Simulating conversations in oral-proficiency assessment: a conversation analysis of role plays and non-scripted interviews in language exams
Judit Kormos

The online version of this article can be found at:
http://ltj.sagepub.com/cgi/content/abstract/16/2/163

Additional services and information for *Language Testing* can be found at:

Email Alerts: http://ltj.sagepub.com/cgi/alerts

Subscriptions: http://ltj.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.co.uk/journalsPermissions.nav

Citations http://ltj.sagepub.com/cgi/content/refs/16/2/163

Simulating conversations in oral-proficiency assessment: a conversation analysis of role plays and non-scripted interviews in language exams

Judit Kormos *Eötvös University*

Several recent studies have investigated the nature of interaction in oral proficiency exams and have concluded that the interview format obscures differences in the conversational competence of the candidates. The present paper examines what opportunities test takers have to display their knowledge of managing conversations in the L2 in two types of tasks: non-scripted interviews and guided role-play activities. The data for the study consists of 30 interviews and 30 role-play activities between near-native examiners and intermediate learners used in language exams in Hungary. The interviews and role-plays have been analysed for the number of topics introduced and ratified by the examiner and the candidate respectively, as well as for the number of interruptions, openings and closings produced by the examiner and the candidate. The findings show that the conversational interaction is more symmetrical in the guided role-play activity with the candidates introducing and ratifying approximately the same number of topics as the examiners. In addition, the examinees have the opportunity to interrupt and hold the floor more effectively during the role-play activity and can demonstrate their knowledge of how to open and close a conversation. These findings suggest that guided role-play activities used in the study exhibit several characteristics of real-life conversations and therefore can be used for assessing the candidates' conversational competence.

I Introduction

Recent multidimensional models of communicative competence (e.g. Bachman, 1990; Celce-Murcia *et al*., 1995; Bachman and Palmer, 1996) have identified oral discourse competence as a distinct component of L2 speakers' communicative language ability. Conversation is one of the basic means of oral interaction; therefore, being able to participate actively and appropriately in a conversation is a skill that many language learners would like to and need to acquire. Among others, conversational competence can be characterized by the ability

Address for correspondence: Judit Kormos, Eötvös University, Department of English Applied Linguistics, Budapest, 1146 Ajtósi Dürer sor 19–21, Hungary; e-mail: kormos@isis.elte.hu

to perform openings, re-openings, closings and pre-closings, to establish and change topics, to hold and yield the floor, to backchannel, to interrupt and to collaborate, as well as to recognize and produce adjacency pairs (Celce-Murcia *et al.*, 1995). Thus the mastery of these components of oral discourse competence constitutes an important measure of L2 proficiency.

As many state-of-the-art language tests attempt to tap the knowledge of the candidates in different language competencies, the assessment of test takers' conversational competence has also gained in importance (e.g. Cambridge First Certificate, Cambridge Certificate of Proficiency and the Hungarian State Language Exam). One problem with the most commonly used forms of oral language tests – the oral proficiency – interviews is that they are unequal social encounters, thus they inherently resemble interviews rather than natural conversation. The traditional interview format of language proficiency exams might prove to be an adequate means for measuring linguistic (grammatical, lexical, etc.) competencies; nevertheless, several researchers (e.g. Bachman and Savignon, 1986; Bachman, 1988; van Lier, 1989; Lazaraton, 1992; Young and Milanovic, 1992; Young, 1995) argue that it does not create a situation in which conversations can occur.

In the past few years a number of studies have examined oral language proficiency exams from a discourse analytic perspective (e.g. Ross and Berwick, 1992; Ross, 1992; Lazaraton, 1992; Katona, 1998). The present study builds on Young and Milanovic's (1992) and Young's (1995) framework of analysis and investigates candidates' opportunities for displaying their conversational competence in two types of task. For this purpose, candidates need to be provided with opportunities to perform openings and closings, to establish, change and reject topics as well as to interrupt the examiner. Thus, the method of analysis adopted for examining these opportunities concentrates on patterns of dominance and contingency, as well as the distribution of rights and duties in two distinct phases of the Hungarian State Language Exam: the non-scripted interview and the guided role-play task. Non-scripted interviews were selected for analysis in order to examine whether interviews in which the questions and topics are not pre-determined provide more opportunities for test takers to initiate new topics and reject topics proposed by the examiners, thus gaining more control over the interaction than scripted interviews studied by Young and Milanovic (1992) and Young (1995). The hypothesis of the study was that the interview setting, regardless of its format, will determine the nature of communicative exchange between the testers and the test takers; consequently it will be unsuitable for measuring conversational

competence. The use of role-plays in language proficiency exams, however, can create a context in which examiners and candidates assume different social roles (e.g. friends, acquaintances, etc.). The study investigates how taking on a different persona in a test setting influences the structure of interaction. It was hypothesized that simulating more natural conversations is possible with the help of this task. Besides a comparative quantitative analysis of dominance and contingency of role-plays and interviews, the paper also considers the effect of the script of the role-play task on the interactional structure.

1 Differences between interviews and conversations

Based on the analysis of turn-taking in conversation by Sacks *et al*. (1974), conversation can be defined as a face-to-face interaction which has not been planned ahead, and the outcome and sequence of which is unpredictable. This definition, however, is not complete since it does not mention the equal distribution of rights and duties in carrying out conversations, which is one of the major factors that distinguishes conversations from interviews (Silverman, 1973; Kress and Fowler, 1979). Therefore Sacks *et al*.'s (1974) system for analysing conversations needs to be complemented with Jones and Gerard's (1967), Silverman's (1973) and Goffman's (1976) models of dyadic interaction. Jones and Gerard (1967) use the concepts *contingency* and *goal-orientation* to model differences between conversations and interviews. In their definition, contingency is the social determinant of the structure of interactions, and goal-orientation is affected by the internal goals of the participants. In Jones and Gerard's (1967) model, interviews are characterized by *asymmetrical contingency*, whereas conversations can be either *reactively* or *mutually contingent*. This means that in interviews both goal-orientation and reactiveness are asymmetrical; the interviewer has high goal-orientation and exhibits a low degree of reactivity, whereas the interviewee's behaviour is characterized by a lower degree of goal-orientation but by a high degree of reactivity. In conversations both of these characteristics are symmetrical, and the reactiveness of both parties is high. The sole difference between reactively and mutually contingent conversations is that in the former the goal orientation of the interactants is low (rambling conversations), while in the latter both interactants have strong internal goals (negotiations and discussions). Thus, the definition of conversation can be expanded to 'an unplanned face-to-face interaction with unpredictable sequence and outcome in which the rights and duties of the interactants are equally distributed and in which speakers' turns are reactively or mutually contingent' (based on van Lier, 1989: 495).

Interviews are similar to conversations in that they are face-to-face interactions with sometimes unpredictable outcomes, but rights and duties are far from equally distributed in them. Silverman (1973) and Kress and Fowler (1979) argue that the interviewer has the right to open and close the conversation, to introduce new topics, and to ask questions, whereas the interviewee is merely allowed to ask questions at rare occasions. As for contingency, interviews are characterized by asymmetrical contingency (Jones and Gerard, 1967). Thus, interviews can be defined as 'unequal social encounters in which the interviewer retains most of the rights in the interaction and in which turns are asymmetrically contingent upon each other'.

2 Language proficiency exams: interview or conversation?

Oral language proficiency exams are usually conducted following the format of interviews, yet test designers and testers usually regard the encounter between examiners and candidates as an instance of conversation (see for example Oller, 1979; Adams, 1980). In evaluations of the ACTFL (American Council on the Teaching of Foreign Languages) oral interview, Bachman and Savignon (1986) and Bachman (1988), however, argued that the interview process does not truly represent the wider context of transactional and interactional discourse. Van Lier (1989) also pointed out that oral language exams are asymmetrical interactions, in which rights and duties are unequally distributed. He concluded that interviews do not reflect the characteristics of conversations, as 'the emphasis throughout is on successful *elicitation of language*, not on successful *conversation*' (van Lier, 1989: 501, emphasis in original). Van Lier's paper was theoretical in nature but has provoked several empirical investigations in this vein.

One of the directions for further research into oral interview discourse outlined by van Lier (1989) was the analysis of the negotiation of meaning between the examiner and the candidate. Ross and Berwick (1992) concluded that a key aspect in understanding interactions in OPIs (Oral Proficiency Interviews) is the exploration of how accommodation takes place between native and non-native speakers. Ross (1992) investigated accommodative questions in 16 OPIs conducted in the framework of an in-house language-training programme at a Japanese company, and argued that 'the manner and quantity of interview accommodation necessary for the interview to take place should be considered in the assessment process' (Ross, 1992: 183). In a study involving 60 candidates in the same setting, Ross and Berwick (1992) analysed 14 types of conversational modifications: four exponents of control and 10 exponents of accommodation. Their findings indicate that, as regards accommodation, OPIs display the features of both conversations and interviews.

Another direction in examining the nature of interaction in OPIs has been the analysis of turn-taking and the distribution of rights and duties. Examining placement tests at UCLA, Lazaraton (1992) found that turn-taking is not locally managed in these encounters, but it works by means of a 'pre-specified system which defines an interaction . . . as an instance of "interview"' (Lazaraton, 1992: 383). Young and Milanovic (1992) and Young (1995) analysed dominance and contingency in oral proficiency interviews that constitute one part of the Cambridge First Certificate Exams. Young and Milanovic adapted Jones and Gerard's (1967) model of goal-orientation, reactiveness and dominance for their analyses and argued that language proficiency interviews are asymmetrically contingent. This seems to indicate that, from this aspect, they have little in common with conversations. Based on a similar method of analysis, Young (1995) also concluded that not only does the interview format not measure conversational competence but it may also 'obscure discourse differences between learners' (Young, 1995: 37). Therefore, one direction for further research is to study how different formats of language proficiency exams can contribute to more accurate evaluation of L2 learners' communicative competence.

The focus of the present study has been to investigate to what extent the assessment of conversational competence is possible in two types of task used in oral language exams. Using an extended version of Young and Milanovic's (1992) and Young's (1995) framework of analysis, the paper analyses the distribution of rights and duties in these tasks and considers the implications of the analysis for measuring conversational competence.

II Method

1 Design of the study

a Subjects   The data for the study was collected in Hungary in 1992 during English examinations offered by the Hungarian State Foreign Language Examination Board which issues official state language certificates. The 30 subjects who volunteered to participate in the study had attended a three-year-long intensive language course organized by the examination board in order to prepare the students for the intermediate level of the State Language Examination. They received a total of 660 hours of instruction during these three years. At the end of the course, the participants took the Intermediate State Language Examination, the level of which is slightly lower than that of the Cambridge First Certificate exam. The participants' ages ranged from 17 to 48; 10 of them were male and 20 female. Despite the fact that

they attended the same language course for three years, there were slight differences in their level of proficiency. Nineteen candidates received a final score for the oral exam between 30 and 40 (see Table 1), which means that they reached the required level to be awarded the certificate (which is roughly equivalent to achieving a 'C' grade in Cambridge First Certificate exams). Five of the participants obtained fewer than 30 points and did not come up to the required standards in the oral exam, and six candidates scored over 40 points indicating that they considerably exceeded the requirements at this level.

The oral exams were conducted by three female examiners and one male examiner, who worked for the testing centre as full-time teachers and examiners. Their native language was Hungarian, but they spoke English at a near-native level. They had all received training in language testing. Two of the examiners had several years of experience, whereas two of them had been working in this field for a shorter period.

A one-way analysis of variance (ANOVA) of the interviewing style of the examiners was carried out to determine whether there was any significant difference as regards the frequency of topics introduced by the examiners, the ratio of the topics introduced by the candidate and ratified by the examiners, as well as the frequency of interruptions performed by the testers in the interview and role-play tasks. No statistically significant differences ($p < 0.05$) were found in these variables. Thus it can be concluded that the four examiners were fairly homogenous in their interviewing style, and their idiosyncratic features did not affect the results of the study in any of the two tasks used.

b Procedure   The exams were conducted by two examiners each, one of whom acted as the interlocutor, and the other one assessed the candidate's performance. Testees had to perform three tasks at the exam: a general non-scripted interview, a guided role-play, and a picture-description activity. The examiners did not have to adhere to strict guidelines in conducting the non-scripted interview, but this phase was supposed to consist of a warm-up and a probing phase, in which the examiner and the candidate discussed 2–3 everyday topics

Table 1   Final scores achieved by candidates in the oral exam

Score achieved at the oral exam	25	27	29	30	31	33	34	35	36	39	40	41	44	46
Frequency		1	1	3	3	1	5	1	2	3	2	2	1	4	1

(e.g. the candidate's job, hobbies, family, etc.). This was followed by a picture description activity and a role-play task, in which the testee and the examiner acted out a situation described on a role-play card. Cues for the role-play were given in the native language of the students (see Appendix 1 for examples of role-plays translated from Hungarian into English). Most of the role-plays included conversations between friends, acquaintances and colleagues (12 out of 15 role-plays), and three tasks were service encounters. The discourse domains of the role-play activities will be discussed in the analysis of the effect of the script on the structure of the interaction.

The oral exams were tape-recorded, video-recorded and later transcribed. The transcriptions were made by near-native speakers of English and have been checked by the researcher. The data were analysed and coded by the researcher based on the transcript and multiple listening to the recording.

2 Analysis

On the basis of the definition of conversations and interviews outlined above, it is contingency, dominance and the power relations of the participants that differentiates these two types of interaction. Therefore the present study has attempted to operationalize these concepts. Since *topic* is the key component in the analysis of contingency and dominance, a framework of analysis for identifying topics is essential. Identification of topics constitutes a controversial area of discourse analysis. Several researchers among discourse analysts (e.g. Maynard, 1980; Brown and Yule, 1983) have pointed out that even participants would judge differently at different points of the conversation what is being talked about, that is, what the topic of the conversation is. Nevertheless, there seems to be a consensus that knowing where a notion commences and ends suffices for identifying a notion. Topics therefore can be defined by finding their boundaries, which are called *topic shifts* (Maynard, 1980).

Consequently, in order to distinguish different topics in conversation, topic shifts need to be identified. The present study combined Brown and Yule's (1983) and Young's (1995) systems for establishing instances of topic shift and used the following criteria:

1) explicit boundary markers such as 'all right', 'so'
2) imperatives or questions in the speech of the interviewer
3) long unfilled pauses (usually exceeding one second)
4) introduction of new information
5) rounding off by repetition or paraphrase in closing of a salient lexical item that was used to initiate the topic

6) high pitch on a new lexical item, clause or sentence as an indication of topic opening
7) low pitch on the same lexical item that opened a topic (or a paraphrase of it) as an indication of closing, loss of amplitude
8) explicit abandonment of the topic.

A reliable system of establishing the boundaries of topics, i.e. topic shifts, was necessary as measuring dominance and contingency is based on the notion of topics. Using Young and Milanovic's (1992) framework of analysis for contingency, which narrows down Jones and Gerard's (1967) definition, contingency was defined as 'property of adjacent turns in dialogue in which the topic of the preceding turn is coreferential with the topic of the following turn' (Young and Milanovic, 1992: 405). The measures of contingency were the number of initiated topics, the number of ratified topics and the proportion of topics initiated by one party that are the topics of the subsequent turns by the other party (Young and Milanovic, 1992: 405). A topic was considered to be ratified if the topic initiation of Participant A was followed by at least one T-unit turn by Participant B in which the same topic continued (based on Young, 1995). (T-unit is 'one main clause plus whatever subordinate clause happens to be attached or embedded within it' (Hunt, 1966: 735).)

In addition to contingency, unequal power relations also distinguish interviews from conversations. Interactants in a dialogue can have several rights. The most basic rights are to open and terminate a conversation as well as to initiate new topics (Sacks *et al.*, 1974). In an interview it is the interviewer's right to open and close the dialogue and to ask questions by which he or she introduces new topics, whereas in a conversation these rights are shared by both participants. Participants in conversations are also entitled to reject or ignore a new topic, that is, not to ratify it. Although greatly dependent on culture, the right for interruption also exists in dyadic interactions; thus, in equal social encounters each party has the right to interrupt the other (Sacks *et al.*, 1974).

Duties in dialogues involve reacting to the other party's initiation of openings and closings. Depending on the context and the given culture, answering questions in some dyadic interactions might also be perceived as a duty by one of the interactants or by both of them. In a conversation between status equals these rights and duties are shared by both parties. In an interview situation, however, one party, the interviewee, is deprived of his or her rights but is heavily burdened by the duties. Therefore, imbalanced power relations are inherent characteristics of interviews. (One of the reviewers pointed

out that the schemata for participating in interviews might be culturally different. Young (1995) supported this claim by arguing that Asian test takers in his study tended to expect the interviewers to dominate the interactions, while Europeans assumed that the interviewee also has to actively participate in managing the exchange).

The present study identifies which speakers retain which rights in the interview and in the role-play activity. The analysis of contingency shows which party has the right to initiate and ratify topics. It also indicates which interactant initiates the openings and closings (pre-closings did not occur in the data). The number of interruptions demonstrates which speaker is entitled to perform interruptions, that is, who has the right to take the floor when the interlocutor has not provided signs that he or she is willing to yield the floor yet (Sacks *et al.*, 1974). This analysis is very similar to the examination of *dominance* in Young's (1995) study, which he defines as 'the tendency for one participant to control the discourse in various ways such as gaining, holding or ceding the floor by means of interruptions and questions' (Young, 1995: 16). Despite his elaborate definition, Young measured dominance only by the quantity of each participant's talk and the number of topic initiations. In order to gain a more precise view of dominance in the interactions to be investigated, the present study complemented Young's measures of dominance and analysed the openings and closings, as well as interruptions in the two types of tasks.

In each transcript the overlaps (// or ///) and interruptions were identified (EI = examiner interruption; CI = candidate interruption), and each instance of topic initiation and ratification was coded (ET = examiner's topic; CT = candidate's topic; RET = ratified examiner's topic; RCT = ratified candidate's topic). The initiation of openings and closings was also indicated (EO = examiner opening; CO = candidate opening; EC = examiner closing; CC = candidate closing). The following dialogue fragment from the non-scripted interview illustrates the analysis of the data (– short pause; + long pause; ++ extended pause; xxxx unintelligible word).

ET	*I*: ///Act? Well, xxx //Do – do they enjoy acting?
RET	*C*: Er – I don't think so be//cause (laughs) – it's a bit
EI	*I*: //That's interesting – Why? Are they too shy?
	difficult to them ///+ Some – er children, for example about
	I: ///I see.
	five children likes it very much//but the others – don't
	I: //Uhum
	like it – xxx xxx.
ET	*I*: Is there going to be a – a summer camp – organized by your school?

    RET *C*: Er – The summer camp is organized by the er – the – council //+
         now the – ///er council of the dis little er
         (*I*: //Uhum   ///xxx xxx xxx)
    CT   council – but it is er not so good for them. – //It's boring.
    RCT *I*: //Why not – and where, – where is it?

III Results and discussion

1 Interviews

Using the measures of dominance and contingency outlined above, the study yielded the following results with respect to the interviews. As Table 2 shows, the examiners initiated 312 topics, which is approximately three times more than the number of topics proposed by the candidates ($n = 85$). The candidates ratified 308 topics out of the 312 initiated by the examiners, which means that they accepted 99% of the examiners' topics, whereas examiners only ratified 44 topics out of 85 initiated by the candidates, which is merely 52% of the total number of topics introduced by the examinees. The examiners interrupted the candidates and took the floor 108 times, while the candidates interrupted the interviewers only 18 times. The examinees talked about 50% more than the examiners. The examiners initiated all the openings and closings in the interview. Table 2 summarizes the results of the quantitative analysis of the non-scripted interviews.

    The fact that the examiners ratified only 52% of the topics shows a low degree of reactivity on the examiners' part, whereas the 99% ratification rate of the examiners' topics suggests that candidates' reactivity is high. To show how this profile is played out in the actual interactions, we will examine a number of excerpts. The following example – non-scripted interview with Candidate 4 – shows two examiner-initiated topic shifts and a candidate-initiated one (conventions and abbreviations as before).

1   **ET**   *I*: ///but can you give an example – what kind of machines //there are?
         CT   *C*: //For example are a – in a foundry ++ in Miskolc – or Diós-győr ++
5   Once have to – change – the old machine into a new one – and ++ three – company bought – that – new machine, – and we installed + we set + and er – set in + work. //+ So that . . .
     EI, ET *I*: //Yes, I I – understand it, and do you find your job interesting?
       RET   *C*: Yes, because it's it's not boring, – everything's changing er
10   ++ er and er + my colleagues are very – pleasant and friendly.
       ET   *I*: And is it a – steady job for you? //+ I mean that er – er can you be sure that you – can keep this job and you will – have this job?
       RET   *C*: (sighs) I – you are right because it's it's very – difficult situ-
15   ation er because of the Hungarian er – ek economy // – is not – the best situation.

Table 2 Measures of contingency and dominance in the non-scripted interview

	Topics		Ratified topics		Per cent ratified		Interruptions		Openings		Closings		Number of words	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C
Mean	10.40	2.83	10.27	1.47	51.57	98.95	3.60	0.60	1	0	1	0	242.53	348.27
Standard deviation	14.11	1.78	3.65	1.11	25.22	2.85	3.17	1.22	0	0	0	0	87.84	108.88
Sum	312.00	85.00	308.00	44.00	–	–	108.00	18.00	30	0	30	0	7276.00	10448.00
Frequency	1.82	0.50	1.80	0.26	–	–	0.58	0.09	–	–	–	–	–	–

Notes: Frequency = (sum of variable/total number of words produced in the interaction $\times$ 100); E = examiner; C = candidate

```
                    + So +
                    I: //uhum
                    because I – can – speak and write English and er German I hope
20                  er they need – a man – like me ///+ so they – don't
                    I: ///yes
                    want – to dismiss me.
```

In line 1 the examiner makes an attempt to bring in a new topic about the types of machine the candidate is working with. Instead of ratifying the examiner's topic, the candidate initiates a new topic by telling the interviewer about the installation of a new machine. The examiner, however, neither accepts the examinee's topic nor tries to repeat the previous topic initiation but shifts to a completely new topic, the one of job satisfaction in line 5. The initiation of the new topic on the examiner's part is carried out by means of interrupting the candidate in the middle of the explanation of why he wanted to put the new machines into operation. The way the examiner manages turn-taking in this excerpt illustrates the extensive right of the testers to take the floor by means of interruptions. The candidate ratifies the examiner's topic similarly to the following one in lines 15 and 16 by answering the question about the security of his job. This example illustrates the general finding of the study that, in the interview-phase, candidate-initiated topics are mainly rejected, whereas examiner-initiated ones are more frequently accepted. This result may not be surprising, since interviewers might often interpret topic shifts as avoidance or the inability to answer and move to another 'test topic'. Personal communication with several examiners in the Hungarian Foreign Language Testing Board (HFLTB) supports this assumption. On the other hand, a recent article on the methodology of posing questions at the exams provided by the HFLTB (Katona, 1997) mentions that quick topic shifts are necessary to get a view of the candidate's competence in a wider variety of topic areas and to prevent testees talking about subjects they have prepared in advance.

The following excerpt – non-scripted interview with Candidate 1 – exemplifies one of the few successful candidate topic initiations in the interview.

1	**ET**	*I*: Ham – or sausage or what is it?
	RET	*C*: Sausage
		I: Sausages – right.
		C: Two little – er cubes.
5		*I*: I see.
	CT	*C*: And er – I'm a scout +
	RCT	*I*: Are you?
		C: Yes, I am. – And //one of . . .
	EI, ET	*I*: Do you enjoy being a scout?
10	**RET,**	*C*: Yes, I'm going too – go a camp this afternoon.
	CT	

RCT		*I*: This after//noon? – Dear me! – All ri///ght.

```
      RCT      I: This after//noon? – Dear me! – All ri///ght.
               C:            //Yes + eh                    ///Well, it will be a
               little camp, just – three days – //and that's because it
15             (I:                                 //Three days – I see)
               + great(?) /// – and – er so that's it.
      ET       (I: ///Uhum – and how – how long – have you been a scout?
      RET      C: For three years.
               I: For three years now // – that – that means you were +
20    CI       (C: //Yes – it's a . . .)
               among the first scouts /// – in Hungary.
               C: Yes + er// . . .
```

The beginning of this excerpt constitutes the end of the previous topic, which was initiated by the examiner and was about preparing scrambled eggs. When the candidate perceives that this topic can be closed, he abruptly introduces a new topic about being a scout. The examiner accepts the topic, together with the following one, which is about the candidate's going to a camp. It is interesting to observe that the candidate himself closes this topic in line 16 by saying 'so that's it'. Although the manner of both of the topic initiations is non-native-like, the acts themselves are successful. Another phenomenon in line 20 is also typical of the management of turn-taking in the interview. Here the candidate starts talking simultaneously with the examiner, but he gives up his turn and lets the examiner finish her indirect question.

As far as dominance and power relations in the non-scripted interview are concerned, the quantitative data suggest that it is solely the examiners' right to initiate the opening and closing of the dialogue. As regards the initiation of topics, candidates seem to be fully aware of their unequal status and attempt to introduce a new topic quite infrequently; furthermore, they are unsuccessful in getting the topic accepted in 51% of the cases. This shows that although candidates have the right to introduce new topics, this right is of much smaller scope than that of the examiners, which is also illustrated by the fact that candidates on average initiate topics approximately four times less frequently than examiners do (see Table 2). Candidates' right to reject a topic seems to be even smaller; rejection is attempted in only four cases (1%) in the data. These figures suggest that it is the duty of the interviewee to accept the topics introduced by the examiner, while the examiner is not in the least obliged to accept the topic of the candidate. The analysis of interruptions reveals that examiners have considerably more authority in taking the floor at places in which transition is not relevant, although candidates retain some rights for interruption as well. As a result, non-scripted interviews are unequivocally dominated by the examiners.

In sum, the analysis of non-scripted interviews supports van Lier's (1989) and Young's (1995) claims that oral language proficiency interviews are not conversations because these dyadic interactions are asymmetrically contingent as the parties display different degrees of reactivity, and the power relations of the interactants are highly unequal. On the other hand, as pointed out above, certain control of topic initiation is essential on the examiners' part to be able to make an accurate judgement of the candidate's linguistic competence. It can be concluded therefore that despite the fact that non-scripted oral proficiency interviews may measure certain components of oral language proficiency reliably, they do not provide candidates with adequate opportunities to display their conversational competence.

2 *Guided role-plays*

The results of the analysis of the guided role-plays (see Table 3) show that candidates initiate approximately 50% more topics ($n = 109$) than examiners do ($n = 73$). The following excerpt – guided role-play task with Candidate 15 – illustrates two subsequent successful topic initiations on the examinee's part.

	I: (sighs) Well, hi.
CO	*C*: Er – hi er – hi, ha Harry – eer + How are you?
	I: Eer – fine, thanks.
CT	*C*: I – mm er – I heard er + eer – youu – er a er I I eer I he – I heard youu –
	you had er – won – er on – on the lottery ++
RCT	*I*: + Erm er – yes – I did. + Just yesterday.
	C: Er I + I – I can – I ca I er – can co–congratulation – //eer mmm . . .
EI	*I*: //Yes, you know I was bankrupt – and er + (sighs) now – I won this – sum of money, and this is a huge sum of money + 9 – million – forints.
	C: Eer mm – eer – yes e-er – youu ha – youu + had have aa – gra – a great er – fortune //– eer what do you
	I: //Uhum – well – yes I did
CT	want what do youu – eer – want – er – to do with this er – ///sum?
RCT	*I*: ///Eerm ++ er I'm not sure – I'm thinking of – buying aa + huge mansion. ++ A great big building. //– A huge house
	(*C*: //great)
	for myself.

Upon acting out meeting a friend, the candidate starts the conversation by introducing the topic of the friend's winning the lottery. The examiner, who plays the role of the friend, accepts the topic and elaborates it by telling the examinee that she came into the money at the right time because she was on the verge of going bankrupt. This theme is developed by the candidate's initiating a further topic about what the friend will do with the money. The examiner ratifies this new topic as well.

Table 3 Measures of contingency and dominance in the guided role-play activity

	Topics		Ratified topics		Per cent ratified		Interruptions		Openings		Closings		Number of words	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C
Mean	2.43	3.63	2.37	3.43	97.62	96.52	1.23	0.47	–	–	–	–	153.43	183.53
Standard deviation	1.14	1.59	1.13	1.38	8.74	9.78	1.19	1.11	–	–	–	–	56.62	52.01
Sum	73.00	109.00	71.00	106.00	–	–	37.00	14.00	13	17	18	12	4603.00	5506.00
Frequency	0.73	1.18	0.41	0.63	–	–	0.36	0.12	–	–	–	–	–	–

Notes: Frequency = (sum of variable/total number of words produced in the interaction × 100); E = examiner; C = candidate

There is no difference in the rate of the ratification of topics: both the examiners and the candidates ratified 97% of their interlocutor's topics. These figures reflect the balanced nature of the interactions in the guided role-play activity. The excerpt below – guided role-play task with Candidate 4 – exemplifies one successful candidate's topic initiation followed by another one on the examiner's part (see Appendix 1 for the script of the role-play task).

	I: Er ++ one four two – eight two nine oh . . .
CO	*C*: Ah – hallo – then it's a right number, I'm Gabriella //– er
	I: //Ah, hello, Gabriella, – this is Ilona speaking . . .
CT	*C*: Hi, Ilona – I call you er – because I would like to invite you – er for a good film. – It goes er – on programme at the – Zrínyi – cinema + and er – I know er – do you like this this kind of films and that's why // – I . . .
EI, RCT	*I*: //What film is it?
	C: + Aah, – it's it's aa – special er + er (laughs) er horror film + and er + er + I'm interested in er because er + I heard er – about it er – very – er good er + er + criticis ++ but er you know – I am er – not love of the – horror films but er I thought – er we can see it together.
ET	*I*: Uhum, – yes, I've heard of it, tooo and I would really er – love to – see it but er – at the moment I'm awfully busy. //++ I think I had – better er – keep on + doing my
CI	(*C*: //Oh, er)
	work.
RET	*C*: Oh, ++ it's terrible – but is it so er – er important?
	I: Yes, I I have to finish something by – the day after tomorrow.
CT	*C*: But er – I think it's er – good er to relax er a little bit + only – for two – or three hours, and then er – maybe – you can er + make it further – but er
	(*I*: //uhum)
	// – er with fresh – energy?

First the candidate introduces the topic of going to cinema, which the examiner accepts by requesting more information about the film the candidate wants to see. The candidate, in turn, provides some details about the film. The examiner, however, rejects the invitation by saying that she has some work to do. This new topic is now ratified by the candidate, who asks about the importance of the work and introduces the topic of having a rest and continuing the work later. Thus, in this short interaction both parties initiate new topics and ratify the topics of their interlocutors.

As regards interruptions, examiners still interrupt more frequently than candidates (see Table 3), but compared to the interruption rate in the interviews (which is approximately three times greater), in the role-play task the examiner interrupts the candidate only 70% more frequently than the candidate does. The more frequent candidate interruptions manifest the candidates' increased rights in this task as opposed to the interviews.

Candidates initiate more openings ($n = 17$) than examiners ($n = 13$), but examiners initiate the closing of the conversation ($n = 17$) more frequently than candidates ($n = 13$). Both of the above role-play task excerpts can serve as examples of candidates' openings. Despite the fact that the script of the tasks often affects who will perform the openings and the closings of the interaction, it seems likely that even in the role-play tasks candidates will wait for the examiner to close the conversation. The reason for this might be that, by closing the conversation, the candidates deprive themselves of further opportunities to display their competence; thus they would be reluctant to terminate the interaction. The following excerpt – guided role-play task with Candidate 19 – illustrates one of the few candidate's closings which was performed as required by the script of the task (see Appendix 1 for the script).

	I: + Yes eer, – they are promising – some improvement in the weather for //the weekend.///
CT	*C*: //Oh xx ///Because my er – children wan wan want tos – to – go an excursion to the mountains.
RCT	*I*: I think er – it would be – just appropriate for – for an outing.
CC	*C*: Oh my God, there's my bus – well, I'm in – I have to hurry – er – bye (*laughs*).
	I: Bye. ++ Thank you, – that was all.

This role-play task is a typical example of a short conversation at a bus-stop where two acquaintances discuss the weather. The candidate in this case is prompted to close the conversation by the arrival of the bus.

Upon examining the individual role-plays it becomes apparent that due to the scripts, roles can often be reversed as in the case of Candidate 10, who got the role of a fairly determined friend who has to persuade the examiner to make an appointment with a dentist. As can be seen in the interaction, the candidate clearly dominates the conversation by introducing new topics and repeatedly rejecting the examiner's topic, which is about her fear of dentists (guided role-play with Candidate 10).

CT,	
CO	*C*: Hi – er – hi, Dennis. – Oh, what's the trouble with you?
RCT	*I*: Well – you know, I have had a terrible night –
	C: Okay, and why? – What's the matter with you – er cheek? – Is there any trouble?
	I: Ah ah – do I look too cheeky? (*laughs*) + Yes, there is in fact you know, well – well, my tooth – my wisdom tooth +
CT	*C*: Okay, and er – er – have you been to a docto . . . the dentist? – //Tomorrow morning?
RCT	*I*: //Oh not yet, you know I'm terribly scared of dentists so I /// . . .
ET	
CI	*C*: ///Okay, but you have to go – //(*laughs*) – this case

        I: //I-ii – yes, – yes I think I should.

CT   *C*: Er I think I know – a very good dentist – //mine

RCT *I*: //Well, do you? – Well, 'cause you know I don't know anybody he///re.

CT   *C*: ///May I give you – er his er – address or telephone number.

RCT *I*: That's very helpful of you – really.

RCT *C*:: Okay, – I think you don't know any dentists here because you are //+ English . . .

ET   (*I*: //No, I don't – I don't and and anyway, – I hate – this situation, and I hate them //xxx

CT   *C*: ///Oh, I'm very sorry you (*laughs*) – Is it a sharp pain in your – tooth?

The fact that both parties initiate and ratify approximately the same number of topics in the role-play activity suggests that both the examiners and the candidates exhibit a high degree of reactivity and that subsequent turns are contingent upon each other. Thus, the tasks used in this study display the characteristics of conversations in that they are reactively contingent, and powers and duties are equally distributed among the participants.

Table 4 summarizes the results of a paired samples two-tailed t-test carried out to determine whether the differences observed in the two types of interaction: the non-scripted interview and the role-play task are significant in the variables investigated in the study.

As Table 4 shows, except for the frequency of candidate and examiner interruptions, all the variables are significantly different in the two tasks ($p < 0.01$). Examiners' topic initiation and ratified examiners' topics are significantly less frequent in the role-play task than in the interview, whereas candidates have significantly more opportunities to initiate new topics and have their topics ratified in the role-play activity. Examiners tend to perform fewer interruptions in the

Table 4 The comparison of the interviews and role-play tasks

Variables compared (frequencies)	*n*	mean	s.d.	t-value	df	*p*
Examiner's topic – interview with	30	1.82	0.58	8.83	29	*0.00
Examiner's topic – role-play	30	0.73	0.40			
Candidate's topic – interview with	30	0.50	0.29	–6.73	29	*0.00
Candidate's topic – role-play	30	1.18	0.63			
Ratified examiners' topic – interview with	30	1.0	0.58	12.65	29	*0.00
Ratified examiner's topic – role-play	30	0.41	0.21			
Ratified candidate's topic – interview with	30	0.26	0.20	–6.11	29	*0.00
Ratified candidate's topic – role-play	30	0.63	0.34			
Examiner interruption – interview with	30	0.58	0.45	2.31	29	0.03
Examiner interruption – role-play	30	0.36	0.35			
Candidate interruption – interview with	30	0.09	0.14	–6.00	29	.555
Candidate interruption – role-play	30	0.12	0.24			

Note: * = significant at the $p < 0.01$ level

role-play task ($p = 0.03$), while candidates interrupt numerically more frequently in this task, although the difference was not statistically significant.

These results seem to support the main initial hypothesis of the study, namely that guided role-play activities in language proficiency exams closely resemble conversations, and they, therefore, produce better measures of conversational competence than the interview phase of the exam.

The above findings clearly indicate the difference in the discourse structure of the two tasks; therefore one might rightly ask whether the performance of the candidates varies in the role-play activity and the interview phase. The paired samples t-test of the total scores of testees in these two tasks shows that the participants achieved significantly higher scores in the interview ($X = 20.9$) than in the role-play ($X = 20.1$) at the $p < 0.01$ level ($p = 0.009$). As these scores mainly reflect the fluency and accuracy of the candidate's output as well as their range of vocabulary, no inferences can be made as regards the assessment of conversational competence. Lacking the measures of the difficulty level of the two tasks, we cannot assume with certainty that the poorer performance of the candidates in these linguistic aspects in the role-play activity was due to the different discourse structure of this phase of the exam.

3 The effect of the script of the role-play task on conversational structure

Having examined the global quantitative results of the comparison of the role-play tasks and the interviews, it is important to explore how the actual script of the role-play affects the number of topic initiations and rejections. Appendix 2 summarizes the results of this analysis.

As regards topic rejections, it has been mentioned above that topics are rarely rejected in the role-play task, which is probably due to the fact that none of the scripts explicitly instruct the candidate not to ratify the examiner's topic. In both cases when rejections were performed, the candidates ignored the complaints of the examiners. In one of the role-plays mentioned above, the candidate did not ratify the examiner's mentioning of her fear of dentists and, in the other task, the candidate simply did not react to the examiner's observation about her standard of living.

The number of topics introduced by the candidates also seems to be strongly influenced by the script of the role-play task. In 20 out of 30 cases the candidates mention exactly the same number of topics as minimally required by the script of the role-play (see Appendix 2). If one considers that only five candidates introduce fewer topics

than prescribed, it can be assumed that scripted role-play tasks are successful in eliciting topic initiations from testees in most of the situations; thus, with their help, this component of conversational competence can be measured in a reliable manner.

The analysis of the interaction in the cases when the candidates introduce fewer topics than required reveals several reasons for this anomaly. Candidates 4, 9, 29 and 30 seem to have failed to mention certain topics due to limited proficiency (none of them passed the oral exam). In the case of Candidates 29 and 30, the discourse domain of the task might not have exerted considerable influence in this aspect, as both role-plays involved familiar situations (warning about crimes in the city and booking accommodation), which are frequently discussed inside and outside foreign language classrooms in Hungary. Candidates 4 and 9, however, had to perform a role-play task with a highly unfamiliar discourse domain, as they were instructed to act out a service encounter in which they had to sell a washing machine. In this task the candidate assuming the role of a shop-assistant had to explain the functioning of the washing machine. Thus, in this case, both the limited proficiency of the candidates and the unfamiliar discourse domain may have contributed to the fact that the candidates were either incapable of, or consciously avoided, introducing certain required topics. In the case of Candidate 8, whose competence was judged to be higher than the required level, we can only speculate why she introduced fewer topics. One reason might be that the role-play task seems to have been meaningful and complete without this particular topic, and she might have forgotten to mention it.

The scripts of many of the role-plays investigated in the study were flexible as regards who introduces certain topics first. In many cases the examiners initiated these topics often in the form of direct or indirect questions. It is interesting to observe that only in six cases did candidates perform these 'optional' topic initiations. Thus, it can be concluded that mainly only those topics are introduced in the interaction by candidates which are explicitly required by the script. Upon examining the individual cases, we find that Role-play 61 has a very flexible script concerning giving advice on what to do to avoid having one's house being burgled while on holiday. In this task either the examiner can bring in the different topics by asking questions, or the candidates can provide the different suggestions without any trigger. When performing this task, Candidate 18 introduced all the required and also the optional topics, and initiated a total of six topics instead of the prescribed three. Unfortunately, only one candidate among the 30 received this task, therefore no comparative data are available. Nevertheless, the influence of the level of proficiency can be excluded, as Candidate 18 was awarded 30 points, which is the

minimum score required to pass the oral exam. The other five candidates who initiated more topics than prescribed all passed the oral exam, but the number of their points varied between 33 and 44 (see Appendix 2), thus no direct evidence of the influence of the level of proficiency can be detected in these cases. Role-play 59, in which the examinees had to persuade their friends to go to the cinema with them, elicited more topics from both of the candidates who performed it. The reason for this might be that, owing to the persuasion element in the script, testees introduced more arguments than minimally required in order to achieve their goals. Candidate 16, who received a task in which he had to give advice to a friend about buying a new car, also initiated more topics than specified in the script. It was evident from the transcript that this examinee showed great interest in cars, which might account for the increased number of topic initiations. Lacking measures of personality for the other two candidates, one can only speculate that their increased willingness to communicate (McCroskey and Richmond, 1987; MacIntyre and Gardner, 1991) can explain why they undertook the optional introduction of the topics.

IV Conclusion

The present study examined how the type of task candidates have to perform in oral exams affects conversational structure and the candidates' opportunities to display their conversational competence. On the basis of the analysis of 30 non-scripted interviews and 30 guided role-play tasks, it was concluded that role-plays can be better measures of the candidates' abilities of managing conversations, that is performing openings and closings, initiating and rejecting topics, as well as interrupting. In non-scripted interviews, due to the unequal distribution of power, candidates rarely have the chance to initiate a new topic and have no right to initiate the opening or the closing of the conversation. Consequently, these components of conversational competence cannot be tested in non-scripted interviews. In contrast, the findings suggest that the guided role-plays in the present study display several characteristics of conversations: candidates have the right and opportunity to introduce, ratify and reject new topics in the course of the interaction. Depending on the script of the role-play task, they can also initiate the closing and opening of the conversation. These results indicate that guided role-plays, scripted with a focus on conversational skills, have the potential to compel candidates to exhibit their conversational competence.

The results of the analysis of how the scripts of the role-play tasks affect conversational structure suggest that rejections are rarely attempted, unless the script provides some opportunities for them. It

therefore seems that either the script of the role-play should contain instructions concerning topic rejections, or the structure of the interaction should be specified in a way that it would get the candidate not to accept the examiner's topics. In the above case, it may be important that the examiners introduce irrelevant topics as part of a conscious technique to elicit this component of conversational competence. As regards topic initiations, it was found that the role-play tasks are successful in eliciting the introduction of the required topics. On the other hand, when there are no explicit instructions concerning who should mention certain subjects, candidates rarely take the initiative to introduce these topics. Thus, if this component of candidates' communicative competence is tested, test designers should word the instructions of the task in a way so that test takers are explicitly instructed to introduce certain topics.

Several directions for further research can be outlined. In order to determine the reliability of the different tasks in measuring conversational competence, there is a need for more controlled studies. On the one hand, it could be investigated how the same subjects perform in one particular role-play activity. By means of combining the research methods of discourse analysis and educational psychology, new insights could be gained into how candidates' global proficiency, as well as certain individual variables (e.g. willingness to communicate) affect the number of topic initiations, rejections, interruptions and conversational closings and openings. On the other hand, it would also be worthwhile to study how subjects of approximately the same level of proficiency perform in role-play tasks with different scripts and with different specifications of the script as regards topic initiations and rejections, as well as closings and openings. Results of research in this field could give guidance to test designers concerning what kind of role-play activities and instructions can be the most efficient in eliciting the above mentioned components of conversational competence.

Further research would be necessary to determine whether the circumstances for engaging in conversations can be established in oral proficiency exams. The analysis of the interaction of two or more candidates (e.g. in the Cambridge Certificate of Proficiency, the Cambridge First Certificate, etc.) could yield new insights into the discourse structure of this task and could provide us with grounds for comparison with role-plays and interviews. It would also be worth studying how tasks requiring different degrees of interactional control from the candidates (e.g. role-plays, interviews and conversations) affect the accuracy, fluency, complexity and appropriacy of their performance.

on earlier versions of this paper. I would also like to thank Lucia Katona for allowing access to the tapes and transcripts of the exams. The writing of this paper has been supported by a grant from the Hungarian Scientific Research Foundation (OTKA No. F025212).

V References

Adams, M.L. 1980: Five co-occurring factors in speaking proficiency. In Frith, J.R., editor, *Measuring spoken language proficiency*. Washington, DC: Georgetown University Press.

Bachman, L. 1988: Problems in examining the validity of the ACTFL oral proficiency interview. *Studies in Second Language Acquisition* 10, 149–64.

——— 1990: *Fundamental considerations in language testing*. Oxford: Oxford University Press.

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

Bachman, L. and **Savignon, S.** 1986: The evaluation of communicative language proficiency: a critique of the ACTFL oral interview. *Modern Language Journal* 70, 380–90.

Brown, G. and **Yule, G.** 1983: *Discourse analysis*. Cambridge: Cambridge University Press.

Celce-Murcia. M., Dörnyei, Z. and **Thurrell, S.** 1995: Communicative competence: a pedagogically motivated model with content specifications. *Issues in Applied Linguistics* 6, 5–35.

Goffman, E. 1976: Replies and responses. *Language in Society* 5, 254–313.

Hunt, K.W. 1966: Recent measures in syntactic development. *Elementary English* 43, 732–39.

Jones, E.E. and **Gerard, H.B.** 1967: *Foundations of social psychology*. New York: Wiley.

Katona, L. 1997: Hogyan kérdezzünk a szóbeli nyelvvizsgán? *Testing Matters* 2, 20–23.

Katona, L. 1998: Meaning negotiation in the Hungarian oral proficiency examination of English as a foreign language. In Young, R. and Weiyun, H.A., editors. *Talking and testing: discourse approaches to the assessment of oral proficiency*. Amsterdam: John Benjamins, 239–67.

Kress, G. and **Fowler, R.** 1979: Interviews. In Fowler, R., Hodge, B., Kress, G. and Trew, T., editors, *Language and control*. London: Routledge and Kegan Paul.

Lazaraton, A. 1992: The structural organization of a language interview: a conversation analytic perspective. *System* 20, 373–86.

McCroskey, J.C. and **Richmond, V.P.** (1991). Willingness to communicate: a cognitive view. In Booth-Butterfield, M., editor, *Communication, cognition and anxiety*. Newbury Park, CA: Sage.

MacIntyre, P.D. and **Gardner, R.C.** 1991: Methods and results in the study of anxiety in language learning: a review of the literature. *Language Learning* 41, 85–117.

Maynard, D.W. 1980: Placement of topic changes in conversation. *Semiotica* 30, 263–90.

Oller, J.W. Jr. 1979: *Language tests at school: a pragmatic approach*. London: Longman.

Ross, S. 1992: Accommodative questions in oral proficiency interviews. *Language Testing* 9, 173–85.

Ross, S. and **Berwick, R.** 1992: The discourse of accommodation in oral proficiency interviews. *Studies in Second Language Acquisition* 14, 159–76.

Sacks, H., Schegloff, E.A. and **Jefferson, G.** 1974: A simplest systematics for the organization of turn-taking in conversation. *Language* 50, 696–735.

Silverman, D. 1973: Interview talk: bringing off a research instrument. *Sociology* 7, 31–48.

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

Young, R. 1995: Conversational styles in language proficiency interviews. *Language Learning* 45, 3–42.

Young, R. and **Milanovic, M.** 1992: Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition* 14, 403–24.

Appendix 1   Examples for guided role-plays

Role-play No. 59.
You call your friend because you want to ask him/her out to go with you to the cinema. There is a film on in a nearby cinema that you know he/she would enjoy watching. If your friend says that he/she has a lot of work to do, try to persuade him or her. Finally agree when and where you would meet.

Role-play No. 63.
You meet an acquaintance on the street. Greet him and mention the nice weather. Complain about the weather yesterday, when it was raining and the wind was very strong. Ask your acquaintance if he/she remembers the weather a few years ago in the winter, when there was a lot of snow. Ask him/her whether she/he knows what the weather is going to be at the weekend because your children want to go on a hike. When you notice that your bus is coming, say good-bye.

Role-play No. 48.
Your colleague Denise has a toothache, you urge her to go and see a dentist. You also offer her some painkiller. You recommend your dentist to her and you suggest asking for an appointment for today in the afternoon. You also offer to do the work instead of her while she is at the dentist.

Appendix 2 The effect of the script on the number of topic initiations

Role-play activity	Candidate topics required by the script	Candidate	Oral proficiency score	Topics		Ratio of initiated and required topics	Ratified topics		Type of interaction; assumed social roles; pragmatic functions; discourse domain
				Examiner	Candidate		Examiner	Candidate	
41	4	29	29	1	3	0.75	1	3	conversation with a tourist in the street giving information and advising; crime in the city, accommodation
		27	31	2	4	1.00	2	4	
47	1	10	35	5	1	1.00	5	1	conversation with a neighbour; giving advice and expressing opinion; arranging furniture; criticism-sensitive topic
		20	33	3	1	1.00	3	1	
		16	41	2	2	2.00	2	2	
48	5	2	46	2	5	1.00	2	5	conversation with a colleague; giving advice; offering help; going to the dentist
		8	44	2	4	0.80	2	4	
		11	36	3	5	1.00	2	5	
49	4	3	39	1	4	1.00	1	4	conversation with a neighbour; rejecting invitations; giving suggestions; going on holiday; driving a car
		7	36	2	4	1.00	2	4	
51	5	19	33	3	5	1.00	3	5	service encounter; candidate is a customer; asking for information; reserving accommodation
		23	30	2	5	1.00	2	5	
		30	29	2	4	0.80	2	4	
53	3	5	36	3	3	1.00	3	3	conversation with a friend; giving suggestions; winning the lottery; buying cars
		22	30	3	3	1.00	3	3	

(Continued over page)

Appendix 2 The effect of the script on the number of topic initiations (continued)

Role-play activity	Candidate topics required by the script	Candidate	Oral proficiency score	Topics		Ratio of initiated and required topics	Ratified topics		Type of interaction; assumed social roles; pragmatic functions; discourse domain
				Examiner	Candidate		Examiner	Candidate	
54	3	14 15 28	34 33 40	3 2 2	3 4 3	1.00 1.33 1.00	2 2 2	3 4 3	conversation with an acquaintance; offering help; unemployment; looking for a job
55	4	6 17	39 33	1 2	4 4	1.00 1.00	1 2	4 4	conversation with a friend; narration; accident; seeing a doctor
58	2	4 9	25 29	4 4	1 1	0.50 0.50	4 4	1 0	service encounter; candidate is a shop assistant; giving information
59	3	25 21	33 44	3 4	4 6	1.33 2.00	3 4	4 5	conversation with a friend; making arrangements; going to the cinema
60	3	12 26	40 30	3 1	3 3	1.00 1.00	3 1	3 3	service encounter; candidate is a costumer; for information; giving information; home insurance; description of the flat
61	3	18	44	2	6	2.00	2	6	conversation with a friend; giving suggestions; safety and security while going on holiday
62	2	13	27	3	2	1.00	3	2	conversation between teacher and parent; candidate is the teacher; narration; accident; hospitals
63	4	1 24	44 35	0 0	4 5	1.00 1.25	0 0	4 5	conversation with acquaintance; socializing; weather