Assessing Speaking

Sari Luoma

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 2003

Printed in the United Kingdom at the University Press, Cambridge

Typeface 9.5/13pt Utopia. *System* QuarkXPress™ [SE]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

ISBN 0 521 80052 8 hardback
ISBN 0 521 80487 6 paperback

Contents

	Series editors' preface	*page* ix
	Acknowledgements	xii
1	Introduction	1
2	The nature of speaking	9
3	Speaking tasks	29
4	Speaking scales	59
5	Theoretical models	96
6	Developing test specifications	113
7	Developing speaking tasks	139
8	Ensuring a reliable and valid speaking assessment	170
	References	192
	Index	201

CHAPTER TWO

The nature of speaking

In this chapter, I will present the way speaking is discussed in applied linguistics. I will cover linguistic descriptions of spoken language, speaking as interaction, and speaking as a social and situation-based activity. All these perspectives see speaking as an integral part of people's daily lives. Together, they help assessment developers form a clear understanding of what it means to be able to speak a language and then transfer this understanding to the design of tasks and rating criteria. The more these concrete features of tests are geared towards the special features of speaking, the more certain it is that the results will indicate what they purport to indicate, namely the ability to *speak* a language.

Describing spoken language

What is special about spoken language? What kind of language is used in spoken interaction? What does this imply for the design of speaking assessments?

The sound of speech

When people hear someone speak, they pay attention to what the speaker sounds like almost automatically. On the basis of what they hear, they make some tentative and possibly subconscious judgements about the speaker's personality, attitudes, home region and native/non-native

speaker status. As speakers, consciously or unconsciously, people use their speech to create an image of themselves to others. By using speed and pausing, and variations in pitch, volume and intonation, they also create a texture for their talk that supports and enhances what they are saying. The sound of people's speech is meaningful, and that is why this is important for assessing speaking.

The sound of speech is a thorny issue for language assessment, however. This is first of all because people tend to judge native/non-native speaker status on the basis of pronunciation. This easily leads to the idea that the standard against which learner pronunciation should be judged is the speech of a native speaker. But is the standard justified? And if it is not, how can an alternative standard be defined?

The native speaker standard for foreign language pronunciation is questioned on two main accounts (see e.g. Brown and Yule, 1983: 26–27; Morley, 1991: 498–501). Firstly, in today's world, it is difficult to determine which single standard would suffice as *the* native speaker standard for any language, particularly so for widely used languages. All languages have different regional varieties and often regional standards as well. The standards are valued in different ways in different regions and for different purposes, and this makes it difficult to choose a particular standard for an assessment or to require that learners should try to approximate to one standard only. Secondly, as research into learner language has progressed, it has become clear that, although vast numbers of language learners learn to pronounce in a fully comprehensible and efficient manner, very few learners are capable of achieving a native-like standard in all respects. If native-like speech is made the criterion, most language learners will 'fail' even if they are fully functional in normal communicative situations. Communicative effectiveness, which is based on comprehensibility and probably guided by native speaker standards but defined in terms of realistic learner achievement, is a better standard for learner pronunciation.

There are, furthermore, several social and psychological reasons why many learners may not even *want* to be mistaken for native speakers of a language (see e.g. Leather and James, 1996; Pennington and Richards, 1986): a characteristic accent can be a part of a learner's identity, they may not want to sound pretentious especially in front of their peers, they may want recognition for their ability to have learned the language so well despite their non-native status, and/or they may want a means to convey their non-native status so that if they make any cultural or politeness mistakes, the listeners could give them the benefit of the doubt because of their background.

Pronunciation or, more broadly, the sound of speech, can refer to many features of the speech stream, such as individual sounds, pitch, volume, speed, pausing, stress and intonation. An important question is whether all of these can be covered under one rating criterion. Moreover, should the focus be on accuracy of pronunciation or expressiveness of the speaker's use of voice, or both? The solutions depend on the purpose for which the scores will be used and the importance of the sound of speech for that purpose. If there are many other rating criteria besides pronunciation, fitting accuracy and effectiveness into a criterion like 'naturalness of pronunciation' may be the only option. If the sound of speech is a main focus in the assessment, evaluating aspects of it separately gives material for more detailed feedback.

A focus on pronunciation accuracy is attractive because it can be judged against a norm and, even if the norm is not easy to define given the discussion above, gross deviations from it are easy enough to notice. Since accuracy is related to comprehensibility, it is often at least one aspect of a pronunciation criterion, but comprehensibility is much more than accuracy. It often includes speed, intonation, stress and rhythm, all of which may be more important for the overall comprehensibility of the talk than the accuracy of individual sounds. If the emphasis in the assessment is on ability to create meaning in discourse, the developers might want to evaluate 'interactional efficiency'. This would encompass the examinees' use of stress and intonation to highlight important phrases, or to suggest in what particular way (e.g. ironically) their words should be interpreted. In yet other contexts, they might want to focus on 'expressiveness' as indicated by the general texture of the talk, the speaker's use of speed and pausing, and variations in pitch, tone and volume. This might be especially relevant in tasks such as creative storytelling or certain kinds of role plays, where liveliness of expression is a central element in task performance. Thus, in designing assessment criteria, the developers need to consider the type of information about the sound of speech that they need. They also have to make sure that their tasks give enough material for rating these features, and that they develop the criteria that serve their needs.

Spoken grammar

Both first and second language learners' progress is often tracked according to the grammatical forms that they can produce accurately (see e.g. Larsen-Freeman and Long, 1991: 38–41 for a discussion on this point). In

general, learners are seen to proceed from knowing a few structures to knowing more and more, from using simple structures to using more complex ones, and from making many errors to making few if any at all. Learner grammar is handy for judging proficiency because it is easy to detect in speech and writing, and because the fully fledged grammars of most languages are well known and available for use as performance standards. However, the grammar that is evaluated in assessing speaking should be specifically related to the grammar of speech.

Written sentences, spoken idea units

A major difference between speech and writing is that speakers do not usually speak in sentences. Rather, speech can be considered to consist of **idea units**, which are short phrases and clauses connected with *and*, *or*, *but* or *that*, or not joined by conjunctions at all but simply spoken next to each other, with possibly a short pause between them. The grammar of these strings of idea units is simpler than that of the written language with its long sentences and dependent and subordinate clauses. This is because speakers are trying to communicate ideas that listeners need to comprehend in real time, as they are being spoken, and this means working within the parameters of the speakers' and listeners' working memory. Idea units are therefore usually about two seconds or about seven words long, or shorter (Chafe, 1985). The units are usually spoken with a coherent intonation contour, and they are often limited on both sides by pauses or hesitation markers. Many idea units are clauses with a verb phrase, a noun phrase and a prepositional phrase, but some of them do not contain a verb, and sometimes an idea unit is started by one speaker and completed by another.

Grammar in planned and unplanned speech

There are of course some situations where complex grammatical features and a high degree of written language influence are not only common but also expected and highly valued. Examples of this include speeches, lectures, conference presentations, and expert discussions where speakers represent their institution or their profession. These situations involve **planned speech** (Ochs, 1979), where the speakers have prepared and possibly rehearsed their presentations in advance, or they express well-

thought-out points and opinions, which they may have voiced many times before. **Unplanned speech**, in contrast, is spoken on the spur of the moment, often in reaction to other speakers. It is particularly in unplanned speech that short idea units and 'incomplete sentences' are common, although even in planned speech, idea units are usually shorter than in writing, because the speakers know that their talk has to be understood by listeners in real time.

The concepts of planned and unplanned speech are closely connected to another factor that affects the grammar of speech, namely the level of formality of the speaking situation. Situations that involve planned speech tend to be relatively formal, whereas unplanned speech situations can range from formal to informal. Formal situations require more written-like language with more complex grammar, whereas informal situations call for more oral-like language with strings of short phrases and short turns between speakers.

For assessing speaking, it is in fact useful to see the differences between spoken-like and written-like language as a continuum, with highly oral language at one end and highly literate language at the other (Tannen, 1982). In addition to grammar, oral and literate speech differ in their pronunciation and choice of vocabulary, among other things. Test designers can design tasks for various places on the oracy–literacy continuum by varying things like planning time and the kinds of speaker roles and role relationships that they include in the tasks.

Two examples

To illustrate the nature of grammar in speech, let us look at two examples of transcribed talk. The first comes from Brown *et al.* (1984). A young British postgraduate is describing what happened when she ordered a snack from room service in an American hotel. The second word, *er*, is a voiced hesitation sound, which could also be spelled *eh* or *uh*. A single plus sign indicates a short pause and two plus signs a longer pause. The speaker is being interviewed by a researcher to give material for a study. In other words, the speakers are relative strangers and the speaking situation is fairly formal.

> and + er + I was pretty exhausted and I phoned up room service and said that I wanted a sandwich + + nothing's ever straightforward in America (laugh) – 'what kind of sandwich' + + I said 'well' er + hummed and hawed +

> and he said 'well + there's a list in your drawer' + 'in your chest of drawers' + + so I had a look at it and gawd there was everything (laugh) you know + and I saw roast beef + so I phoned back and said I would have a roast beef sandwich (laugh) + and a glass of milk + so an hour later + + nothing happened you see + so I phoned him up again and yes + they were coming + and in walked this guy with a tray + an enormous tray and a steel covered + plate + dinner plate you see + so I lifted that up + and I've never seen anything like it + + there was three slices of bread lying on this plate + and there was I counted eight slices of roast beef + hot roast beef + with gravy and three scoops of mashed potato round the outside + an enormous glass of milk and a glass of water
>
> (Brown *et al.*, 1984: 17)

Brown *et al.* point out that this is a very competent storyteller who structures long turns confidently. Even so, the chunks of language are mostly clause-sized, they are strung together with the conjunction *and* or follow one another without conjunctions, and the vocabulary is rather simple. There are short phrases, pauses, repetitions and reformulations. On two occasions, the speaker does not follow number concord. A non-native speaker in a test situation might be marked down for such a performance. Similarly, the shortness of phrases and the absence of 'advanced' vocabulary might affect the rating. Yet this is a natural sample of native speaker storytelling.

The second example is from unplanned and informal dialogue. Three British female students (S01–S03) are chatting in the kitchen of a house they are renting.

1 <S01> Does anyone want a chocolate or anything?
2 <S02> Oh yeah yes please
3 <S03> Yes please
4 <S02> [laughs]
5 <S03> [laughs]
6 <S01> You can have either a Mars Bar, Kit-Kat or erm cherry Bakewell
7 <S03> Oh erm it's a toss-up between [<S02> [laughs]] the cherry
8 Bakewell and the Mars Bar isn't it?
9 <S01> Well shall I bring some in then cos you might want another one
10 cos I don't want them all, I'm gonna be
11 <S03> Miss paranoid about weight aren't you?
12 <S01> Yes but you know

13 <S03> You're not fat Mand
14 <S01> I will be if I'm not careful
15 <S02> Oh God
. . .

(Carter and McCarthy, 1997: 85)

This is typical casual talk. Most of the turns consist of one short meaning unit and speakers change quickly. In her longest turn, Student 1 uses the causal connector *cos* (lines 9 and 10) and, at the last juncture, simple stringing along. Other than that, the coherence in the discourse is created by thematic linking. On line 11, Student 3 shortens her turn by omitting the subject and the verb, *you are*, but her meaning is still fully comprehensible. The use of phrases like *you know* and *it's . . . isn't it* make the turns characteristically spoken-like and informal.

The internal structure of idea units

Many spoken idea units are clauses, grammatically speaking, but the way that idea units are structured is often slightly different from standard written clauses. Two structures that clearly belong to spoken-like language use are **topicalisation** and **tails**.

Topicalisation, or thematic fronting, gives special informational emphasis to the initial element of a clause in informal speech, as in *Joe, his name is* (Quirk and Greenbaum, 1976). Topicalisation breaks the standard word order of written language. In speaking, the word order does not seem 'broken' in any sense, however, since the aim is to emphasise the topic. It is a very frequent feature of informal talk, and McCarthy and Carter (1995: 211) suggest that the explanation is that it has significant interpersonal meaning. It often indicates that an important topic of conversation is to follow. Thus, their example of *That house in the corner, is that where you live?* is presumably an introduction into a discussion on the house or the neighbourhood, something that the speaker is reminded of upon seeing the house.

Tails, in turn, are noun phrases that come at the end of a clause. In a way they are the mirror image of topicalisation, in that they repeat a pronoun that has been used earlier in the clause. By using tails, speakers can emphasise the comment they make at the beginning of the clause, and still make it clear what they are talking about, as in *It's very nice, that*

road through Skipton to the Dales (McCarthy and Carter, 1995). The comment that the speaker expresses at the beginning of the clause is often an evaluation, such as *he's quite a comic, that fellow, you know*, but not always, as in *'cos otherwise they tend to go cold, don't they, pasta*. Tails emphasise the point made at the beginning of the clause, and at the same time, they create an informal tone in the talk.

Both topicalisation and tails follow clear patterns, which can be formed into 'rules' for talk. The patterns are characteristically spoken-like, but not traditionally taught in language classes or talked about in grammars. They create an impression of naturalness and interpersonal involvement in spoken discourse, and if examinees use them appropriately they could be rewarded for it. However, they cannot be punished for not using them, because they are not obligatory in any context.

To summarise the discussion on spoken grammar, speech is organised into short idea units, which are linked together by thematic connections and repetition as well as syntactic connectors. The most frequent connectors are coordinating conjunctions (*and*, *or*, *but*, etc.). Some speaking situations call for more literate grammar with complete clauses and subordination. These are typically formal speaking situations, which may involve prepared talk such as a presentation.

Speakers may emphasise points by topicalisation, which means starting their turn with the main topic and making the word order unusual, or tails, which means using the natural emphasis of the beginning of their turn for a comment or an evaluation and putting the noun that they are making their comment on at the end of the clause. This gives talk a spoken flavour. It adds interpersonal and evaluative tones, which is typical for spoken discourse.

Words, words, spoken words

Many rating scales for speaking include descriptions of vocabulary use, and at the highest levels these often talk about being able to express oneself precisely and providing evidence of the richness of one's lexicon. This can indeed be important in professional contexts or when trying to convey detailed information. Well-chosen phrases can also make descriptions or stories vivid, and learners who can evoke the listener's feelings deserve to be credited for their ability. However, very 'simple' and 'ordinary' words are also very common in normal spoken discourse, and using these naturally in speech is likewise a marker of highly advanced

speaking skills (see e.g. Read, 2000). Moreover, there is a core of phrases and expressions that are highly typical for speaking, which contribute to the listener's impression of the speaker's fluency. They work at the interpersonal level by keeping the conversation going and developing the relationship between the speakers. This aspect of word use should also be rewarded in assessing speaking.

Specific and generic words

Some forms of written language require the use of specific words to make it clear what is being talked about. For example, a written instruction for how to adjust an office chair states: *Use the ball adjustment to move the lumbar support to a position where it supports the back.* If the same instruction were given orally in a hypothetical set of video-taped instructions, similar words might well be used, but with added visual support. In an interactive speaking situation, the same instructions would probably sound quite different. The speakers would use many generic words such as *this one* / *that one*, *the round thing*, *move*, *put*, *fine*, and *good*. The instruction-giver and the chair-user would probably exchange several turns to make sure that the task got done properly.

Generic words are very common in spoken interaction. Even though they are not precise, they are fully comprehensible in the speaking situation because they talk about people, things or activities that can be seen or because they are familiar to the speakers. They make spoken communication quick and easy, and few people would find anything strange about this in their mother tongue. Generic words may also come naturally to second-language learners, but in a foreign language context where learners have few opportunities to speak the language outside the classroom this feature of spoken language may be harder to notice and learn. Assessment designers can help this by including descriptions of effective use of generic words in rating scales. This sends the message to learners and raters that generic words are important for the naturalness of talk.

Another common feature of interactive and relatively informal talk is the use of vague words like *thing*, *thingy*, *thingummy* and *whatsit* when the speaker cannot think of the word he or she needs to use. Channell (1994) has investigated the use of these words in English, but she refers to other researchers' examples for French and presumes that all languages have a set of such words. Vague words help the speaker go on regardless of the missing word, and at the same time they appeal to the listener to

understand and supply it if they can. They are natural in informal talk, and if learners use them appropriately they deserve to be rewarded for it.

Fixed phrases, fillers and hesitation markers

Speakers also need to know words, phrases and strategies for creating time to speak. These are sometimes called fillers or hesitation markers, and they include expressions such as *ah*, *you see*, *kind of*, *sort of*, and *you know*, as well as whole expressions such as *That's a good question*, or *Now let me see*. Speakers often also use repetition of their own words, or of those used by the previous speaker, to achieve the same purpose, i.e. to keep the floor while formulating what they want to say. These expressions are very common in native speaker speech, but for some reason their appearance in test performances by foreign language learners is sometimes frowned upon. When writing assessment scales, test developers should perhaps consider if examinees who manage to use such expressions successfully in a test situation should be rewarded for it instead.

Fixed conventional phrases are also used for other purposes in talk than creating time. Examples of these include responses like *I thought you'd never ask* or *I'm doing all right, all things considered.* The phrases either always have the same form, or they constitute a formula where one or two slots can be filled by various terms (e.g. *What a nice thing to say*, *What a horrible thing to say*). They have been called lexicalised sentence stems by Pawley and Syder (1983), and lexical phrases by Nattinger and DeCarrico (1992). They are easy for speakers to use because they come almost automatically when a relevant situation arises and because, once a speaker begins such a phrase, saying it will give them time to judge the situation, perhaps plan how they want to put what they want to say next, or think of something else to say.

Word use in studies of assessing speaking

There are a few studies that support the relevance of the above-mentioned characteristics of speech for assessing speaking. Towell *et al.* (1996), for instance, show that learners' use of lexical phrases is connected with a listener's experience of the speaker's fluency. That is, if two learners use an approximately similar lexicon in their speech, but one of

them also uses a range of fixed phrases while the other does not, the one who uses the phrases is perceived to be the more fluent of the two. And if a learner uses a wide range of fixed phrases, listeners tend to interpret that as proof of a higher level of ability than when a learner is using a few stock phrases in all kinds of contexts.

Hasselgren (1998) investigated learners' use of filler words with three groups of speakers: British native speaker schoolchildren of 14–15 years of age, and two ability groups of Norwegian schoolchildren of the same age, high and low. Hasselgren called the verbal phenomenon she investigated **'smallwords'**, which she defined as 'small words and phrases, occurring with high frequency in the spoken language, that help to keep our speech flowing, yet do not contribute essentially to the message itself' (p. 4). Her results support the case that the more smallwords a learner uses, the better their perceived fluency.

Nikula's (1996) study of a range of similar expressions, which she considered under the heading of 'pragmatic force modifiers', adds the observation that even advanced learners produce a much narrower range of 'spoken-like' expressions and discourse markers than native speakers. She studied the speech habits of her non-native speakers also in their mother tongue, and was thus able to prove that the difference was not caused by personal or cultural communication style but was truly related to language ability. Together, these studies strongly support the case that the use of spoken-like words is important in speaking performance.

Slips and errors

Normal speech contains a fair number of slips and errors such as mispronounced words, mixed sounds, and wrong words due to inattention. If the listeners notice, they tend to pardon native speakers because they believe them to 'know', but in the speech of second or foreign language learners slips and errors mysteriously acquire special significance. Their slips can signal lack of knowledge, and this seems to be important for many listeners. While there are some errors that only learners make, such as using *no* + verb to express negation in English (*I no write*) or violating simple word order rules, there are others that are typical for all speakers. Assessment designers may have to provide special training to raters to help them outgrow a possible tendency to count each 'error' that they hear.

Processing and reciprocity

Bygate (1987) summarises the above features of spoken language use by contrasting them with writing. He suggests that the differences can be explained with reference to two sets of conditions: processing and reciprocity. Processing conditions are connected with time, the crucial difference being that, while writers can generally take as much time as they need to produce their text and readers can pace their reading (on a separate occasion) to their needs and interests, the processes of speaking and listening are most often intertwined and happen under the pressure of ever ticking time. The solution to this is reciprocity, by which Bygate means that speakers react to each other and take turns to produce the text of their speech together. This helps the speakers with the processing demands of speech, but it also has a social dimension in that their phrases and turn-taking patterns create and reflect the social relationship between them.

Speaking as meaningful interaction

Speaking and spoken interaction

Teaching and testing experts often talk about speaking as a technical term to refer to one of the various skills that language learners should develop and have. This type of speaking tends to be seen as something that *individuals* do. It is legitimate, and for educational purposes useful, to see speaking in this way too, because it is true that individuals speak, and an important part of language use is personal. Nevertheless, it is also important to remember that speaking forms a part of the shared social activity of talking.

In a typical spoken interaction, two or more people talk to each other about things that they think are mutually interesting and relevant in the situation. Their aim can be to pass the time, amuse each other, share opinions or get something done, or they can aim to do several of these and other things at once. The point in their interaction is that they do these things together. Each participant is both a speaker and a listener; they construct the event together and share the right to influence the outcomes – which can be both shared and individual.

The openness of meanings in interaction

When people talk and listen to each other, they are driven by a quest for meaning, but meanings are not always clear and explicit. Moreover, people know that anything that is said has not just one meaning but many: it says something about some topic or other, but it also indicates the speaker's attitude towards the topic and towards the other participant(s) and reflects the speaker's knowledge about the history of the topic, his or her views about what might be happening next, and more. As discussed earlier in this chapter, this kind of non-explicitness appears in many verbal forms, and it has many motivations.

The openness of meanings is not only a convenience in speech; it is also an effective strategy for speakers. They can avoid committing themselves to a statement or attempt to find out how the listener feels about the topic before proceeding. They can try to find out what the listener already knows, what he or she is prepared to accept or understand, and what the best strategy might be to persuade the listener to accept their point of view. For example, someone may introduce the topic of going to the movies and listen for reactions before raising the idea that this group of people might want to go out to a particular show that weekend. A member of the group who has other commitments may then say that she likes the idea but does not know yet because something urgent may come up with work or something. This is strategically a highly skilful way of using language, and speakers, at least in a language that they live in, use vague expressions for these purposes automatically, because they are a fundamental part of spoken communication.

Language learners' attempts at vagueness may cause peculiarities in discussions. They may simply sound strange because they do not know how interpretations are appropriately left open. Alternatively, they may use the right kind of strategies, but the listener may fail or refuse to recognise their intention. The natural appearance of open meanings in a discussion involving a learner is a clear sign of highly advanced speaking skills, as it proves that the learner is able to produce successful indirect utterances and that the listener is willing and able to interpret and act upon this in the context of the interaction. This kind of naturalness may not be easy for raters to notice unless their attention is specifically drawn to it through training, rater instructions and wordings of assessment scales.

Variation within spoken language use

Although spoken language as a whole can be contrasted with written language, there is also a lot of systematic variation within spoken language use. The analysis of this is a part of discourse analysis, which is a vast area of research in applied linguistics. A thorough introduction to the field is not attempted here (for this, see e.g. Schiffrin, 1994; McCarthy and Carter, 1997). In the sections below, I discuss three significant areas in spoken discourse for assessing speaking: purposes of talk, the speaking situation and speaker roles. They help assessment developers think about what *kinds* of talk need to be included in their assessment, and thus focus the assessment on the right construct.

Talking to chat and talking to inform

One way in which speech events differ from each other is the purpose for which the people are talking to each other. With this approach to analysing talk, Brown *et al.* (1984) characterise two extremes: chatting or listener-related talk, and information-related talk. They stress that this is not a clear-cut dichotomy but rather a dimension along which sections of talk will be situated. Moreover, both types of talk can occur in one and the same speech event; in fact, this is what normally happens. Information-related talk often comes sandwiched between social chat, and a social chat can easily turn into a serious discussion.

Brown *et al.* (1984) define **chatting** as the exchange of amicable conversational turns with another speaker. The primary purpose is to make and maintain social contact, to oil the social wheels, and thus chatting forms a large part of anyone's social life. Skilful chatting involves finding a fluid stream of topics that the speakers find sufficiently interesting to take up, and on which they can find a shared angle. The topics are not necessarily discussed very deeply, and it is more important to create a positive atmosphere and to agree than to express oneself precisely or to be completely truthful. Chatting in one's first language can only really become strenuous on a social rather than linguistic dimension. However, we are not all equally socially gifted, and not all equally good at chatting. Yet, as Brown *et al.* point out, chatting in the first language is so closely connected to personality and individual communication styles that it cannot really be taught.

In language teaching, however, some focus on at least the basic phrases for chatting is necessary, and when chatting occurs it involves the learn-

ers' personalities and their social behaviour. It also involves their culture, as appropriate topics for chatting differ between cultures. This causes some difficult dilemmas for assessing speaking.

If the assessment situation involves chatting in the target language, as it often does during the initial and final stages of the interaction and maybe in other stages as well, the developers have to consider how far it is necessary or justified to assess a learner's personality or social skills. It is perhaps realistic to accept that it is impossible to exclude the social aspects of personality from the assessment completely, but it may be possible to avoid highlighting some sociability aspects of chatting. One way in which this is attempted in many tests is by telling raters not to assess the initial stages of a test interaction. In some assessments, however, chatting might be the main focus of assessment, especially in learning contexts if it has been taught recently. Then it would be important to make sure that the participants know what kind of talk they should aim for to do well on the test.

The other end of Brown *et al.*'s (1984) dimension of kinds of talk, **information-related talk**, refers to speech aimed at transferring information on a particular topic. People's talk at work mostly belongs to this end of the continuum, for example policemen talking to witnesses, nurses and doctors talking to patients and to each other, or factory workers interacting with each other. Information-related talk is also very much a part of teaching–learning situations, and these kinds of tasks are very often included in assessment situations as well. As with chatting, Brown *et al.* make the point that native speakers vary in their ability to produce information-related talk, but in contrast to chatting they feel that the techniques for more effective information-related talk can be taught.

The most important point about information-related talk is getting the message across and confirming that the listener has understood it. Establishing common ground, giving the information in bite-sized chunks, logical progression, questions, repetitions and comprehension checks help speakers reach this aim. These features should therefore appear in examinee performances on information-related tasks, and they may help explain why some of them do better on the test than others. Once the developers analyse some learner performances to find out exactly how the performances at different ability levels differ, for example whether weaker performers fail to establish common ground or do not sequence the information logically, they can use these concepts in rating scales to indicate how raters can tell performances at different levels apart.

Apart from basic information-structuring skills, information-related talk also requires other skills for organising communication and making it easy to follow. Brown and Yule (1983) discuss five types of information-oriented tasks for language learning, including telling a story from pictures. In this task, speakers need to be able to identify the main characters and refer to them consistently, describe the main events and activities, and mention any significant changes in characters, time or locations. Stories become more difficult to tell the more characters there are who are difficult to tell apart, so that a story involving three girls is more difficult than one involving two girls, which in turn is more difficult than a story involving a girl and a boy. They also become more difficult the more events there are, and the more changes in characters, time or locations that the story involves. Good storytelling routines are important for speakers, as one of the most common types of chatting involves personal stories about accidents or embarrassing situations (Rintell, 1990; Jones, 2001). To be able to convey the nature of the situation and the speaker's emotions, learners need to have routinised the basic storytelling skills discussed above. I will return to information-related tasks and task difficulty in Chapter 3.

Talking in different social situations

One set of features that has an influence on what gets said in a speech event and how it is said is the social and situational context in which the talk happens. Hymes (1972) has helpfully summarised these concerns into a framework that forms the acronym SPEAKING. The framework has so many categories because it is meant to be applicable to a large variety of social situations, but all of them may not be relevant for every situation.

The SPEAKING framework lists the potential social and contextual factors influencing speech as:

Situation	The physical setting (for instance a classroom) and the nature of the event (for instance an end-of-term test of speaking).
Participants	Speaker, hearer, audience, etc.; for instance, two examinees, an interlocutor and an assessor (whether present in the situation or absent, only listening to the interaction afterwards from tape).

Ends	Conventional outcomes of the event, if any. For instance, accomplishing whatever task is the goal of the event, or producing a test score and verbal feedback. The ends also include the individual participants' goals, such as exposing the strengths and weaknesses of the examinees' speaking ability, showing one's ability to speak a foreign language at its best, or making fair and equitable assessments.
Act sequence	The form and content of speech acts: the content of what is said, and the way it is said; how each act is spoken, and the sequence of acts in the discourse.
Key	Tone, manner, or spirit of act; for instance, supportive, friendly, open, formal, impersonal, tentative, withdrawn.
Instrumentalities	Channel or mode, e.g. spoken, written, pre-recorded. Forms of speech: dialects, accents, and varieties used.
Norms	Norms of interpretation and norms of interaction, such as right/responsibility to initiate topics, ask questions, express views, ask for clarification, explain, elaborate.
Genre	Categories such as a joke, lecture, description, instruction, storytelling, presentation.

Assessment developers can use this framework when they make initial plans for their test. It will help them describe the test construct in some detail. Later in the development work, the framework can guide the comparison of individual test administrations against each other, which is important for fairness. If there are clear differences, the scores may not be comparable. The categories can be used to compare talk in the test with speaking situations that the examinees are likely to meet outside the test. This is significant because the assessment developers probably want to predict the examinees' ability to cope with the non-test situations on the basis of their test results. If there are differences, the predictions may not be safe.

The importance of any differences is a value judgement, however. Hymes's framework can make the analysis of the differences more systematic, and thus help make this judgement more informed. The key questions that the assessment developers have to answer are: is there a