The 'free conversation' and the assessment of oral proficiency

Francine Chambers
University of Southampton

Brian Richards
University of Reading

Introduction

Of the various tasks used to elicit language in an oral examination, the 'free conversation' seems to have face validity among teachers and candidates, and is used in many oral tests to obtain information on a range of personal details, views and activities. It is not generally considered a controversial task if compared, for example, with oral narrative of pictures (Chambers and Richards, 1993). However, our very familiarity with this task, as well as its semblance of authenticity, may account for the absence of any major challenge to its validity.

The status of the 'free conversation'

The National Criteria for French stated that for speaking at Higher Level, candidates should be expected to demonstrate:

> . . . the skills listed under Basic speaking over a wider range of clearly defined topic areas [. . .] and to conduct *a sustained free conversation* [our emphasis] (i.e. a conversation which has not been rehearsed) on one or more subjects as specified in the syllabus. (DES, 1985:3)

The use of the phrase 'free conversation' to describe what are, in fact, structured interviews where the teacher does most of the questioning and the candidate supplies most of the information, is obviously inaccurate unless the format of tasks can genuinely allow the *exchange* of *previously unknown* information. Nevertheless, at GCSE the quality of the conversation is a major factor in defining a Higher Level performance and MEG is the only examining group which gives more marks to any task (a narrative) other than the conversation. Other groups (WJEC, SEG) award an equal proportion of marks to the conversation and the total for the role-plays, while NEAB allocates the conversation more marks than the two Higher Level role-plays (ratio of 3:2).

It is unfortunate, therefore, that so many problems should be associated with this task. Some of these have to do with its authenticity, as indicated above. Others are due to the difficulty of achieving reliable marking (Richards, 1983; Richards and Chambers, 1992, 1993). This is sometimes attributed to the lack of expertise of teacher-markers (Hurman, 1990), although we believe the lack of well-defined and valid marking criteria is a key factor (Richards and Chambers, 1992). Most assessment criteria for the interview at Higher Level include some reference to linguistic quality and complexity, for example, but precise definition of these is lacking (Chambers and Richards, 1992). In addition, a mismatch between examiners' expectations of candidates and what native speakers would achieve is also a threat to validity. Our purpose in this article, therefore, is to explore the true scope of the GCSE conversation through a comparison with French native speakers on a similar task.

A comparison of native and non-native speaker performance

In the course of an experimental study investigating oral assessment at GCSE (Richards and Chambers, 1992), a detailed analysis of taped GCSE conversations was carried out in order to obtain objective measurements which could be compared with the marks given by teachers. We found that while teachers' impressions of features such as range of vocabulary were very accurate, there was an alarmingly low level of agreement with our objective measure of 'complexity of structures'. Following these comparisons and subsequent interviews with teachers, it became apparent that in the absence of detailed linguistic criteria which exemplified categories in the marking scheme such as 'complexity of language', 'fluency', or 'use of idiom', teachers used their own subjective criteria which did not match the language produced by candidates (Chambers and Richards, 1993). A comparison with the language of native speakers seemed therefore to be a worthwhile undertaking in order to ascertain whether the GCSE criteria matched the linguistic potential of the task. We were interested to see what kinds of linguistic complexity would occur and how native speakers would deal with GCSE topics. The intention was not to consider the interviews of native speakers as a bench-mark for relative beginners, but if competent speakers did not use certain linguistic features, it would be unreasonable to expect these even in the very best non-native performances.

Our data consist firstly of 28 audiotaped conversations of GCSE Higher Level candidates who subsequently obtained grades A to D (A = 4, B = 7, C = 7, D = 10). These lasted between 3 and 5 minutes and averaged 118 words and 23.9 clauses.

The second source is 25 five-minute videotaped interviews with French 15 year olds covering the same topics, and averaging 437 words and 56.9 clauses.[1] The format adopted for the French teenagers was as close as possible to that of the GCSE examination. The discourse is therefore typical of informal interviews rather than conversation. Interviews were carried out in two schools near Paris, one in a middle-class suburb, the other in an overspill housing area. The interviewer was herself French, but with long experience of conducting oral examinations in British secondary schools.

Both the native and non-native interviews were transcribed for computer analysis using the conventions of the Child Language Data Exchange System (CHILDES) (MacWhinney and Snow, 1990). An illustration is provided in the Appendix.

The range of topics covered

The topics discussed in the interviews with the French teenagers were those specified by the GCSE syllabuses:

- details about family and home life
- school life
- plans for further study
- leisure activities including sports
- their town and its amenities
- a journey abroad
- plans for the summer holiday.

Each interview was timed to last exactly five minutes – the instructions at GCSE state that two or three of these topics should be covered in a conversation lasting no more than seven minutes. In spite of this, we found that with all the native speakers more than three topics had to be introduced to keep the interaction going. Eleven interviews discussed six topics (the average), and only one covered as few as four. Seven interviews dealt with seven topics. It seems clear, then, that the GCSE guidelines restrict both interviewer and interviewee unnecessarily, and in a way which is unauthentic. This practice is designed so that candidates are discouraged from relying on pre-prepared material or producing recitations. However, if the number of topics was increased, this danger might actually be lessened, and it might help the interviewer to be more free-ranging in the conduct of the interview and follow the internal logic of the interview rather than adhere to a rigid plan. The fact that six topics could be dealt with easily and naturally in five minutes with native speakers suggests that the number should be increased in GCSE.

There was no conscious decision by the interviewer to introduce more than three topics but it felt very unnatural to pursue

a topic just for the sake of obtaining more language, a practice which is common in oral examinations. Trying to exhaust one topic before starting the next can lead to a very stilted exchange, and we tried not to impose a predetermined sequence of topics. This meant taking cues from the interviewee to link topics and also going back and forth from one topic to another. Most interviews started with a question asking where pupils lived and how long they had lived there. Depending on their response, this could lead to a discussion about the town or their school. The topic of home, family and relatives was often approached obliquely and never used as a starter to the interview, while the question about parents' occupation was asked only when it seemed appropriate to the topic, not as a manifestation of curiosity. For example, one girl who had said that she had only recently moved to Paris was asked why. She replied that this was because of her father's job. It was then natural to ask what her father did, before returning to the main topic which was why she preferred living in Toulouse (see Appendix). On another occasion, a girl was talking about holidays with her family. It was therefore natural to ask about brothers and sisters. There are many more examples of how the topics were intertwined and temporarily abandoned, but this can only be achieved if the interviewer is not obliged to deal with a limited range of topics.

Some themes proved inappropriate with native speakers of this age. Asking for descriptions of their house, flat, or even bedroom, as it is current practice in British examinations, became difficult in the very first interviews. The interviewer felt that this was not socially acceptable as she detected a strong element of surprise leading to very brief answers. There is clearly an invasion of privacy when examiners probe into details of candidates' home life. The only natural way of dealing with this matter occurred when interviewees had recently moved house, and could therefore make comparisons and give opinions. This situation occurs only rarely, however, and one can argue that this topic does not offer much scope, which may account for the stilted rehearsed descriptions produced by candidates in GCSE.

It was also impossible to ask French teenagers to describe their routine before going to school and to fit it into the interview in any coherent manner. The purpose of this often seems to be to encourage the use of reflexive verbs, and interestingly few of the reflexive verbs used by the French speakers, apart from *se promener*, overlapped with those taught in Britain. Equally, the interviewer felt unable to ask French pupils to describe what they were wearing, as if she were blind. As illustrated by the following extract, some of the best prepared English candidates produce unnatural semi-monologues describing, for example, how they get dressed in the morning. This candidate scored 2 out of 3 in the Higher Speaking test and obtained a grade A overall in the examination:

Teacher: A quelle heure te lèves-tu le matin?
Pupil: Euh <je lève> je me lève à sept heures et quart.
Teacher: Oui.
Pupil: Et <je> je mets mes vêtements.
Teacher: Oui qu'est-ce que tu mets comme vêtements pour venir au collège?
Pupil: Une chemise blanche.
Teacher: Oui.
Pupil: Une cravate des pantalons noirs.
Teacher: Oui.
Pupil: Et des chaussures.

Sometimes the demands of eliciting language and the need to cover the syllabus may leave teachers and examiners insensitive to the artificiality of such exchanges.

In the interviews with native speakers, the topic used most frequently as a starter was the town and its amenities. Very often remarks about the lack of entertainment locally led to talking about leisure activities. School life and plans for next year were also productive; since the French teenagers were about to change school at the end of the year, most of them could offer information about the courses they would be following.

Evidence of linguistic complexity

We tried to glean from the syllabuses and the assessment guidelines issued by examining groups which linguistic features are anticipated, and we elicited from teachers the features they look for when assessing complexity. We then quantified these features both in the performances of the non-native and native speakers. If teachers and examiners believe that certain linguistic (e.g. grammatical) features must be present in order to attain certain grades without reference to authentic native speaker performance or actual pupil performance, there is a risk that assessment is based on taken-for-granted notions of grammatical difficulty which reflect the progression adopted in textbooks. There is a risk of regarding features which are not appropriate to the task as criterial. For example, it seems reasonable, given the well-documented differences between spoken and written French in formal and informal situations (Blanche-Benveniste and JeanJean, 1987; Gadet, 1989), to have different expectations about grammatical structures and other linguistic features in a speaking test and a writing test. This should be reflected in the assessment.

It is of course very difficult to avoid the circularity of the situation: teachers concentrate their efforts on aspects of French that pupils find difficult to master and consequently value these most highly in language production. As candidates are taught how to perform assessment tasks, the best candidates will produce difficult features because they have been told that they are valued, thus giving their presence a spurious validity. This can introduce a distortion in the criteria and the dominance of the written word. This is where a study of native performances on tasks given to FL learners can help us to formulate expectations more in tune with genuine communicative needs.

Although credit is given for complexity of language at Higher Level, the vagueness of criteria makes it impossible to establish which features of language are expected at different levels of performance. The ULEAC syllabuses for 1995, for example, values 'the intrinsically difficult grammatical and structural features' (p. 11), but what does 'intrinsically' mean? A specimen tape from SEG includes the comment 'One mark (out of two) for complexity because she does use fairly complex structures'. This begs the question: which linguistic features are complex?

In the absence of unequivocal definitions, one has to assume that teachers operate on implicit notions of complexity which reflect the conventional wisdom of the profession, i.e. what is difficult for a native English speaker learning French. Fifteen of the 24 teachers who took part in our experimental marking study (Richards and Chambers, 1992) were asked which complex structures they would give credit for. Opinions varied, with some who felt that the notion of complexity was not appropriate at that level, but the most common suggestions were:
  – relative clauses
  – appropriate use of the perfect and imperfect tenses
  – nonfinite clauses: *en* + present participle; *après avoir . . .*; *avant de . . .*

In fact it must be pointed out that the present participle and the perfect infinitive are for receptive use only in some GCSE syllabuses. However, productive use of the perfect, imperfect and future tenses, as well as the relative pronouns *qui, que, ce qui, ce que* is usually expected of Higher Level candidates together with direct and indirect object pronouns and the pronouns *y* and *en*.

We have compared GCSE candidates with native speakers on: relative pronouns; object pronouns; verb usage. For both types of pronoun we have been able to use another study as a reference point. O'Connor Di Vito (1991) compared the grammar syllabus in French textbooks in the USA with the distribution and frequency of the same features in four types of French text ranging from the spoken informal to the written formal. Her informal spoken data consists of recorded interviews with four university-educated native French speakers aged 18 to 25. There are obviously differences in comparison with our interviews, carried out with examination guidelines in mind, and using younger subjects, but O'Connor Di Vito's study offers a useful perspective since there is a paucity of similar corpus studies.

Relative clauses

Table 1 indicates that French teenagers use significantly[2] more relative clauses than the GCSE candidates (6.2% versus 0.9% of all clauses). Table 2 shows the raw numbers for individual relative pronouns. There were no examples of either *qui* or *que* for the GCSE candidates who obtained grades D to B. However, they do occur occasionally for grade A candidates. By contrast,

relative clauses introduced by *qui* amounted to 59.6% of all relative clauses for the French pupils (Table 3).

Table 1: Relative clauses as a percentage of all clauses

GCSE candidates (N = 28)			French children (N = 25)			O'Connor di Vito (N = 4)	
Total	%	Range	Total	%	Range	Total	%
6	0.9	0–4	89	6.2	0–17	147	7.6

Table 2: Total relative pronouns used by each group of teenagers

	GCSE candidates (N = 28)	French children (N = 25)
qui	2	53
que	1	20
ce qui	0	5
ce que	1	5
où	2	6

Note: There were no examples of *dont* or relative pronouns following a preposition.

Table 3: Relative pronoun types as a proportion of all relative pronouns.

	French children	O'Connor di Vito
qui	59.6%	56%
que	22.5%	14%
ce qui	5.6%	7%
ce que	5.6%	16%
où	6.7%	4%
other	–	3%
Total	100%	100%

The overall frequency of relative clauses for the French teenagers (6.2%) is close to O'Connor di Vito's figure of 7.6%. However, there did appear to be some differences in the types of relative clause used (Table 3), with differences in *que* and *ce que*. Nevertheless, the age and sophistication of the native speakers undoubtedly influenced the data, as well as individual differences within O'Connor di Vito's small sample. Surprisingly, five out of the 25 French teenagers we interviewed used no relative pronouns at all. The higher proportion of *que* for the French teenagers (22.5% of relative clauses) may be influenced by the popularity of the phrase: *(Il) y a des [NOUN] + qui/que . . .* Nearly one in five of the clauses introduced by *qui* or *que* followed this pattern, a feature of informal spoken French:

   – *Y en a un que j'ai vu* (French pupil 3)
   – *Y a des choses que j'ai aimées* (French pupil 6)
   – *Y a des choses qui passent pas* (French pupil 21).

Concerning *ce qui* and *ce que*, O'Connor di Vito remarked that they were not always given much prominence in the textbooks she analysed. Because *ce que* accounted for 16% of the relative clauses in her corpus (1991:390), she argued that more emphasis should be given to this form in textbooks. However its frequency in our data (5.6%) does not support this suggestion. Nevertheless, it is interesting to note how *ce qui* and *ce que* often occur at the beginning of an utterance as a topicalisation device:

   – *Ce que j'ai aimé, c'est . . .* (French pupil 7)
   – *Moi, ce qui m'intéresse c'est que . . .* (French pupil 3)

Object pronoun usage

An analysis of the frequency of the direct and indirect object pronouns, the reflexive pronouns and the pronouns *y* and *en*, was carried out for the French teenagers and GCSE candidates.

Table 4: Relative frequency of direct object (DO), indirect object (IO), reflexive pronouns (RP) and *y* and *en*

	GCSE candidates	French children	O'Connor di Vito
DO	12.1%	8.5%	31%
IO	0%	8.1%	26%
RP	33.3%	23.9%	23%
en	0%	6.1%	13%
y (excl. *il y a*)	0%	12.6%	–
il y a	54.6%	40.8%	–
Total *y*	54.6%	53.4%	7%

Table 5: Frequency of clauses with 1 or 2 preverbal object pronouns

	GCSE candidates		French children		O'Connor di Vito
	Number	%	Number	%	%
1 object pronoun	66	9.4	225	15.9	20
2 object pronouns	0	0	11	0.7	1

Table 4 shows that apart from *y*, whose high frequency is accounted for by the use of *il y a*, it is the reflexive pronouns which are used most frequently by both groups of teenagers, while the French adults use more direct and indirect object pronouns. There were no indirect object pronouns for the GCSE candidates. The French children, on the other hand, used *me*, *nous* and *vous* but not *lui* or *leur*, the absence of *te* being due to sociolinguistic appropriateness—no child addressed the interviewer using the *'tutoiement'*.

It is worth noting with regard to pronouns that on all but two of the comparisons between the two groups of pupils there were statistically significant differences. The two exceptions were reflexive pronouns and *y*. More extensive analysis of speech in a variety of contexts is still required, but these differences, and the absence of certain pronouns in the native speaker transcripts, as well as the rarity of dual preverbal pronouns (Table 5), suggests that traditional priorities need to be re-examined.

Finite verb forms

GCSE syllabuses clearly state which tenses and moods are expected of candidates, and also differentiate between what is required for production and comprehension. These differ slightly but at Higher Level NEA requires, for example, a productive knowledge of the following: present, imperfect, perfect with *être* and *avoir* and of reflexive verbs, future, imperative. Other forms such as the pluperfect, the past historic, the conditional, the present participle, the perfect infinitive (*après avoir joué*), and the passive voice have to be understood but not used. It would appear from our teacher interviews, therefore, that some teachers go beyond what is stated in the syllabus in their assessment of 'complexity'. It was interesting that there were no instances of *après* + *perfect infinitive* for the French teenagers, but two in the GCSE group (e.g. *après m'être levé et habillé . . .*) which sound quite unnatural and appear to be part of well-rehearsed scripts.

Table 6: Tense, voice and mood as a percentage of all finite verbs				
	GCSE candidates (N = 28)		French children (N = 25)	
	Number	*%*	*Number*	*%*
Present	650	76.0	1,007	73.5
Perfect	108	16.7	193	14.0
Imperfect	14	2.2	122	8.9
Pluperfect	0	0	2	0.2
Future	28	4.3	13	1.0
Future with *aller*	6	1.0	33	2.4
All future	34	5.2	46	3.4
Subjunctive	1	0.2	8	0.6
Conditional	9	1.4	14	1.0
Passive	2	0.3	18	1.3

An analysis of all finite verbs was carried out to compare frequencies for tense, mood and voice (Table 6). As would be expected, the present indicative represents the largest share of all verb forms. This is followed by the perfect, and together these two tenses represent 93% of all finite verbs for the English children and 88% for the French children. Two differences are worth commenting upon. Firstly, to express a future event, the GCSE candidates tend to use the future tense (4.3% of all verbs) rather than *aller* (1%). By contrast the native speakers used the phrase with *aller* (2.4%) more often than the inflected form (1%). The difference in usage of *aller* is statistically highly significant.

Secondly, the imperfect appears significantly more often in the native corpus (8.9%) than in the non-native corpus (2.2%). This is in spite of the fact that the use of the imperfect varies considerably among native speakers – when narrating a journey to England, the part of the interview which perhaps lent itself most to the use of the imperfect, some used only the perfect while others used a combination of both the imperfect and the perfect. In the non-native corpus the imperfect featured mainly in the performances of the four candidates who obtained a grade A. Examining groups differ in their requirements regarding the imperfect, but it would appear from the above that its production is an appropriate and potentially useful linguistic objective for able candidates.

Discussion and conclusions

Firstly, in respect of their content, the comparison we undertook highlights the fact that the interviews, although dealing with the same topics, had different purposes with each group. Those with the native speakers had more of a communicative intent – the elicitation of unknown information which often proved to be of genuine interest. Social and sociolinguistic safeguards inhibited any attempt at purely linguistic or syllabus-based elicitation. This sensitivity to conversational appropriateness may be judged purist by teachers who are simply doing their best to encourage the production of language which can be assessed. However, one is forced to ask whether the authenticity and validity of the GCSE oral interview would not be enhanced if the interviewer did not know the candidates. It may be true that candidates prefer to be interviewed by their own teachers, although to our knowledge this has never been demonstrated, but it is not clear that they will have the opportunity to provide adequate evidence of their communicative competence under such conditions.

Van Lier (1989) claimed of the Oral Proficiency Interview that '. . . the emphasis throughout is on successful elicitation of language, not on successful conversation' (1989:101) (our emphasis). This echoes Richards' (1983) analysis of CSE 'conversations' which drew attention to teachers' frequent use of linguistic correction, praise, and other comments. Teachers who were interviewed by Richards stated that one of their main concerns was putting pupils at their ease, giving encouragement and supporting the shy and less able, rather than engaging in authentic conversation. However, some of these strategies were found to be conversationally inhibiting rather than enabling.

Secondly, the lack of clarity and agreement concerning the linguistic features which can be assessed in an interview suggests that further analysis of the real performance of learners at different levels of proficiency needs to be carried out in order to develop more coherent criteria. In addition, native performances on the various assessment tasks used in examinations have much to offer. This is not to say that the levels of achievement for the foreign-language learners should be modelled entirely on native performances. Rather the latter offer a yardstick against which the validity of tasks and the features which discriminate between levels of foreign-language proficiency can be cross-checked.

Acknowledgements

Notes

1. The videotaped interviews of French native speakers are available from the Department of Arts and Humanities in Education, University of Reading, Bulmershe Court, Earley, Reading RG6 1HY.
2. Claims of differences are limited to those which have been shown to be statistically significant after the application of appropriate statistical tests. Full details of these, and distributional statistics of all variables are available from the second author.

References

Blanche-Benveniste, C. and Jeanjean, C. (1987). *Le Français parlé*. Paris: Didier Erudition.

Chambers, F. and Richards, B. J. (1992). Criteria for oral assessment. *Language Learning Journal*, **6**, 5–9.

Chambers, F. and Richards, B. J. (1993). Oral assessment: the views of language teachers. *Language Learning Journal*, **7**, 22–6.

DES (1985). *General Certificate of Secondary Education. The National Criteria: French*. London: DES.

Gadet, F. (1989). *Le Français ordinaire*. Paris: Armand Colin.

Hurman, J. (1990). Deficiency and development in the French National Criteria for Speaking. *Francophonie*, **1**, 8–12.

MacWhinney, B. and Snow, C. A. (1990). The Child Language Data Exchange System: an update. *Journal of Child Language*, **17**, 457–72.

NEA. General Certificate of Secondary Education. Syllabus for the 1991 Examination. French. Manchester: Joint Matriculation Board.

O'Connor di Vito, N. (1991). Incorporating native speaker norms in second language materials. *Applied Linguistics*, **12**, 4, 383–96.

Richards, B. J. (1983). *Oral assessment in Modern Languages*. Unpublished MEd dissertation, University of Bristol.

Richards, B. J. and Chambers, F. (1992). *Assessing open-ended oral tasks in foreign languages: an empirical study of reliability in the GCSE examination*. University of Reading: Working Paper of the Department of Arts and Humanities in Education.

Richards, B. J. and Chambers, F. (1993). *Oral assessment in modern languages: summary of findings* (Working Paper No. 3). Reading: University of Reading, Department of Arts and Humanities in Education.

Van Lier, L. (1989). Reeling, writhing, drawling, stretching and fainting in coils: Oral proficiency interviews as conversation. *TESOL Quarterly*, **23**, 3, 489–508.

Appendix

Extract from an interview with one French teenager transcribed verbatim for computer analysis using the CHILDES software.

@Begin
@Participants: CHI pupil 3, TEA interviewer.
*TEA: alors tu veux me parler un peu de la ville de Sèvres?
*CHI: ben enfin je la connais pas beaucoup parce que j(e) suis là que depuis septembre.
*CHI: #euh bon moi j(e) trouve que c'est une petite ville.

*TEA:	tu habitais où avant?		*TEA:	et tu habitais à Toulouse avant.
*CHI:	à Toulouse.		*TEA:	ç(a) a été un grand changement.
*TEA:	ah à Toulouse!		*TEA:	tu as déménagé <à cause> [//] parce que tes parents ont dû +/?
*CHI:	bon c'est bien parce que c'est à proximité de Paris avec le train.		*CHI:	oui mon père a été muté à Vanves.
*CHI:	on y est rapidement.		*TEA:	à Vanves oui qu'est ce qu'il fait ton père?
*CHI:	bon sinon y a quand même un centre culturel qu(i) est assez bien qui propose pas mal d'activités mais #euh sinon j(e) trouve que y a +...		*CHI:	il travaille à l'INSEE.
			*TEA:	c'est quoi ça?
			*CHI:	pour les statistiques.
*CHI:	c'est pas très vivant.		*CHI:	l'Institut + national + de + statistiques.
*TEA:	tout est fermé le soir?		*CHI:	voilà.
*CHI:	oui #euh puis bon y a pas énormément d(e) magasins quand y en a i(ls) sont souvent chers.		*TEA:	et tu te plaisais à Toulouse?
			*TEA:	oui beaucoup.
*CHI:	c'est ce que j'ai ressenti en arrivant.		*TEA:	pourquoi?
*TEA:	et alors le centre culturel quelles sortes d'activités est ce qu'il propose?		@ End	
*CHI:	ben i(l) propose du sport déjà avec la mairie et #euh +...		**Key**	
			()	incomplete word
*CHI:	mais moi ce qui m'intéresse c'est que i(ls) font des travaux manuels de la poterie du dessin de la sculpture et bon ça m'intéresse ça.		#	pause
			+	joins separate words into a single unit
			+...	uncompleted utterance
*TEA:	<tu vas> [//] c'est c(e) que tu vas faire toi.		+/	interruption
*CHI:	oui.		[//]	self-correction of bracketed material