modality') expresses how likely something is to happen or how frequently something happens. It can appear through modal operators (e.g. might) and mood adjuncts (e.g. possibly). Modulation (or else, 'deontic modality') expresses obligation / necessity (something should(n't) be done) or inclination (how willing I am to do something).
7. Κονδύλη & Λύκου (2009) provide a complete description of each one of all these lexicogrammatical features, by offering examples and commenting on the theoretical framework they fit into.

### Acknowledgement

I would like to sincerely thank Professor Dendrinou for urging me to participate in this project, whose development and completion is mainly due to her fervent enthusiasm, diligent work and guidance.

### References

- Bygate, M. (1999). 'Quality of language and purpose of task: patterns of learners' language on two oral communication tasks'. *Language Teaching Research* 3/3: 185-214.
- Bygate, M. (2009). 'Teaching and testing speaking'. In M. H. Long & C. J. Doughty, (Eds.), *The handbook of language teaching*. West Sussex, Malden: Blackwell Publishing Ltd., 412-440.
- Council of Europe, (2001). *Common European framework for reference for languages: learning teaching and assessment*, Cambridge: Cambridge University Press.

- Dendrinos, B. (2007). 'Information booklet: the Greek state examination system to certify foreign language proficiency.' Athens: RCEL. Accessed at: <http://rcel.enl.uoa.gr/rcel2/texts/B1-Specs.pdf> and <http://rcel.enl.uoa.gr/rcel2/texts/B2-C1-Specs.pdf> on 20 Feb 2010.
- Eggs, S. (1994). *An introduction to systemic functional linguistics*. London: Pinter.
- Fulcher, G. & Davidson, F. (2007). *Language testing and assessment. an advanced resource book*. New York: Routledge.
- Halliday, M.A.K. (2002). *On grammar*. London: Continuum.
- Halliday, M.A.K. & Hassan, R. (1976). *Cohesion in English*. London: Longman.
- Halliday, M.A.K. & Hassan, R. (1985). *Language, context and text: aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.
- Halliday, M.A.K. & Mathiessen, C.M.I.M. (1999). *Construing experience through meaning*. London: Cassell.
- Halliday, M.A.K. & Mathiessen, C.M.I.M. (2004). *An introduction to functional grammar (3<sup>rd</sup> ed.)*. London: Arnold.
- Huang, J. & Mohan, B. (2009). 'A functional approach to integrated assessment of teacher support and student discourse development in an elementary Chinese programme'. *Linguistics and Education*, 20: 22-38.
- Hughes, C. (2009). 'Assessment as text production: drawing on systemic functional linguistics to frame the design and analysis of assessment tasks'. *Assessment and Evaluation in Higher Education*, 34/5: 553-563.
- Hunta, A. Kohonen, V. Kurki-Suonio, L. Luoma, S. (1997). *Current developments and alternatives in language assessment*. Jyväskylä (Finland): University of Jyväskylä.
- Karavas, K. (2009). *The KPG speaking test in English: a handbook*. Athens: RCEL.
- Knapp, P. & Watkins, M. (2005). *Genre, text, grammar. technologies for teaching and assessing writing*. Sydney: UNSW Press.
- Κονδύλη, Μ. & Λύκου, Χ. (2009). «Γλωσσολογική περιγραφή των θεμάτων του ΚΠΓ: Η οπτική του κειμενικού είδους και της λεξικογραμματικής». *Research Periodical*, 1-17, accessed at: [http://rcel.enl.uoa.gr/periodical/article\\_periodical\\_01.pdf](http://rcel.enl.uoa.gr/periodical/article_periodical_01.pdf) on 12 May 2011.
- Kormos, J. (1999). 'Simulating conversations in oral-proficiency assessment: a conversation analysis of role plays and non-scripted interviews in language exams'. *Language Testing*, 16/2: 163-188.
- Long, M.H. and Doughty, C.J. (Eds.) (2009). *The Handbook of language teaching*. West Sussex, Malden: Blackwell Publishing Ltd.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Martin, J. (1998). 'Linguistics and the consumer: the practice of theory'. *Linguistics and Education*, 9/4: 411-448.
- Milanovic, M. & Saville, N. (1996). 'Introduction'. In M. Milanovic & S. Saville (Eds.), *Performance testing, cognition and assessment: selected papers from the 15<sup>th</sup> language testing research colloquium*. Cambridge: Cambridge University Press, 1-17.
- O'Sullivan, B. (2008). *Modelling performance in tests of spoken language*. Frankfurt: Peter Lang.
- O'Sullivan, B., Weir, J.C. & Saville, N. (2002) 'Using observation checklists to validate speaking-test tasks'. *Language Testing*, 19/1: 33-56.
- Pavlou, P. (1997). 'Do different speech interactions in an oral proficiency test yield different kinds of language?'. In A. Hunta, V. Kohonen, L. Kurki-Suonio & S. Luoma (Eds), *Current developments and alternatives in language assessment*. Jyväskylä (Finland): University of Jyväskylä, 185-201.
- Shohamy, E. (1988). 'A proposed framework for testing the oral language of second/foreign language learners'. *Studies in Second Language Acquisition*, 10: 165-179.

- Young, R. (1995). 'Conversational styles in language proficiency interviews'. *Language Learning*, 45/1: 3-42.
- ΥΠΕΠΘ. (2003). *Κρατικό πιστοποιητικό γλωσσομάθειας: ενημερωτικό έντυπο*. Αθήνα: Υπουργείο Εθνικής Παιδείας και Θρησκευμάτων.