

Guidelines for Writers of Listening Tests

© BRITISH COUNCIL, HUNGARY

1 Introduction

These Guidelines were originally written for secondary school and university teachers, who had been trained in language testing to write test items for the British Council School-leaving Examination Reform Project. We hope that they will be also useful for English teachers in general, and will help them write better classroom listening tests. And even if teachers do not have to or do not wish to write their own listening test tasks, understanding how listening tests are designed should help language teachers understand the ready-made tests and textbook exercises better. These teachers will also prepare their students for modern European English examinations in a more efficient way.

On the next pages, you will find both general and specific guidelines for the construction of Listening test items and tasks. Each set of guidelines includes the following subtopics: *text selection*, *task selection*, *task design*, and *rubrics*.

Guidelines like these should normally be attached to a set of test specifications. Each examination should have a set of specifications which describes what the test is aiming to test and how it does this. Specifications are normally produced in a handbook together with sample test papers so that those who are interested in the test can obtain a clear idea of what the test is like. For example, you should find the information about the test's purpose, who the test-takers are (or are likely to be), what is the difficulty level of the test (for example, in terms of the Common European Framework), which language skills and sub-skills are being tested, and how they are being tested, how many sections there are in one paper, how much time is given for doing a task and a test as a whole, which task types should be used, how many points are awarded for each item and each task, what is the length of a (listening) text, what is the language of the rubrics (the students' mother tongue or the target language), how rubrics should be worded, what criteria should be used for marking, and so on.

Test writers of international, national, or classroom tests have to be familiar with the test specifications they are writing items for. Even classroom teachers can write their own specifications which can be a simple and practical outline of the test. They may indicate how a 45-minute lesson (during which the test will be administered) will be divided up, what the aim of the test is, how much time will be given to each section of the test, what language skills will be tested, and what

the items will look like, i.e. whether the students will have to write short answers or longer texts, circle the given choices or match the parts of the texts, etc.

Specifications for testing listening skills in the classroom setting may look something like this:

Listening test

Date: *20th April 2010*

Aim or purpose of the test: *diagnostic: to identify weaknesses in students' listening skills*

Age group: *18/19*

Sex: *20 male and 10 female;*

Difficulty level: *intermediate, approximately at the B1 level.*

First/second foreign language: *first foreign language (English)*

Study programme: *Grammar school (600 45-minute lessons of English of instruction so far)*

Number of test sections/parts: *2*

Skills tested: *listening for specific information and listening for the main points*

Target language situation: *listening to the radio and understanding the main points and some specific details*

Task type: *integrative; both objectively and subjectively assessable*

Test method: *completion/short answer*

Text type: *radio news/radio interview with a writer*

Text length: *3 minutes each (listened to twice)*

Number of items: *10 per task*

Rubrics: *According to a standard template, in English. Students will have been familiarised with these in class. One example is given at the beginning;*

Time: *Students have 30 seconds to read the task before they start listening and 30 seconds after the listening to finalise their answers. Total test time: 30 minutes*

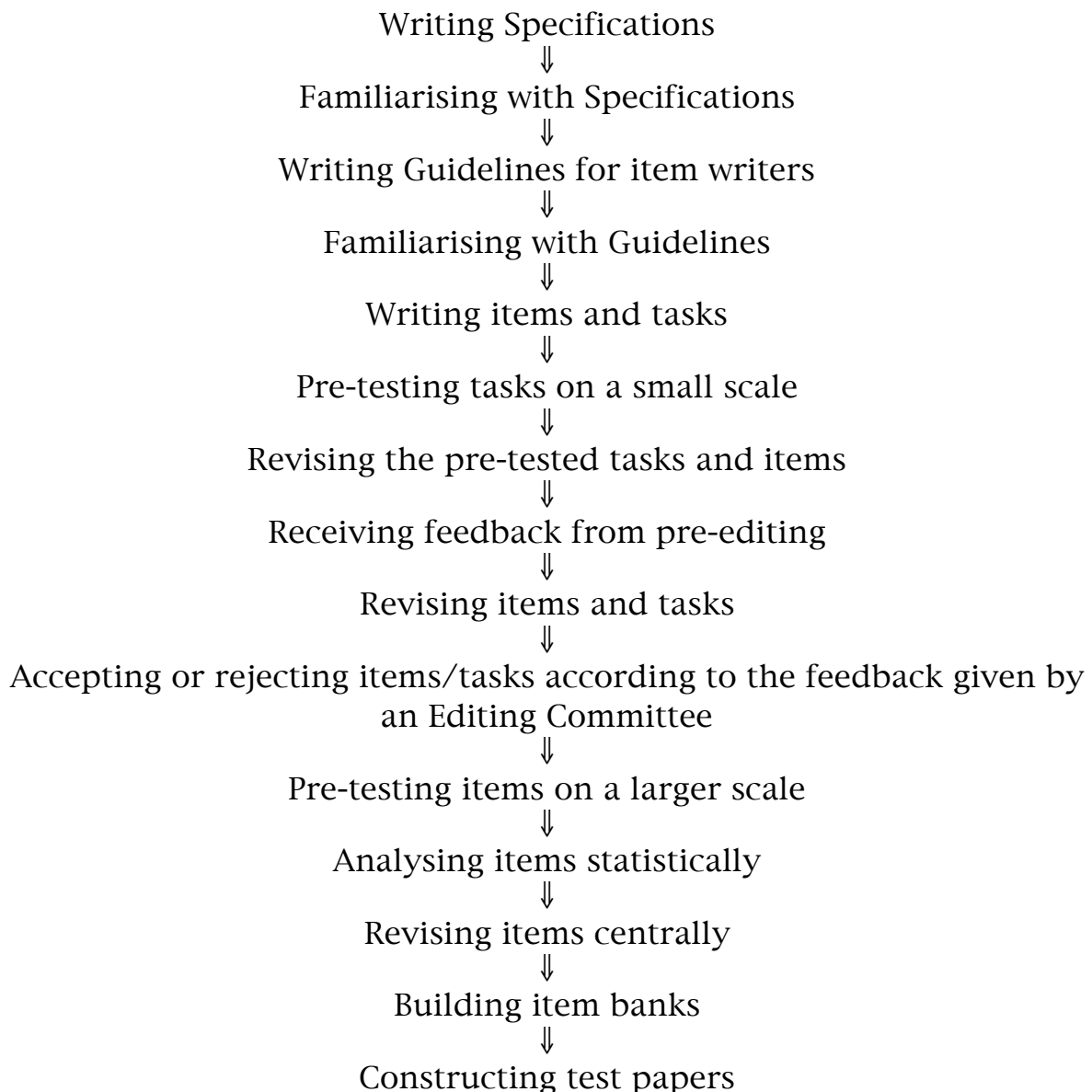
Criteria: *Answer key (acceptable and unacceptable) is provided; meaning more important than spelling;*

Length of answers: *in the short-answer task, answers longer than 4 words will be penalised*

It is clear, however, that for standardised, large-scale tests that are intended to be widely distributed and therefore widely generalised, test specifications must be much more formal and detailed. The test construction process, of which item writing is only a part, must follow internationally accepted procedures. These procedures were also

respected in producing the tasks for this book.

There are several phases that test construction has to go through:



It is clearly very difficult for a classroom teacher to design tests according to the procedure mentioned above in all its stages. However, what all teachers should remember is that nobody can write good tasks on their own and nobody can write tests without mistakes. It is therefore essential to ask a few teacher colleagues to **do** the test tasks and give you their feedback on task suitability, item difficulty, instructions, etc. Any feedback a teacher receives on his or her test tasks is invaluable. If you want to get some more information on the items and tasks, then you should pre-test your own testing materials on students of similar abilities that the test is intended to assess. These you

may find in your colleagues' classes or at a neighbouring school.

After marking the tests, it is a good idea to analyse the results, not only the task results but also item scores. In this way, you will be able to find out which items performed well and which not and why. Next time, you may avoid making the same mistake(s) again.

After compiling a number of suitable and well-written tests, you may start building your own bank of tasks. For example, a bank for testing listening skills may be organised according to topic, text type, text source, task type, listening skill, task difficulty, and so on. Such a collection of tasks will make assessment easier for you, more reliable and fairer for the students.

2 General Guidelines

2.1 Text selection

2.1.1 All texts must be authentic.

2.1.2 Texts must not be taken from printed material in other examinations or English language teaching materials.

2.1.3 The topic of the texts must be accessible to the students' age group.

2.1.4 Texts must not be offensive, distressing or violent.

2.1.5 Texts must be of a suitable length.

2.1.6 Texts must be of a suitable level of difficulty.

2.1.7 A text must have a title unless it is used as part of the item.

2.1.8 The same text can be used with more than one task. For example, a listening text with different tasks could be used for intermediate and advanced levels.

2.2 Task selection

2.2.1 Have a clear idea what skill and sub-skill(s) you intend to test and select the task and texts/prompts accordingly.

2.2.2 The task type must be one which is familiar to the students who are tested.

2.2.3 Tasks must be accessible to the students' age group.

2.2.4 Tasks must not be offensive, distressing or violent.

2.2.5 The time allocated for each task must be sufficient for somebody who has the ability being tested to complete the task comfortably within the time limit.

2.2.6 Students must be able to see easily how the task relates to the text.

2.2.7 No more than one testing technique should be used in each task.

Do not switch from multiple choice to single-word answers in a single task, for example.

2.3 Task design

- 2.3.1 There should be a minimum of five items in a task.
- 2.3.2 Items must follow the text sequence.
- 2.3.3 Items must be spread evenly through the text.
- 2.3.4 It must not be possible to answer any item without reference to the text. This must be checked carefully.
- 2.3.5 In non-sequencing tasks, items must not be interdependent – students should not need one answer in order to find another.
- 2.3.6 Items must not overlap.
- 2.3.7 Two items must not have similar answers.
- 2.3.8 Items must have a complete answer key on a separate sheet. All possible good answers and predictable wrong answers should be provided.
- 2.3.9 The items in each task must be numbered sequentially.
- 2.3.10 One item in each task must have an answer provided as an example. The example item, written in italics, should be marked 0 (zero) and should precede the other items.
- 2.3.11 The example answer must be provided in the form in which candidates are expected to write their answers and, where appropriate, it must be hand-written.
- 2.3.12 “Find the wrong answer” type items are not acceptable in multiple-choice tasks.
- 2.3.13 Each item must score one point.

2.4 Rubrics

- 2.4.1 Rubrics should be in English. They must be clear, simple and brief. Redundancies must be deleted. Do not use exclamation marks and meta-language.
- 2.4.2 Conform to standard rubrics, for example, the ones in this textbook.
- 2.4.3 Indicate clearly what students have to do i.e. what they have to write and where.
- 2.4.4 The rubrics should be in bold.
- 2.4.5 Put the total number of points available for each task (in other words, the number of items) at the end of the task, in bold.

2.5 Checklist

Text

- 2.5.1 Is the text type worth testing?
- 2.5.2 Is the text user-friendly in style and acceptable in content?
- 2.5.3 Is the length of text appropriate for your students?
- 2.5.4 Is the level of difficulty appropriate? Do any alterations to the text conform to the Guidelines?

Task

- 2.5.5 Is the rubric clear, brief and in bold?
- 2.5.6 Has the example been provided for each task?
- 2.5.7 Is there a complete answer key for each task?
- 2.5.8 Is the task type familiar to students?
- 2.5.9 Is the task type accessible and acceptable?
- 2.5.10 Is the number of items acceptable?
- 2.5.11 Are the items spread evenly through the text?
- 2.5.12 Are all items properly labelled: questions with numbers, options with letters?
- 2.5.13 Can each question be answered independently?
- 2.5.14 Do items follow the text sequence?
- 2.5.15 Do answer keys reflect the form in which the answer is expected?
- 2.5.16 Does the answer key include alternative answers (if applicable)?

2.6. Acknowledgements

The following reference materials have been used in compiling these guidelines:

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge Language Teaching Library. Cambridge: Cambridge University Press.

Milanovic, M. (Ed.). (1996). *User's Guide for Examiners*. Council of Europe, CC-LANG, 10.

3 Listening Paper

3.1 Text selection

- 3.1.1 Use spoken texts that sound like spoken English. Where, for example, pauses or fillers would be normal in conversation, they should also be used in the recording. You may record the listening texts from radio stations off-air, or from audio files available on the Internet. You may also record your own texts, using a good microphone and a digital recorder.
- 3.1.2 Avoid using texts from television programmes. These texts have been recorded for TV viewers who listen and watch at the same time. Therefore, these texts consist of two sources of information, one is the picture and the other is the sound. If one is missing, comprehension becomes very difficult or even impossible.
- 3.1.3 Select texts whose topic is accessible to the students' age group.
- 3.1.4 Try to select texts that are likely to be interesting for students. Be aware that humorous texts, for example radio comedy shows, might be enjoyable and easy for you but very difficult for students.
- 3.1.4 Avoid offensive and distressing topics, such as war, fatal diseases, family tragedies, etc.
- 3.1.5 Avoid well-known audio recorded novels or any other well-known radio programmes as text sources.
- 3.1.6 Avoid listening materials that were recorded and published for teaching.
- 3.1.7 Select texts of suitable length. If the text is a part of a longer text, make sure it stands alone and does not contain references to previous or following parts of the original texts.
- 3.1.8 Do not use more illustrations than necessary. When illustrations are used, they should be clear and photocopiable.
- 3.1.9 Do not use texts with content covered in detail in any listening exercise included in a language textbook. You do not want to test knowledge of content, but ability to understand recorded texts.
- 3.1.10 Select texts that do not assume a high level of background knowledge (social, cultural, factual etc.).
- 3.1.11 If you decide to produce a text of your own, you must seek the advice of a fellow teacher and, if at all possible, of a native speaker as to the accuracy and the acceptability of the vocabulary and grammatical structures used in the spoken texts.
- 3.1.12 Ensure that in case of sequencing tasks the text lends itself to sequencing.
- 3.1.13 Do not use listening texts with many speakers participating in a conversation, especially if it is not absolutely clear who says

what, or whether the relationships are not explicitly stated.

3.1.14 Ideally each speaker's voice should be clearly distinguished from the others. For example, a male could be paired with a female, or a younger with an older person.

3.2 Task selection

3.2.1 Have a clear idea about the purpose of each listening task, which listening skill is to be tested in a particular task (e.g. listening for the main idea or listening for specific information), and why a particular task type is used.

3.2.2 The task should reflect the way in which people usually listen to the type of text used.

3.2.3 The students should be familiarised with the task type in advance.

3.2.4 Tasks should not test students' memory.

3.2.5 Consider applying more difficult texts with easier tasks and vice versa.

3.3 Task design

3.3.1 Produce tasks from tapes, not tape-scripts.

3.3.2 There should be a minimum of five items in a task.

3.3.3 The wording of the items should be such that it is easy to understand and does not require too much reading.

3.3.4 Make sure items cannot be answered correctly without listening to the text.

3.3.5 Use clear signposts to help students in tasks where texts consist of a number of segments.

3.3.6 Provide an example, written in italics and marked with (0), at the beginning of each task. Example answers should be given in a hand-written form where appropriate.

3.3.7 Allow enough time for students to familiarise themselves with items before listening to the text. Intermediate: 4 seconds per item; advanced: 3 seconds per item.

3.3.8 Allow enough time for students to finalise their answers after the text is finished: 3 seconds per item.

3.3.9 Make sure that each question can be answered independently; that is, the answer to one question does not depend on the answer to another question (except in sequencing tasks).

3.3.10 Select items in such a way that there is enough time for students to answer before having to concentrate on the text again for the next item.

3.3.11 Spread items evenly. Do not have long stretches of text

where no item needs to be answered.

- 3.3.12 Sequence items following the sequence of ideas in the text.
- 3.3.13 Number items and use letters to mark options.
- 3.3.14 Make sure there is only one possible answer to each question. If more than one answer is required, this fact must be clearly indicated, and they must count as different items.
- 3.3.15 Do not focus only on one type of factual information (numerals, countries, nationalities etc.).
- 3.3.16 Do not base items on proper names and avoid items the answer to which depends on understanding proper names (e.g. people's names, geographical names, etc.).
- 3.3.17 Provide four options for multiple choice items but make sure they do not require too much reading.
- 3.3.18 Provide more options than pairs in matching tasks.
- 3.3.19 Make sure you base completion items on factual information, and not on subjective judgment. The missing words should be concrete nouns and verbs rather than adjectives because adjectives usually have synonyms and many other plausible replacements.
- 3.3.20 In sentence completion tasks the sentences candidates hear should not be the same as the sentences they have to complete.
- 3.3.21 In productive tasks (i.e. in completion tasks and short-answer question tasks) avoid items whose answer is very difficult to spell.
- 3.3.22 Design one task to fit on one page or two facing pages.
- 3.3.23 Design tasks with clear and consistent layout.
- 3.3.24 Provide answer keys reflecting the form in which answers are expected. Give all possible or acceptable answers. Unacceptable answers that are likely to occur should also be indicated. In tasks where the correct answer may be phrased in various ways, the answer key should provide content-related instructions concerning acceptable answers.
- 3.3.25 Weigh items equally (1 item = 1 point).

3.4 Rubrics

- 3.4.1 Use simple, standardised rubrics.
- 3.4.2 Word all instructions in English.
- 3.4.3 Present all instructions, identical with the taped rubrics, on the test sheet.
- 3.4.4 Inform candidates what they are going to hear: the text type and/or source.
- 3.4.5 Inform candidates in the rubrics when they are expected to

- answer, and how many times the text will be played.
- 3.4.6 Indicate clearly what candidates are expected to do i.e. what they have to write and where.
- 3.4.7 Give guidance on the length of appropriate answers.
- 3.4.8 Rubrics must be written in bold.
- 3.4.9 Indicate the total number of points available for each task at the end of the task in bold.

List of reference books

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge University Press.
- Anderson, A. & Lynch T. (1988). *Listening*. Oxford University Press.
- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford University Press.
- Blundell, L. & Stokes, J. (1981). *Task Listening*. Cambridge University Press.
- Buck, G (2001) *Assessing listening*. Cambridge University Press
- Harrison, A. (1983). *A Language Testing Handbook*. Macmillan.
- Heaton, J. B. (Ed.). (1982). *Language Testing*. Modern English Publications.
- Heaton, J.B. (1998). *Writing English Language Tests*. Longman.
- Hughes, A. (2003). *Testing for Language Teachers*. Second edition. Cambridge University Press.
- Rixon, S. (1986). *Developing Listening Skills*. Macmillan.
- Rost, M. (1990). *Listening in Language Learning*. Longman.
- Rost, M. (1991). *Listening in Action*. Prentice Hall.
- Scarborough, D. (1984). *Reasons for Listening*. Cambridge University Press.
- Underwood, M. (1989). *Teaching Listening*. Longman.
- Ur, P. (1992). *Teaching Listening Comprehension*. Cambridge University Press.
- Weir, C. (1993). *Communicative Language Testing*. Prentice Hall.
- Weir, C. (1993). *Understanding and Developing Language Tests*. Prentice Hall.
- Practice test booklets and practice books for the Cambridge FCE and CAE examinations