

Speaking Test Materials: Let's give them something to talk about!

Julie Norton, The University of Leicester, UK

Introduction

Previous articles in *Folio* (Rinvolucri, 1999, Thornbury, 1999) have critically examined the topical content of ELT coursebooks and have raised questions about the types of topics that should be included to challenge and motivate students to participate in class. More generally, the cultural content of ELT coursebooks has been debated (Bell and Gower, 1998; Gray, 2000), and it is widely recognised that no teaching materials can achieve the perfect 'fit' and appeal to all learners and teachers globally (Maley, 1998). Reshaping and reinterpreting texts in the classroom are viewed as "a key element in the construction of new meanings and in the creation of the culture of the classroom" (Gray, 2000:275).

What about oral testing materials? Whilst teachers can adapt or omit the activities in coursebooks which they deem to be culturally inappropriate or irrelevant for their learners, examiners in speaking tests working under real-time constraints and with scripted interlocutor rubrics do not have this flexibility. Indeed, adapting materials in a testing situation could constitute a breach in examiner conduct and result in major problems for test standardisation. Candidates in speaking tests are thus placed in a position where they must either think of something to say about the chosen topic, or risk negative assessment of their oral ability in English. After all, a "will-not talk" candidate may easily be confused with a "cannot talk" candidate, as van Lier points out (1989: 501).

In light of the above comments, the title of a recent book on the validity of the oral proficiency interview as a measure of conversational ability in English, "The Art of Non-Conversation" (Johnson, 2001) may come as no surprise; and may possibly even raise a smile and may resonate particularly with oral examiners who have struggled to elicit a language sample suitable for assessment purposes from reticent candidates. More seriously, this pithy title could quite rightly set alarm bells ringing and provoke debate about the rigid frameworks imposed upon candidates in speaking tests, call into question the relevance of the topical content of these tests, and cause concern about the validity and reliability of the oral proficiency interview

as a test of spoken English. This position is ratified by Young and Milanovic's (1992: 421) research on the Cambridge Speaking Tests: 'The oral proficiency interviews that we have examined here bear very little resemblance to the collaborative management of talk by both parties that we believe to be the structure of non-testing situations.'

This article aims to investigate the topical content of speaking test materials, and to comment upon their cultural appropriateness. The article focusses on the performance of Japanese candidates in the following Cambridge Speaking Tests: First Certificate in English (FCE) for intermediate level learners; Certificate in Advanced English (CAE) for upper-intermediate level learners; and Certificate of Proficiency in English (CPE) for advanced level learners. Firstly, background information on the study is presented. The topics included in the Speaking Test materials are reviewed, then examples of topics which appear problematic for Japanese candidates of these tests are critically discussed with reference to the interview data. Some of the topics initiated by the candidates themselves during the Speaking Tests are then presented to examine to what extent these topics coincide with those prescribed in the actual test materials. Finally, implications for the development of oral testing materials are explored.

Background to the Study

Data Collection and Description

The Speaking Tests were recorded during examining sessions held in the United Kingdom in 1995-1996. The FCE data were videotaped by the test administrator, the University of Cambridge Local Examinations Syndicate (UCLES) during a piloting of the revised FCE Speaking Test. I audiotaped the CAE and CPE interviews. Ten FCE interviews, ten CAE interviews and seven CPE interviews are included in the data sample discussed here. Candidates were informed that the Speaking Tests were being recorded for research purposes. As an experienced oral examiner, I can verify that this did not appear to affect candidate performance in the tests. Data were transcribed according to the transcription conventions presented in Psathas (1995).

Format, Tasks and Assessment in the Speaking Tests

At the time of data collection, two examiners and two candidates participated in the FCE and CAE interviews, but a single examiner and individual candidate was the common format for the CPE Speaking Tests (only one example of a paired CPE Speaking Test features in this data sample). It is worth noting that the CPE Speaking Test was revised in June 2003 and has now adopted a paired format in line with the other Speaking Tests.

Each Speaking Test consists of an Introduction, three separate tasks and a closing. The Introduction is a warm-up phase, which allows the examiner(s) to find out some personal information about the candidate(s), and is intended to help candidates relax. The information elicited during this stage can be used to select an examination pack appropriate to the candidates' interests. Many oral examiners, however, select the examination pack before meeting the candidates for ease of organisation, working through the packs in the sequential order in which they appear, and omitting packs which they dislike. After the Introduction, candidates are asked to comment upon photographs. This task is intended to give individual candidates the opportunity to take a long turn – candidates are allowed one minute to compare and contrast photographs on a particular theme. In this data sample, the tasks in the CPE Speaking Tests are thematically related, but not at FCE and CAE levels. In the FCE and CAE Speaking Tests, candidates are requested to perform a two-way collaborative task in the third stage of the test. CPE candidates perform similar ranking activities in this stage of the test. The final stage of each Speaking Test involves a Discussion between examiner(s) and candidate(s).

Candidates in this study were assessed on the following scales: fluency; grammatical accuracy and range, pronunciation, task achievement (CAE only), interactive communication and vocabulary resource.

The Topics

As may be expected, a limited number of "general interest" topics, similar to those identified in Reda's (2003: 260) exploration of a corpus of ELT coursebooks, feature in the Cambridge Speaking Tests included in this data sample. Unsurprisingly, these topics reflect those covered in examination preparatory coursebooks, and are presumably considered 'safe'; that is, culturally inoffensive and accessible to candidates from a diverse range of backgrounds. The following topics feature in the speaking test materials included in this data sample:

Table 1: Topics in the FCE Speaking Tests

Eating out
Jobs

Travel
Leisure time
School improvements
Rooms
Work environments
Youth clubs
Holidays
Strange pictures
Relationships

Table 2: Topics in the CAE Speaking Tests

Crowded places
Feelings
Festivals
Funding worthy causes
Weddings
Women's jobs
Work experience
Protective clothing
The seaside
Social problems
Service station facilities

Table 3: Topics in the CPE Speaking Tests

Freedom and responsibility
Culture
Tourism
Education

As the above tables reflect, CPE candidates are expected to be able to discuss a range of more abstract and academic topics than candidates in lower level Speaking Tests.

Problems for Japanese Candidates

Youth Clubs

The topic of youth clubs causes problems for two female Japanese FCE candidates who seem slightly confused about the concept of the youth club presented in the photographs and have difficulty equating this with the type of clubs organised for young people in Japan. This leads to some "uncomfortable moments" (Erickson and Schultz 1982: 104; Lazaraton 1991: 24), as participants appear to operate with differing cultural assumptions of a youth club, and the examiner possibly interprets the candidates' reticence as unwillingness to participate fully in the test. The candidates are presented with photographs of various activities and facilities (including sports, games, a library, a café, a computer room, a disco) that can be incorporated into a new youth club, which is to be built by the local council. Candidates are then asked to choose the three most popular or useful facilities or activities for young people to include in the youth club. The candidates' reticence in this task may result from their unfamiliarity with the type of youth club which is depicted. Clubs for young people in Japan

are normally organised by schools and take place either before or after school. They usually focus on one activity, such as a sport or drawing, and it is compulsory for Japanese school children to attend at least one club activity. This concept of a youth club is not compatible with the photographs of the youth club presented in the examination materials, which incorporate many different activity types together. Neither candidate explains this to the examiner, however, and it becomes difficult for the candidates to respond to the examiner's prompts about how successful this type of club would be in Japan.

Example 1

- 375 E: But how/ how popular/ would a/ a youth centre/ like this/ be/ in Japan <do you think?>
 376 think?>
 378 H.M: Oh? / in Japan
 380 N.K: <Erm>
 382 E: Would it be popular with young people?
 384 N.K: <Yes/ I think> / there are (xxx) sports clubs/ is very popular
 386 H.M: <Yes/ I think so> uhm
 388 E: Uhm
 390 N.K: ... computer school is not popular/ but/ it's useful
 392 E: Uh-huh
 425 E: Okay/ and er/ what age range/ do you think/ it would be popular with/ what kind of age of young person/ would like to go to a youth club
 428 (3.0)
 430 H.M: <Erm>
 432 N.K: <I think/ erm/ maybe university> students
 434 H.M: Yes/ and high school/ and er so on/ yes
 436 E: Students/ what about younger teenagers/ what can younger teenagers do
 438 N.K: Younger teenagers have to study or (hhh) preparing enter high school
 440 E: But you can't study all the time
 442 N.K: (hhh)
 444 H.M: Yes of course/ uhm (FCE.3: L425-444)

The prevalent view in Japan is that final year junior high school and high school students must completely devote themselves to study. Candidate N.K's laugh (L442) may indicate her reluctance to contradict the examiner on this point, and accounts for the initial suggestions that older students could participate in this type of club. This is finally alluded to at the end of the test, but not before considerable fruitless questioning and the examiner's dismay at their lack of elaboration on a topic which is assumed to be within their experience.

Example 2

- 521 E: <Yeah> so: / it seems that only university students have a good time/ yeah
 523 H.M: [Yes (hhh)]

525 N.K: [Yes (hhh)]

527 E: ... the other ones study/ okay/ thank you/ that's the end of the test (FCE.3: L521-527)

The supposedly innocuous topic of youth clubs perhaps suggests how easily cultural bias can enter the speaking test materials.

Strange Pictures

The topic of 'strange pictures' in the two-way collaborative task does not appear to motivate two male Japanese FCE candidates in their early twenties to talk. Indeed, this topic leads to very stilted interaction, characterised by lengthy pauses. The two candidates are asked to talk about what the strange pictures may represent, and then to choose one picture each for their respective homes and decide where to put it. The nature of this task involves candidates closely scrutinising the pictures and this may have a detrimental effect upon the interaction, as eye contact is limited with candidates focussing on the pictures rather than each other. The lexical resources involved in this task also seem quite demanding, as the pictures present abstract and surreal images, such as Dali's time pieces.

Example 3

- 234 Y.Y: Strange (xxx) (5.0) I prefer this one/ because/ this is impossible (1.0) situation
 236 T.S: Ah yeah (hhh)
 238 (4.0)
 240 Y.Y: Yeah/ I also like this
 242 T.S: Chess
 244 Y.Y: Chess
 246 (7.0)
 248 T.S: Yeah I don't know how to play/ er/ there must be/ they must be poster or yes (FCE.4: L234-248).

Given the age of these candidates, this topic may well be beyond their personal experience: they are unlikely to have had to select art for their homes in real-life contexts. In addition, they may have limited experience of talking about art, because this is not a usual topic of daily conversation, and seems far more challenging than some of the other topics included in the FCE materials in this data sample, such as holidays and jobs (see Table 1 above). Indeed, this topic may well have been more appropriate for a CPE level Speaking Test.

Which Topics do Candidates Want to Talk About?

Unsurprisingly, candidates are more likely to develop the interaction when they are allowed to talk about themselves or about topics within their personal experience: topics that they know something about and on which they have views and opinions. Candidate R.M., for example, volunteers personal information about his part-time job in London during the Introduction stage

of the FCE Speaking Test:

Example 4

- 130 E: Wh- what do you do in your free-time
 132 R.M: Er/ I/ (coughs) I teach/ er/ I teach English
 people Japanese language
 134 E: Oh really
 136 R.M: Yeah/ three days/ a week
 138 E: Yes
 140 R.M: ... at the University of London (FCE.5: L130-140)

During these relatively rare moments in the Speaking Tests included in this data sample, we appear to get more 'conversation' and less 'interview'. The importance of such impromptu, spontaneous and natural contributions cannot be overlooked if the aim of a speaking test is to measure conversational ability in English, and it is clear that test developers need to build more opportunities into the test materials for this type of interaction to occur.

Candidate S.K. (CAE.1) produces more coherent and fluent discourse when she is allowed to discuss the familiar topic of Japanese festivals (Example 5), compared to her performance when discussing the photographs depicting 'Feelings' in the one minute extended turn (Example 6):

Example 5

- 486 S.K: Okay/ er/ in my prefecture/ there is a big
 festival/ er/ in the middle of August
 490 And/ er/ we call it (xxx)/ it last three days/ er/
 nowadays/ it's becoming like
 491 a carnival/ because/ er/ it's very new one/ er/ it
 just started after World War Two/ and
 492 er/ people wearing Japanese traditional/ erm/ kind
 of clothes/ it's not kimono/ we
 493 call 'happi' (CAE.1: L486; L490-493).

Example 6

- 270 S.K: this: / okay/
 271 this one/ (hhh) er: / I don't know/ maybe (2.0) they
 are singing/ this is singing
 272 (1.0) or (2.0) or maybe it's admiring someone (hh)
 / but I don't know/ and er/ uhm
 273 (3.0) I think/ er: / maybe they are meeting
 someone/ very special person/ and/ er/ they
 274 wanted to meet some/ and they're now/ and
 finally they met/ and how can I say/ it's
 275 very/ er (1.0) er/ they are very/ glad/ not glad/
 yeah glad/ yes I think so/ uhm (CAE.1: L270-275).

This suggests that allowing candidates to talk about familiar topics can help 'scaffold' their L2 ability, a view which is supported in the literature (Zuengler, 1993). This should certainly be taken into account in the development of speaking test materials to ensure

that fair and accurate testing procedures are adopted.

Table 4 presents some of the other topics initiated by Japanese candidates in this data sample:

Table 4: Topics Initiated by Japanese Candidates in the Speaking Tests

- Japanese culture and traditions
- The Hanshin earthquake
- Japanese musical instruments (the 'koto')
- The Japanese education system
- Japanese holidays ('Golden Week')
- Japanese traditional dress (Japanese 'socks' or 'tabi')
- Japanese literature ('The Tale of Gengi' – the first Japanese novel)
- Difficulties in learning English
- Japanese views on travel abroad
- The sensitivity of Japanese people

Japanese candidates evidently value the opportunity to discuss topics related to their daily lives in Japan. They enjoy informing the examiner about aspects of Japanese culture and traditions which they assume will be unfamiliar. They perceive a genuine communication gap when discussing these topics, because they feel they have to explain what 'tabi' are, or what you do with a 'koto'. This generates more discussion and greater participation from candidates, and this seems a more natural, far less 'painful' way to elicit a language sample suitable for assessment purposes compared to the nerve-racking, one-minute extended turn.

Implications for the Development of Speaking Test Materials

The content of speaking test materials must be given serious consideration if we are to elicit a greater number of 'conversational' contributions from candidates. Young and Milanovic (1992), in their investigation of the Cambridge Speaking Tests, found that topics related to non-native speakers' personal experience of work were sustained longer than more general topics, such as 'learning' or 'having a good time'.

The Cambridge Speaking Tests allow time in the opening stage for the exchanging of names and general introductions. Lazaraton (1991: 117) claims that this is 'a crucial part of oral competence', and thus merits inclusion in the test format. The small talk initiated in the opening stage of the test can also be viewed as an attempt to offset the constraints of the testing procedure and create a more relaxed atmosphere. Furthermore, the examiner discovers personal information about the candidates which should perhaps influence the selection of a particular examination pack. According to Selinker and Douglas's Discourse Domain Model (1985: 199), second language learners acquire their language through different domains of discourse, which usually involve specific

content areas. If this is true, a candidate's performance may vary from one discourse domain to another and selecting a topic of interest to the candidate may scaffold their L2 ability. The importance of content selection is reinforced by Zuengler's study (1993: 423) which found that 'inequalities in content knowledge can lead to different patterns of active participation by interlocutors.' Although Zuengler's (1993: 423) research did not involve oral proficiency interviews, implications for oral testing can be inferred:

'... when NNSs are engaged in talking about something that they know more about than their interlocutors, their greater content knowledge can override any limitations they may have in their oral proficiency, and enable them to be the talkers in the conversation. In so doing, these conversations may provide the best opportunity for learners to perform this aspect of what Young (1992) calls their "interactional competence".'

Conclusion

This article has examined the topical content of the Cambridge Speaking Tests and shown the importance of topic choices if individual candidates are to engage with the test materials and demonstrate their proficiency in English. Effective speaking test materials must crucially and minimally allow examiners to elicit a language sample suitable for assessment purposes. Speaking tests are likely to capture more accurate evidence of oral proficiency, and be more appealing for candidates if they allow opportunities for more personal exchange of information, and create 'genuine' communication gaps. Allowing candidates more responsibility to select topics which interest them could potentially offset the one-sided nature of the interview by reducing examiner control, and encouraging greater candidate participation. Whilst recognising the importance of test standardisation, it is incumbent upon test developers to explore new and innovative ways to create choice and flexibility in speaking test materials. This is in line with Tomlinson's (1996) and Maley's (1998) views on the importance of developing teaching materials which offer learners choices, and would seem a fruitful direction for future research into the development of speaking tests.

References

Bell, J. and Gower, R. (1998) 'Writing course materials for the world: a great compromise' in Tomlinson, B. (Ed.) *Materials Development in Language Teaching*, Cambridge: Cambridge University Press.

Erickson, F. and J. Shultz. (1982) *The Counselor as Gatekeeper. Social Interaction in Interviews*. New York and London: Academic Press Inc.

Gray, J. (2000) 'The ELT coursebook as cultural artefact: how teachers censor and adapt' *ELT Journal* 54/3: 274-282.

Johnson, M. (2001) *The Art of Non-conversation*, New Haven and London: Yale University Press.

Lazaraton, A. (1991) 'A Conversation Analysis of Structure and Interaction in the Language Interview.' *Ph.D. thesis*, University of California.

Maley, A. (1998) 'Squaring the circle reconciling materials' in Tomlinson, B. (Ed.) *Materials Development in Language Teaching*, Cambridge: Cambridge University Press.

Pasathas, G. (1995) *Conversation Analysis. The Study of Talk-in-Interaction. Qualitative Research Methods Series 35*. Sage Publications Inc.

Reda, G. (2003) 'English coursebooks: prototype texts and basic vocabulary norms', *ELT Journal* 57/3: 260-268.

Rinvolutri, M. (1999) 'The UK EFL sub-culture and dialect', *Folio* 5/2.

Selinker, L. and Douglas, D. (1985) 'Wrestling with "context" in interlanguage theory.' *Applied Linguistics* 6/2: 190-204.

Thornbury, S. (1999) 'Window-dressing vs cross-dressing in the EFL sub-culture', *Folio* 5/2.

Tomlinson, B. (1996). 'Choices', *Folio* 1/3.

Van Lier, L. (1989) 'Reeling, Writhing, Drawing, Stretching and Fainting in Coils: Oral Proficiency Interviews as Conversation.' *TESOL Quarterly* 23/3: 489-508.

Young, R., and Milanovic, M. (1992) 'Discourse Variation in Oral Proficiency Interviews.' *Studies in Second Language Acquisition* 14/4: 403-24.

Zuengler, J. (1993) 'Encouraging Learners' Conversational Participation: The Effect of Content Knowledge.' *Language Learning* 43/3: 403-32.

Acknowledgements

I am extremely indebted to UCLES for allowing me to use their data.

Julie Norton is a Lecturer in Education, at CELTEAL, the School of Education, at the University of Leicester. She teaches on MA courses and supervises doctoral students in Applied Linguistics and TESOL. Her doctoral research at the University of Cambridge investigated the performance of Japanese candidates in the Cambridge Speaking Tests. Her research interests include: intercultural pragmatics; discourse analysis and oral testing. She has taught English in Europe and Japan.

‘There is no doubt that comparative frameworks can serve a useful function for a wide variety of test stakeholders: for test users – such as admissions officers, employers, teachers, learners – frameworks make it easier to understand the range of assessment options available and help users to make appropriate choices for their needs; for applied linguists and language testers frameworks can help define a research agenda and identify research hypotheses for investigation; for test providers frameworks not only help with product definition and promotion, but also with planning for future test design and development. But we need to understand that they have their limitations too: they risk masking significant differentiating features, they tend to encourage oversimplification and misinterpretation, and there is always a danger that they are adopted as prescriptive rather than informative tools. They need to come with the appropriate health warnings!’

**Dr Lynda Taylor in Cambridge ESOL
Research Notes, February 2004, P.5**