13 Testing grammar and vocabulary

Testing grammar

Why test grammar?

Can one justify the separate testing of grammar? There was a time when this would have seemed a very odd question. Control of grammatical structures was seen as the very core of language ability and it would have been unthinkable not to test it. But times have changed. As far as proficiency tests are concerned, there has been a shift towards the view that since it is language skills that are usually of interest, then it is these which should be tested directly, not the abilities that seem to underlie them. For one thing, it is argued, there is more to any skill than the sum of its parts; one cannot accurately predict mastery of the skill by measuring control of what we believe to be the abilities that underlie it. For another, as has been argued earlier in this book, the backwash effect of tests which measure mastery of skills directly may be thought preferable to that of tests which might encourage the learning of grammatical structures in isolation, with no apparent need to use them. Considerations of this kind have resulted in the absence of any grammar component in some well-known proficiency tests.

But most proficiency tests which are administered on a large scale still retain a grammar section. One reason for this must be the ease with which large numbers of items can be administered and scored within a short period of time. Related to that, and at least as important, is the question of content validity. If we decide to test writing ability directly, then we are severely limited in the number of topics, styles of writing, and what we earlier referred to as ‘operations’, that we can cover in any one version of the test. We cannot be completely confident that the sample chosen is truly representative of all possibilities. Neither can we be sure, of course, that a (proficiency) grammar test includes a good sample of all possible grammatical elements. But the very fact that there can be so many items does put the grammar test at an advantage.

Even if one has doubts about testing grammar in a proficiency test, there is often good cause to include a grammar component in the achievement, placement and diagnostic tests of teaching institutions. It seems unlikely that there are many institutions, however ‘communicative’ their approach, that do not teach some grammar in some guise or other.
Wherever the teaching of grammar is thought necessary, then consideration should be given to the advisability of including a grammar component in achievement tests. If this is done, however, it would seem prudent, from the point of view of backwash, not to give such components too much prominence in relation to tests of skills, the development of which will normally constitute the primary objectives of language courses.

Whether or not grammar has an important place in an institution’s teaching, it has to be accepted that grammatical ability, or rather the lack of it, sets limits to what can be achieved in the way of skills performance. The successful writing of academic assignments, for example, must depend to some extent on command of more than the most elementary grammatical structures. It would seem to follow from this that in order to place students in the most appropriate class for the development of such skills, knowledge of a student’s grammatical ability would be very useful information. There appears to be room for a grammar component in at least some placement tests.

It would be particularly useful to know for individual students (and for groups of students) what gaps exist in their grammatical repertoire. Unfortunately, as explained in Chapter 3, for the moment at least, to obtain reliable and more than rudimentary diagnostic information of this kind would involve the development and administration of long and unwieldy tests. Computer-based testing may bring progress in this field in the near future.

WRITING SPECIFICATIONS

For achievement tests where teaching objectives or the syllabus list the grammatical structures to be taught, specification of content should be quite straightforward. When there is no such listing it becomes necessary to infer from textbooks and other teaching materials what structures are being taught. Specifications for a placement test will normally include all of the structures identified in this way, as well as, perhaps, those structures the command of which is taken for granted in even the lowest classes.

SAMPLING

This will reflect an attempt to give the test content validity by selecting widely from the structures specified. It should also take account of what are regarded for one reason or another as the most important structures. It should not deliberately concentrate on the structures which happen to be easiest to test.
WRITING ITEMS

Many, if not most, language testing handbooks encourage the testing of grammar by means of multiple choice items, often to the exclusion of just about any other method. Chapter 8 of this book gave a number of reasons why testing within institutions should *avoid* excessive use of multiple choice. In this section, there will be no examples of multiple choice items. Though all the techniques can be given a multiple choice format, the reader who attempts to write such items can often expect to have problems in finding suitable distractors.

The techniques to be presented here are just three in number: paraphrase, completion, and modified doze. Used with imagination, these should prove sufficient for most grammar testing purposes. They have in common what should be appreciated as a highly desirable quality: they require the student to supply grammatical structures appropriately and not simply to recognise their correct use. Ways of scoring are suggested in the next section.

PARAPHRASE

These require the student to write a sentence equivalent in meaning to one that is given. It is helpful to give part of the paraphrase in order to restrict the students to the grammatical structure being tested.

Thus:

1. Testing passive, past continuous form.
   
   When we arrived, a policeman was questioning the bank clerk.
   When we arrived, the bank clerk ..........................  

2. Testing present perfect with *for*.
   
   It is six years since I last saw him.
   I ................................................................. six years.

COMPLETION

This technique can be used to test a variety of structures. Note how the context in a passage like the following, from the Cambridge First Certificate in English *Testpack 1*, allows the tester to elicit specific structures, in this case interrogative forms.

In the following conversation, the sentences numbered (1) to (6) have been left incomplete. Complete them suitably. Read the whole conversation before you begin to answer the question.

(Mr. Cole wants a job in Mr. Gilbert’s export business. He has come for an interview.)

**Mr Gilbert:** Good morning, Mr Cole. Please come in and sit down. Now let me see. (1) Which school ........................ ?
Mr Cole: Whitestone College.
Mr Gilbert: (2) And when .................................................? 
Mr Cole: In 1972, at the end of the summer term. 
Mr Gilbert: (3) And since then what .................................? 
Mr Cole: I worked in a bank for a year. Then I took my present job, selling cars. But I would like a change now. 
Mr Gilbert: (4) Well, what sort of a job ..................................? 
Mr Cole: I’d really like to work in your Export Department. 
Mr Gilbert: That might be a little difficult. What are your qualifications? 
(5) I mean what languages .................besides English? 
Mr Cole: Well, only a little French. 
Mr Gilbert: That would be a big disadvantage, Mr Cole. (6) Could you tell me why ...........................................? 
Mr Cole: Because I’d like to travel and to meet people from other countries. 
Mr Gilbert: I don’t think I can help you at present, Mr Cole. Perhaps you ought to try a travel agency.

MODIFIED CLOZE

Testing prepositions of place

John looked round the room. The book was still .......... the table. The cat was ........ the chair. He wondered what was ........ the boxthe .......... telephone.

And so on.
Testing articles
(Candidates are required to write *the*, *a* or NA — No Article.)

In England children go to … … school from Monday to Friday. ………… school that Mary goes to is very small. She walks there each morning with ………… friend. One morning they saw ………… man throwing …………. stones and …………. pieces of wood at …………. dog. …………. dog was afraid of …………. man.

*And so on.*

Testing a variety of grammatical structures
(The text is taken from Colin Dexter *The Secret of Annexe 3.*

When the old man died, …………. was probably no great joy …………. heaven; and quite certainly little if any real grief in Charlbury Drive, the pleasantly unpretentious cul-de-sac …………. semi-detached houses to which he …………. retired.

Testing sentence linking
(A one-word answer is required.)

The council must do something to improve transport in the city. …………. they will lose the next election.

**SCORING PRODUCTION GRAMMAR TESTS**

The important thing is to be clear about what each item is testing, and to award points for that only. There may be just one element, such as the definite article, and all available points should be awarded for that; nothing should be deducted for non-grammatical errors, or for errors in grammar which is not being tested by the item. For instance, a candidate should not be penalised for a missing third person s when the item is testing relative pronouns; opened should be accepted for opened, without penalty.

If two elements are being tested in an item, then points may be assigned to each of them (for example present perfect form and since with past time reference point). Alternatively, it can be stipulated that both elements have to be correct for any points to be awarded, which makes sense in those cases where getting one element wrong means that the student does not have full control of the structure.

For valid and reliable scoring of grammar items of the kind advocated here, careful preparation of the scoring key is necessary. A key for items in the previous section might be along the following lines.
**Paraphrase**

1. Must have *Was being …………ed* — 1 point
2. *I haven’t* [or *have not*] — 1 point
   
   *seen* — 1 point
   
   *for six years* — 1 point

Alternatively, all three elements could be required for a point to be given, depending on the information that is being looked for.

**Completion** The Cambridge First Certificate Testpack, from which the completion items were taken, is rather less explicit than one might wish (though no doubt in an actual examination scorers would be given more detailed instructions). Thus:

Whole mark for each question formed correctly with auxiliary and inversion (except No. 6 ‘why you want to’ etc.) and with tense suitable to context. Bonus for continuous tense in No. 3. The American phrase ‘quit college’ for ‘leave school’ allowed.

**Modified cloze** As was mentioned in Chapter 8, it is helpful if candidates are required to put their responses in numbered spaces at the right-hand side of the page. Checking against the key is then very straightforward.

<table>
<thead>
<tr>
<th>Prepositions</th>
<th>Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. on</td>
<td>1.NA</td>
</tr>
<tr>
<td>2. under/beneath</td>
<td>2.The</td>
</tr>
<tr>
<td>3. in/inside</td>
<td>3.a</td>
</tr>
<tr>
<td>4. near/beside</td>
<td>4.a</td>
</tr>
<tr>
<td>5. NA</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Various structures</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. there</td>
<td>7.a</td>
</tr>
<tr>
<td>2. in</td>
<td>8.The</td>
</tr>
<tr>
<td>3. of</td>
<td>9.the</td>
</tr>
<tr>
<td>4. had</td>
<td></td>
</tr>
</tbody>
</table>

Sentence linking otherwise.

**Testing vocabulary**

**Why test vocabulary?**

Any doubts about the advisability of testing grammar apply equally to the testing of vocabulary. Clearly knowledge of vocabulary is essential to the development and demonstration of linguistic skills. But that does not necessarily mean that it should be tested separately.
Similar reasons may be advanced for testing vocabulary in proficiency tests to those used to support the inclusion of a grammar section (though vocabulary has its special sampling problems). However, the arguments for a separate component in other kinds of test may not have the same strength. One suspects that much less time is devoted to the regular, conscious teaching of vocabulary than to the similar teaching of grammar.

If there is little teaching of vocabulary, it may be argued that there is little call for achievement tests of vocabulary. At the same time, it is to be hoped that vocabulary learning is taking place. Achievement tests that measure the extent of this learning (and encourage it) perhaps do have a part to play in institutional testing. For those who believe that systematic teaching of vocabulary is desirable, vocabulary achievement tests are appreciated for their backwash effect.

The usefulness (and indeed the feasibility) of a general diagnostic test of vocabulary is not readily apparent. As far as placement tests are concerned, we would not normally require, or expect, a particular set of lexical items to be a prerequisite for a particular language class. All we would be looking for is some general indication of the adequacy of the student’s vocabulary. The learning of specific lexical items in class will rarely depend on previous knowledge of other, specified items. One alternative is to use a published test of vocabulary. The other is to construct one’s own vocabulary proficiency test.

**WRITING SPECIFICATIONS**

How do we specify the vocabulary for an achievement test? If vocabulary is being consciously taught, then presumably all the items thereby presented to the students should be included in the specifications. To these we can add all the new items that the students have met in other activities (reading, listening, etc.). A subsequent step is to group the items in terms of their relative importance.

We have suggested that a vocabulary placement test will be in essence a proficiency test. The usual way to specify the lexical items that may be tested in a proficiency test is to make reference to one of the published word lists that indicate the frequency with which the words have been found to be used (see *Further reading*).

**SAMPLING**

Words can be grouped according to their frequency and usefulness. From each of these groups items can be taken at random, with more being selected from the groups containing the more frequent and useful words.

**ITEM WRITING**

**Recognition** This is one testing problem for which multiple choice can be recommended without too many reservations. For one thing, distractors
are usually readily available. For another, there seems unlikely to be any serious harmful backwash effect, since guessing the meaning of vocabulary items is something which we would probably wish to encourage. However, the writing of successful items is not without its difficulties.

Items may involve a number of different operations:

1. **Synonyms**

Choose the alternative (A,B,C,D) which is closest in meaning to the word on the left of the page.

\[
\text{gleam} \quad A. \text{gather} \quad B. \text{shine} \quad C. \text{welcome} \quad D. \text{clean}
\]

The writer of this item has probably chosen the first alternative because of the word *glean*. The fourth may have been chosen because of the similarity of its sound to that of *gleam*. Whether these distractors would work as intended would only be discovered through pretesting.

Note that all of the options are words which the candidates are expected to know. If, for example, *welcome* were replaced by *groyne*, most candidates, recognising that it is the meaning of the stem (gleam) on which they are being tested, would dismiss *groyne* immediately. On the other hand, the item could have a common word as the stem with four less frequent words as options:

\[
\text{shine} \quad A. \text{malm} \quad B. \text{gleam} \quad C. \text{loam} \quad D. \text{snarl}
\]

Note that in both items it is the word *gleam* which is being tested.

2. **Definitions**

*loathe* means

A. dislike intensely
B. become seriously ill
C. search carefully
D. look very angry

Note that all of the options are of about the same length. It is said that test-takers who are uncertain of which option is correct will tend to choose the one which is noticeably different from the others. If *dislike intensely* is to be used as the definition, then the distractors should be made to resemble it. In this case the item writer has included some notion of intensity in all of the options.

Again the difficult word could be one of the options.

One word that means to *dislike intensely* is

A. growl
B. screech
C. sneer
D. loathe
3. **Gap filling (multiple choice)**

Context, rather than a definition or a synonym, can be used to test knowledge of a lexical item.

The strong wind …………… the man’s efforts to put up the tent.

A. disabled  
B. hampered  
C. deranged  
D. regaled

Note that the context should not itself contain words which the candidates are unlikely to know.

**PRODUCTION**

The testing of vocabulary productively is so difficult that it is practically never attempted in proficiency tests. Information on receptive ability is regarded as sufficient. The suggestions presented below are intended only for possible use in achievement tests.

1. **Pictures**

The main difficulty in testing productive lexical ability is the need to limit the candidate to the (usually one) lexical item that we have in mind, while using only simple vocabulary ourselves. One way round this is to use pictures.

Each of the objects drawn below has a letter against it. Write down the names of the objects:

A  ………………………………..  
B  ………………………………..  
C  ………………………………..  
D  ………………………………..  
E  ………………………………..  
F  ………………………………..

This method of testing vocabulary is obviously restricted to concrete nouns which can be unambiguously drawn.
2. **Definitions**  
This may work for a range of lexical items:

A ............. is a person who looks after our teeth.  
............ is frozen water.  
............ is the second month of the year.

But not all items can be identified uniquely from a definition: any definition of say *feeble* would be unlikely to exclude all of its synonyms. Nor can all words be defined entirely in words more common or simpler than themselves.

3. **Gap filling**  
This can take the form of one or more sentences with a single word missing.

Because of the snow, the football match was ............ until the following week.

I ............ to have to tell you this, Mrs Jones, but your husband has had an accident.

Too often there is an alternative word to the one we have in mind. Indeed the second item above has at least two acceptable responses (which was not intended when it was written!). This problem can be solved by giving the first letter of the word (possibly more) and even an indication of the number of letters.

**Postscript**

This chapter should end with a reminder that while grammar and vocabulary contribute to communicative skills, they are rarely to be regarded as ends in themselves. It is essential that tests should not accord them too much importance, and so create a backwash effect that undermines the achievement of the objectives of teaching and learning where these are communicative in nature.

**READER ACTIVITIES**

1. Construct items to test the following:  
   Conditional: *If ...... had ....*, ..... *would have .....*  
   Comparison of equality  
   Relative pronoun *whose*  
   Past continuous: ...... *was -ing, when ......*  

Which of the techniques suggested in the chapter suits each structure best? Can you say why?
2. Produce two vocabulary tests by writing two items for each of the following words, including, if possible, a number of production items. Give each test to a different (but comparable) group of students (you could divide one class into two). Compare performance on the pairs of items. Can differences of performance be attributed to a difference in technique?

beard sigh bench deaf genial
greedy mellow callow tickle weep

Further reading

West (1953) is a standard word list.