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*