

# Strategies for Eliciting Language in Examination Conditions

**Mark Griffiths<sup>1</sup>**

*Trinity College London*

There is a long tradition in oral language examinations of employing structured interviews with pre-scripted, fixed examiner contributions. This chapter examines an alternative format, the semi-structured interview, in which examiners adapt their test plans around candidates' input, in order to elicit target language and communicative skills related to a chosen level. The Trinity College London Graded Examinations in Spoken English (GESE) are the subject of study. Formally calibrated to the Common European Framework of Reference (Council of Europe, 2001), GESE examinations are set out in 12 levels, which Trinity refers to as 'Grades' (e.g. 'Grade 1' to 'Grade 12'.) By analysing video recordings of 32 spoken examinations from the 12 GESE levels, this research identifies a range of examiner techniques that have evolved as a result of the examiners' training and experience in adapting test plans to elicit target language skills. 14 examples of different elicitation techniques were identified from 374 minutes of recorded material. Of particular interest is the finding that contrary to what we might expect from oral proficiency testing interviews, direct questions are only used part of the time, with examiners adapting and developing prompts, creating elicitation strategies that represent a range of conversation patterns and roles.

Key words: semi-structured, oral, examination, elicitation, prompt.

## **1. Introduction and context**

The context for this chapter is the oral language examination for non-native speakers of English. Such assessments usually take the form of an interview. Van Lier (1989) asserted that the oral proficiency interview comes as close as possible to being an acceptable surrogate for real-life oral assessment (*ibid*, p. 494). In the early 21<sup>st</sup> century, we live in a time when academic and commercial demand has never been higher for oral proficiency assessments that offer insights into candidates' communicative strengths. However, whilst a range of oral assessments has been developed across the commercial and academic fields, not all assessment interviews are structured in the same way, in particular where examiner role and input are concerned. Many of these oral language examinations are designed around highly-structured interviews in which a series of prompts is delivered to the candidate(s) by the examiner. In such examinations, the examiner prompts are typically direct questions or instructions to the candidate(s) to perform a verbal task. All questions, instructions and examiner interaction are decided upon prior to the examiner and candidate(s) meeting and before the oral communication begins.

---

<sup>1</sup> mark.griffiths@trinitycollege.co.uk

There also now exist oral examinations based solely on pre-recorded examiner prompts which are delivered by computer, without the examiner being present or participating in the conversation that is being generated. Proponents of these pre-written assessment frameworks are likely to point to the benefits of employing this method of examination and interview structure, which may include: economies of scale in terms of test production; potential ease of examination administration; tightly-controlled production-line consistency and uniformity of examiner input and examiner response, regardless of candidate contributions. This uniformity of examiner talk is often identified as a major contribution to an assessment's reliability, based on each candidate being exposed to the same stimulus and response.

Yet the formalised, pre-scripted assessment framework may also be problematic. For example, one may liken the pre-scripted examination to that of an *interview by autocue*, with the examiner's contributions limited to reading the words of a non-present speaker rather than being a full conversational interactant. In turn, the candidate is assigned a passive/reactive role in the interaction, the content and quantity of their input being measured and scaled by the pre-written 'autocue'. The assessment may be standardised in terms of examiner input and response, but this rigid standardisation may also limit the scope of what the candidate is exposed to linguistically and what he/she is permitted to produce in the target language, with inevitable consequences for what linguistic data is available for assessment. Additionally, one might reasonably expect the purpose of an oral proficiency examination to include assessment of a speaker's communicative competences. If we take Canale and Swain's (1980) helpful four-part model of communicative competences - linguistic, sociolinguistic, discourse and strategic competence - we can reasonably expect a candidate's linguistic competence to be assessed in a pre-scripted exam, as a candidate should demonstrate control of lexis and grammatical structures. However, it is not immediately easy to see how a candidate's sociolinguistic, discourse and strategic competences can be viewed or assessed if an autocue examination does not build in locally-assembled unplannedness or unpredictability of sequence and outcome (Goffman 1981), or if the candidates do not receive feedback tailored to their contributions, or have limited or no opportunity to shape the development of the discourse. This again raises questions regarding what exactly the assessment process is telling us about the candidate's conversation skills. Morrow (1979) suggested that reliability should be 'subordinate to face validity' and given the weaknesses identified above, one may ask whether the pre-scripted examination leaves itself open to accusations of prioritising reliability over validity and whether pre-scripted oral assessments are valid assessments of real-world conversation skills.

Some examinations attempt to avoid potential threats to construct validity by including paired-speaking tasks in assessments, in which the candidates may have some scope for generating a short exchange. However, their candidate

interlocutor is a fellow test-taker, usually of the same approximate learner level and often of the same L1, meaning that where the range of communicative competences might be tested, they are done so in relation to speakers who may have similar learner and error profiles rather than with a native speaker or non-native speakers from a different language or culture, again raising doubts regarding what the candidates' English output is evidencing. Also, the exchanges in paired speaker tasks are usually limited in turns available to the candidates, with the content, aims and timing of the task being controlled by the examiner. This has implications for the candidates' opportunity to display discourse and strategic competence.

Exactly what a pre-scripted assessment tells teachers, parents, employers and the candidates and whether this represents an assessment of real-world linguistic skills is not a matter for this chapter. Likewise it is not the concern of this chapter to offer a detailed critique of the strengths and weaknesses of formal, pre-scripted tests. We can only surmise that oral proficiency assessments do not all have the same interview structure or test purpose, which sets the scene for exploring an alternative exam format. This alternative is the focus of the following section and of the research in this paper.

## **2. Focus of the study**

This section and the rest of the chapter focus on the semi-structured interview. The purpose of this examination format remains the same as the formal, structured interview, aiming as it does at eliciting language and communication skills of a specific syllabus. However, the conversation format allows for the interviewer to discuss groupings of topics and questions with candidates in a range of ways. The example semi-structured examination under investigation is Trinity College London's Graded Examinations in Spoken English (GESE). The examination format is a 1-1 unscripted conversation with an examiner, within which the candidate demonstrates specific language items and his/her mastery of the target language and skills of the chosen level as set out in the examination syllabus. The GESE examinations are set out in 12 levels, which Trinity refers to as 'Grades' (e.g. Grade 1 to Grade 12.), all of which have been formally mapped to the Common European Framework of Reference (Council of Europe, 2001; Papageorgiou 2007). Grade 1 is calibrated at CEFR pre-A1, with Grade 2 at CEFR A1 and the scale continuing to Grade 12, which is calibrated to C2. The Grade 1 level of the examination is 5 minutes long with the examination increasing in length to 25 minutes at Grades 10-12 (C1-C2). To briefly summarise what is expected of the candidate, throughout the 12 levels of the examination, a Conversation Phase is included, which in lower levels is centred on the exchange of basic personal facts and information, extends into a discussion of 6 listed conversation subject areas by A2.2 on the CEFR, to stating preferences and opinions by B1, and the candidate discussing their relationship to, experience and opinions of conversation subject areas listed in the syllabus from B2 and

above. These conversation subject areas are intended at lower levels to allow candidates to talk about concrete every-day themes, and as candidates progress towards C2 on the CEFR, the subject areas generate more discursive conversations, with candidates performing a range of language functions, including reporting and contrasting ideas and opinions and developing, justifying and defending points of views. Additional tasks are introduced at various levels, with the introduction of the Topic Phase at A2.2 in which candidates are expected to prepare a topic to discuss. At B2, an Interactive Phase is introduced, requiring the candidate to take more responsibility for initiating, maintaining and driving the interaction, with the candidate finding out further information and making comments in response to an examiner prompt. At C1, the candidate's chosen topic becomes the focus of both a Topic Presentation Phase (a formal monologue presentation on a discursive theme with appropriate counter-points and conclusions) followed by a Topic Discussion Phase in which the examiner and candidate discuss the topic and the candidate responds to questions and challenges from the examiner.

With regard to interaction, the GESE examination suite is based on enabling the candidate to 'participate in a genuine two-way exchange within the linguistic limits set by the syllabus' (Trinity College London 2009, p.11). The unscripted nature of the examination means that either participant – the candidate or the examiner – is able to contribute at any point, rather than turns being fixed by a pre-scripted examination format. The bidirectionality of the interaction is underlined by the expectation that from A1 upwards, candidates will ask the examiner questions. This starts at A1 with a request for basic information using 'Where, how, have you got, etc.' and an increasing emphasis is placed on the proactive role of the candidate in the examination throughout the suite. In these respects, the examination is markedly different in structure from the formal, pre-scripted assessments, with the semi-structured interview believed to provide ample opportunity to generate and promote natural, multi-turn conversations. ('Conversation' is a word which is frequently used throughout the GESE syllabus.)

Turning to the role of the examiner, the semi-structured nature of the GESE examination means that the examiner is expected to act as a facilitator rather than an interrogator, aiming to enable a genuine two-way exchange, with either the candidate or the examiner able to contribute at any point. Given this expectation, restricting examiner contributions to the reading aloud of a pre-written script regardless of candidate contribution is seen as contrary to the aims of the examination. Instead, the examiner is expected to use a test plan – a selection of communication strategies and prompts specifically designed to elicit the range of target (grammatical, lexical, functional) language items and communicative skills of a particular level. Examiners have delegated responsibility to utilise the elicitation techniques in their test plans in flexible sequence as the examination progresses, with the conversation being locally assembled. Which items they use

and in which sequence examiners use them depends on the candidate's contributions and which target language and skills they have not yet demonstrated.

An additional aim of the semi-structured format utilised in this examination is that of giving candidates the opportunity to not only demonstrate a range of language items specific to the level being assessed but also to demonstrate communicative competences and language skills that reflect real-world exchanges outside the examination room. All four of the Canale and Swain (1980) communicative competences have been mapped onto the GESE syllabus, with grammatical competence being required from Grade 1; discourse and strategic competences evident from Grade 3 (A2); and sociolinguistic competence evident from Grade 4 (B1) (Wall & Taylor 2011; O'Sullivan, Taylor & Wall 2011).

The absence of examination pre-scripting may lead some to question the reliability of the examination. Trinity College London provides the examiners with specific training on the design of prompts to elicit specific language items/key skills. The guidelines given by Trinity on the creation of prompts (confidential internal documents) strictly direct examiners to locate their prompts within the language of the level being examined, the full range of which must be the focus of the examination. No language from levels above the level being examined may be used. Trinity College London also publishes that it only recruits experienced, well-qualified teachers, with high standards of initial and on-going examiner training. It also carries out live monitoring of examining and uses samples of recorded examinations for monitoring and feedback in its programme of standardisation and training.

### **3. Research aims**

To date, little has been published about the range of examiner prompts and elicitation strategies that arise from this semi-structured interviewing format and there are many lines of enquiry that one could pursue relating to examiner and candidate behaviour. The focus of this study is the use of examiner prompts within the semi-structured interview. Given that examiners have been delegated responsibility for the design and implementation of their test plans, the research described here explores the language-eliciting prompts that examiners have devised through and how these prompts are introduced into the examination. How do the examiners use the space and flexibility provided by the semi-structured format to best elicit the target language and how does examiner talk influence candidate responses in oral proficiency interview tests (Kasper, 2004)? The study is an investigation into real examples of semi-structured interviewing from all levels of the GESE exam suite and addresses the following research question:

- Which prompts do examiners develop and draw on to elicit appropriate candidate contributions in the forms of level-specific language items and communicative skills?

Given the potential for investigation in this field, it is acknowledged that other aspects of the semi-structured interview such as inter or intra-examiner variation, inter-rater reliability, or the possible impact of interactional features on final examination scores could all be valuable lines of enquiry, along with questions regarding reliability or validity of the test format. However, they are beyond the scope of this study. There is also no concern here with Trinity's choice of test specifications, purpose, or format.

#### 4. Methodology

The analysis was carried out on a sample of 32 video-recorded oral examinations taken from the Trinity exam archive. The examinations varied in length according to the Trinity level chosen. Only whole examinations were used, rather than excerpts. All the recordings were filmed between October 2009 and February 2010 and were conducted using the 2010 Trinity GESE Syllabus (which at the time of writing is still current). The number of examinations in each grade, their relationship to the Common European Framework of Reference (CEFR), the length of the examinations, number of samples used and total sample size in minutes per Trinity Grade (i.e. level) are given in Table 1:

Table 1. Details of video samples used

Grade	CEFR	Length (minutes)	Number of samples used	Total number of minutes
12	C2	25	2	50
11	C1	25	1	25
10	C1	25	2	50
9	B2.3	15	1	15
8	B2.2	15	2	30
7	B2.1	15	1	15
6	B1.2	10	5	50
5	B1.1	10	4	40
4	A2.2	10	4	40
3	A2.1	7	3	21
2	A1	6	3	18
1	A0	5	4	20
TOTAL:			32	374

A total of 7 examiners appeared across the 32 examinations. It was not possible to use an even number of samples from each examiner due to the nature of filming and other practicalities. Rather than round the number of samples for each examiner down to the lowest equivalent number, it was decided that the greater the range of samples analysed, the greater the likelihood of capturing a

range of examiner prompts. All examiners in the sample data have been working for Trinity for between 5 and 10 years in the role of International ESOL examiners and can be regarded as highly experienced in their roles. The candidates varied in age from 7 years old for the Grade 1 exam, up to 33 for one of the candidates taking Grade 12. The numbers of samples from each examiner are given in Table 2:

Table 2. Numbers of samples from each examiner

Examiner	Number of examinations	Total time in minutes
1	9	110
2	9	92
3	6	83
4	4	26
5	3	23
6	1	25
7	1	10

The focus of the analysis was the examiners' contributions and the candidates' responses as part of the duologue. All of the recordings were listened to by the researcher and examiner prompts were noted. These included prompts that appeared to come directly from the examiner test plans (which were analysed separately) and prompts that seemed to have been adapted from the test plan as the examination was proceeding, to extend the candidate's contributions or elicit specific language items or communicative skills. It was anticipated that as the focus of each exam was different, depending on the level being taken, there would be variation in the types of prompts elicited.

Phonetic and other features such as timing and pausing were also not noted here as it was not felt that they would add anything further to our understanding of the choice of examiner prompts.

The examples of examiner prompts were analysed manually and then categorised according to prompt type. The labels chosen for each category reflected the strategy apparently being employed in the use of the prompt. Finally, sample tokens of each examiner strategy were tabulated to illustrate the range and types of examiner prompt used. These results are presented in the next section.

## 5. Results and discussion

To begin with, the reader's attention is drawn to the fact that the data should only be described and evaluated in the context of a language examination, in which a live examiner (rather than a computer) has the aim of eliciting specific language items and communication skills from a candidate in a one-to-one context, and in a limited time frame. Comparisons should not be drawn with classroom teacher talk, the purpose of which serves a wide range of other functions (explanatory,

checking understanding, setting up activities, giving feedback, etc.). Examiner talk should be seen as a distinctive type of discourse.

Returning to the data, the 207 examples of examiner prompts emerging from the data were reduced to 14 types distinguished by their structure and apparent examiner aim. The 14 types did not seem to fall into any kind of level-related pattern, and as the study was not able to access identical sample lengths from each of the 12 levels, and with a relatively low example prompt-to-prompt type ratio, frequencies of occurrence are not provided here as there would be little reliable statistical inference to be drawn from such data. However, the 14 different example types drawn from across the available data are listed below along with the next response from the candidate and it is the range of example types and their responses that are the focus here.

#### Example prompt types in the data:

(1) *Examiner asks a direct question using incorrect information*

E: Are they red socks?

C: No, they're blue socks (Grade 1 – A0)

(2) *Examiner models the answer and then prompts for full sentence response*

E: My birthday's in July. And yours?

C: My birthday's in July too (Grade 3 – A1)

(3) *Examiner offers a basic statement*

E: I finish work at 5.30.

C: I finish work at 3.30.

(4) *Examiner uses a direct question re: candidate ability*

E: What jobs can you see?

C: He's a policeman. She's a nurse. She's a doctor. (Grade 3 – A2.1)

(5) *Examiner uses a direct question to encourage comparisons*

E: What's the difference between a friend and a best friend?

C: A best friend is more important than a friend simple. (Grade 4 – A2.2)

(6) *Examiner presents an open instruction*

E: Tell me about your interest in cars

C: I've been interested in cars maybe since I was 14-15 years ago. Yes, my father used to teach me to drive. Yes I used to sit in him and used to hold the steering wheel. (Grade 5 – B1.1)

(7) *Examiner elicits a description, avoiding closed question*

E: How do you describe your style?

C: I don't have a style. I clothes, I like it and I buy it. But I'm not looking for it so much. (Grade 6 – B1.2)

(8) *Examiner uses a direct question to elicit an evaluation, avoiding a closed question*

E: How important is money for you?

- C: Not very. I need money to live, but I don't need to be a rich. (Grade 6 – B1.2)
- (9) *Examiner presents a simple statement of fact*  
 E: Some people do shopping on the internet.  
 C: I surf the net, use MSN, Facebook. But I don't buy anything. And you? (Grade 6 – B1.2)
- (10) *Examiner questions the previous statement, encouraging expansion*  
 E: Really?  
 C: Yes, and in the future, they will all be touch screen phones. (Grade 7 – B2.1)
- (11) *Examiner expresses an opinion as a statement, but without a question*  
 E: I sometimes wish I'd been born later.  
 C: Me too. I think it would be great to be born a few years ago. For example I would be able to grow up with modern technology. (Grade 9 – B2.3)
- (12) *Examiner appears to feign lack of knowledge, encouraging candidate to bridge the knowledge gap*  
 E: I don't know much about racism.  
 C: Well, it's an issue that has been viewed quite negative by people in the UK and Greece with immigrants in both countries experiencing many problems. (Grade 12 – C2)
- (13) *Examiner appears to deliberately mis-paraphrase candidate's statement*  
 E: So you're basically saying that immigrants work in conditions and for salaries that are damaging to the economy.  
 C: No, that's not what I'm saying. I'm saying that people might think that immigrants are not healthy for a country's economy. (Grade 12 – C2)
- (14) *Examiner stays silent. The candidate must continue the conversation.*  
 E:  
 C: Nevertheless, it's not their decision to make. People should make their own choice about voting. What do you think? (Grade 12 – C2)

A number of issues emerge from this categorisation exercise. The first has to do with a notion that many would take for granted; that candidate language in an oral examination is always a response to examiner questions. It is clear that this is not the case in this examination, where trained and experienced examiners are allowed to vary their elicitation strategies to elicit target language and communicative skills. For example, whilst questions appear in Examples 1, 4, 5, 7 and 8, non-question form elicitation strategies exist in Examples 2, 3, 6, 9, 10, 11, 12, 13 and 14. While the data do not represent statistical tendencies or frequency of use, it is interesting to note that professional elicitation strategies in a semi-structured interview consist of more than simply a list of questions or instructions. This range of elicitation strategies is discussed next.

The first elicitation strategy to discuss is that where questions were used, they were always direct questions. Indirect questions were absent from the whole data set – not just from the 14 examples listed above. We know that in conversations between native speakers of English, indirect questions are used commonly as a

pragmalinguistic device to attend to listeners' face needs and convey appropriate levels of politeness. Yet there were no instances of indirect question requests such as 'Can you tell me..', or 'Could you explain..', even at the higher levels. The reason for this is unclear, but one possible explanation may be that the examiner in effect modifies his/her expected level of candidate sociopragmatic comprehension (Ross 1995). That is, the examiner may make no assumptions that candidates will understand the sociopragmatic meaning of the indirect question form and they may even suspect that it might lead to candidate confusion. The examiner performs downward linguistic accommodation, avoiding complex sentences and politeness forms to facilitate listener comprehension of form over pragmatic intent. Whether this is an overt strategy by the examination board or overspill of the examiners' own teaching experience into the examination room would be a matter for future investigation.

Alternatively, the absence of indirect question forms may be a function of the examiner using direct questions for particular effect: Example 6, 'Tell me about..' not 'can you tell me' may be seen as a more powerful and efficient method for eliciting the required target language and maximising candidate talking time in a limited time-frame assessment. The use of such grammatical forms to communicate force in an L2 setting has also been observed by Kasper (2004). What is interesting about the use of direct questions is the range of communicative purposes: Example 1 requires the candidate to correct the examiner; Example 4 requires a list of vocabulary; Example 5 encourages comparisons; Example 7 elicits a description and Example 8 elicits an evaluation.. The examiners' prompts also require a response from the candidate that specifically utilises the language of the level. For example, in Grade 1, candidates are expected to know colours, items of clothing, and adjective/noun word order in English. Likewise, Example 5 is taken from Grade 4, in which the candidates are expected to make comparisons and use comparative adjective and adverb. We see from such examples that direct questions are used by examiners in combination with tasks related to specific target language items, structures and functions. We see also that although the questions themselves may appear relatively closed, the candidate is free to choose how they wish to answer: they can often personalise the response and they can choose which aspect of the information they wish to communicate.

A second elicitation strategy present in the data is that of encouraging elaborations through direct requests for expansion. In Example 6, with the candidate having already mentioned their interest in cars, the examiner responds with a 'Tell me about..' prompt to encourage expansion. Example 8 sees the examiner use the exclamation 'Really?' to express interest or surprise. Of note is that the examiner's exclamation is not followed by a further examiner remark, indicating that the examiner has given the floor to the candidate to elaborate.

A third elicitation strategy, seen in Examples 3, 9, 11, 12, is for the examiner to make various statements to encourage elaboration. Nakatsuhara (2007) has

also observed how statements are effectively used by interviewers as implicit demands for more information, often resulting in a candidate elaborating their response and this strategy appears in various forms: Example 3 is a basic short statement, as is Example 9, which embeds a personal opinion or sentiment into the statement; however, the examiner invites contrast by making a statement about only *some* people. In Example 12, the examiner appears to feign lack of knowledge, encouraging the candidate to bridge the knowledge gap. It appears that examiner statements are pragmatically understood as an invitation to elaborate or expand and take the floor, in particular when the examiner displays a lack of knowledge or incorrect knowledge of a subject.

A fourth example of an elicitation strategy, seen in Example 13, shows the invitation to elaborate taking a sociopragmatically different form, with the examiner using reformulation (the re-phrasing of what has been said). Heritage (1985, p. 115) compares L2 reformulations with those used in news interviews, pointing out that reformulations are understood as alternatives to going on to a next question and are routinely treated as invitations to elaborate further. In Example 13, the examiner follows up a previous statement with a reformulation of the candidate's words, 'So you're basically saying that immigrants work in conditions and for salaries that are damaging to the economy.' This seems to be a deliberate misrepresentation of the candidate's words or opinions. This is immediately corrected by the candidate, who then elaborates, which was clearly the examiner's aim.

A fifth and final elicitation strategy and perhaps the most sociopragmatically extreme version of elaboration is the use in Example 14 of silence. Jaworski (1997) points out that we need to go beyond the understanding of silence as an 'absence of sound'. LoCastro (1995) describes silence as a linguistic resource to signal pragmatic inferences in interactional contexts. In the context of language examining, it is clear that the examiner is choosing silence as an intentional strategy for cueing expansion from the candidate, the pragmatic inference being that the candidate has not satisfied what Grice (1975) would identify as the maxim of quantity. Interestingly, it was noticeable across all the data that silence was used by examiners in examinations at B2 and above. This may be related to the CEFR B2 descriptors which characterise a B2 candidate as being able to 'initiate discourse, take his/her turn when appropriate.' In Example 14, we see that this was effective in encouraging the candidate to continue the turn and expand on his/her point, inviting comment from the examiner.

## **6. Concluding remarks and future research**

It seems appropriate at this point to do some redefining. Throughout this study I have referred to examiner prompts, which have been shown to be broadly divided into direct questions, statements and indirect cues for elaboration. However, given that silence is also in evidence in the data it would seem more accurate to

summarise these as *elicitation techniques*, of which prompts are one part and silence is another<sup>2</sup>. The examples presented illustrate a range of examiner techniques that were found in the data. We see examples of direct questions, direct instructions, examiner statements, examiner false statements for correction, encouraging comparisons, eliciting descriptions, eliciting evaluations, feigning lack of knowledge, encouraging contradiction, encouraging expansion, deliberate mis-paraphrasing in reformulations and silence.

Based on this evidence, there would appear to be a case for further investigating how semi-structured examinations may be employed to generate candidate language and whether they afford the candidate an enhanced opportunity to display a range of communicative skills and core linguistic competences. It would appear from the type of interaction present in the data that a semi-structured exam format that utilises a range of elicitation techniques may widen the aperture through which the examiner views the range of the candidates' communication skills, enhancing the validity of an assessment of communication skills.

It is clear from the data presented here that the semi-structured interview used in this suite of examinations has encouraged a range of examiner elicitation techniques to be introduced into the assessment room. The data do not give us insight into any relationship between examiner experience and range of prompt type and we cannot give any account of the training given to the examiners by the examinations board. Yet it is hard to resist the conclusion that this range of examiner techniques represents a broader and arguably more authentic range of oral examiner input than structured interviews and this could be the subject of further research. A further potential research strand could be to investigate candidates' and examiners' perceptions and attitudes towards semi-structured examinations.

It would also be useful to explore whether semi-structured examinations encourage candidates to maximise their contributions. Rather than restricting the candidate to a passive, reactive role, following a set of instructions, answering questions only or making statements which are not responded to, does the semi-structured format encourage the candidate to participate more fully in exchanges, demonstrate more conversational control and even take more responsibility for driving the conversation? In addition, it would be interesting to observe whether certain prompt types are more commonly used in particular levels and whether there is any variation in the frequency of use of certain elicitation techniques within or between levels. We also do not have clear any data regarding how effective different elicitation techniques may be at eliciting language specific language items or communicative skills. If so, what are the implications for

---

<sup>2</sup> In other examination contexts, we might choose to include body language as a further elicitation technique.

providing examiners with an improved vantage point from which to view candidates' language skills?

Future research could also compare the effectiveness of the examiner techniques listed above with structured, pre-scripted interviews where the examiner is unable to divert from the list of questions and instructions. Are examination formats that prescribe test items beforehand and do not allow examiners flexibility, testing with the same purpose as examinations in a semi-structured format? Can the same range of language and communicative skills be elicited using a prescribed question/instruction exam format as can be by using a semi-structured format? Are the elicitation techniques of structured and semi-structured examinations natural or realistic? Do the elicitation techniques in both formats ever occur in natural conversation? Do these elicitation techniques represent the types of interactions that candidates will encounter in the real world and are these qualities in an examination that are worth striving for? If you introduce the data to a group of researchers, rather than the single researcher in this study, will they identify a wider range of elicitation techniques than those identified in this research? It would be interesting to hear how other examination boards respond to the data here regarding the effective use of test prompts and what this tells us about styles of oral language examining.

It would also be useful to investigate the pragmatics that underpin semi-structured interviews when the format of the scripted interview is removed and conversation is more fluid. What strategies do candidates use? What happens when communication breaks down? Do candidates rely on a range of strategies that differ from those used in a scripted and structured language assessment? Which examiner prompts and elicitation techniques are favoured by examiners using a semi-structured format or seen as more effective at each grade? It would be useful also to collect qualitative data from examiners regarding when and why they use the range of strategies, techniques and prompts at their disposal, and what the process is in finalising the prompts they include in their test plans. Do they have preferred techniques that they see as particularly effective? How can we ensure reliability in a non-prescribed question format? Does a semi-structured interview support or hinder a teacher's communicative classroom aims?

In a small-scale study such as this it is clear that the data set is not sufficient to answer many of these questions. However, as the first study of its type in this area, one can hope that enough light has appeared through the cracks to illuminate examiner practice and the types of techniques used in such semi-structured interviews. It would be a positive step if this study stimulated further research interest into an area which has to date been under-investigated. It appears that there are multiple elicitation routes and techniques that can be utilised during an oral interview to elicit target language and communicative skills, and there is much potential for developing new techniques and informing new practices.

## References

- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning Teaching, Assessment*. Cambridge: Cambridge University Press.
- Goffman, E. (1981). *Forms of talk*. Oxford: Basil Blackwell.
- Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds). *Speech Acts*. (pp. 41–58). New York: Academic Press.
- Heritage, J. (1985). Analyzing news interviews: Aspects of the production of talk for an overhearing audience. In T. van Dijk (Ed.) *Handbook of Discourse Analysis* (Vol. 3, Discourse and Dialogue, pp. 95-117). London: Academic Press,
- Jaworski, A. (1997). Introduction: an overview. In A. Jaworski (Ed.) *Silence: interdisciplinary perspectives* (pp. 3-14). Berlin: Mouton de Gruyter.
- Kasper, G. (2004). Speech acts in (inter)action: repeated questions. *Intercultural Pragmatics* 1(1), 125-133.
- LoCastro, V. (1995). Second language pragmatics. In E. Hinkel (Ed.) *Handbook of Research in Second Language Teaching and Learning* (Vol. 2, pp. 319-344). New York: Routledge.
- Morrow, K. (1979). Communicative language testing: Revolution or evolution? In C. J. Brumfit, & K. Johnson (Eds.), *The communicative approach to language teaching* (pp. 143-159). Oxford: Oxford University Press.
- Nakatsuhara, F. (2007). Inter-interviewer variation in oral interview tests. *ELT Journal* 62(3), 266-275.
- O’Sullivan, B., Taylor, C. & Wall, D. (2011). Establishing evidence of construct: A case study. Presentation delivered at the European Association of Language Testing and Assessment (EALTA) conference, Siena, May. Retrieved 15 February 2012 from <http://www.ealta.eu.org/conference/2011/programme.html>
- Ross, S. (1995). Aspects of communicative accommodation in oral proficiency interview discourse. *Unpublished PhD dissertation*. University of Hawaii at Manoa.
- Papageorgiou, S. (2007). *Relating the Trinity College London GESE and ISE examinations to the Common European Framework of Reference. Piloting of the Council of Europe Draft Manual*, final project report. London: Trinity College London, retrieved on February 24th 2012 from [www.trinitycollege.co.uk/resource/?id=2261](http://www.trinitycollege.co.uk/resource/?id=2261)
- Trinity College London (2009). *Graded Examinations in Spoken English (GESE) Syllabus from 1 February 2010*. London: Trinity College London.
- van Lier, L. (1989). Reeling, writhing, drawling, stretching, and fainting in coils: oral proficiency interviews as conversation. *TESOL Quarterly* 23, 489–508.

Wall, D. & Taylor, C.). Communicative Language Testing (CLT): Reflections on the 'issues revisited' from the perspective of an examinations board. Forthcoming.