



Figure 1: Cognitive processing model adapted from Geranpayeh and Taylor (Eds) 2013

The listening ability tested by Key and Preliminary exams spans across several levels on the CEFR, i.e. lower and higher A2 for Key, and lower and higher B1 for Preliminary. At the lower A2 level, the listener is expected to 'understand phrases and expressions related to areas of most immediate priority (e.g. very basic personal and family information, shopping, local geography, employment), provided speech is clearly and slowly articulated' (Council of Europe 2001:32). At the higher level of A2, the listener is expected to 'understand enough to be able to meet the needs of a concrete type provided speech is clearly and slowly articulated' (Council of Europe 2001:32).

As far as B1 level descriptors are concerned, the lower-level B1 descriptor states that a listener 'can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure etc., including short narratives' (Council of Europe 2001:66). At the higher B1 level, the listener 'can understand straightforward factual information about common everyday or job related topics, identifying both general messages and specific details, provided speech is clearly articulated in a generally familiar accent' (Council of Europe 2001:66).

The Cambridge English view is that test takers at A2 and B1 levels need to focus a great deal of attention at the more local levels of processing (input decoding, lexical search and syntactic parsing) and have little spare attentional capacity to give to the wider areas of meaning construction and discourse construction (Geranpayeh and Taylor (Eds) 2013). This had also been reflected in the design of the previous listening tasks of the Key and Preliminary tests, and had been revealed when the underlying task constructs were analysed. The analysis of the Key and Preliminary listening component also revealed one significant issue as far as construct coverage is concerned, namely the lack of test items that demand listening for gist.

During the analysis of the cognitive validity of the listening component of Cambridge English Qualifications, the extent to which different levels of cognitive processing are targeted in Key and Preliminary Listening sub-tests was investigated (Geranpayeh and Taylor (Eds) 2013). Following the analysis, it was concluded that, as far as lower-proficiency tests are concerned, there is a strong focus on perceptual-level processing. However, it should be borne in mind that the performance of lower-proficiency listeners, both in test conditions and in the real world, is largely conditioned by their ability to successfully employ compensatory strategies which enable them to infer general meaning even if the input has not been completely mastered. It would therefore seem relevant to include a number of items that would allow the test takers to demonstrate their ability to report the main point made by the speaker without grasping the full content of the message. In other words, what was missing from Key and Preliminary Listening sub-tests were items that demand listening for gist.

Following the publication of *Examining Listening* (Geranpayeh and Taylor (Eds) 2013), a number of changes to the Listening component of the Key and Preliminary tests have been implemented during the revision process. They are detailed in the two following sections of this paper.

Revised Key Listening test

A summary of changes made to the Key Listening test, including a comparison with the current version (until end of 2019) of the test, can be found in Table 1. The table charts the revisions through two trialling sessions.

Table 1: Revised Key Listening test format

<i>Current: Key Listening (until end 2019)</i>		<i>Trialling version 1: Key Listening</i>		<i>Trialling version 2: Key Listening Revised test format (from January 2020)</i>	
<i>Timing</i>	<i>Content</i>	<i>Timing</i>	<i>Content</i>	<i>Timing</i>	<i>Content</i>
22 minutes (approx.) plus 8 minutes transfer time	Part 1 Five discrete 3-option multiple-choice items with visuals. 25–60 words.	29 minutes (approx.) plus 6 minutes transfer time	Part 1 Six discrete 3-option multiple-choice items with visuals. Short neutral or informal dialogues. 40–60 words.	25 minutes (approx.) plus 6 minutes transfer time	Part 1 Five discrete 3-option multiple-choice items with visuals. Short neutral or informal dialogues. 40–60 words.
	Part 2 Longer informal dialogue. Matching task. Five items and eight options. 150–170 words.		Part 2 Longer informal dialogue. Matching task. Five items and eight options. 160–180 words.		Part 2 Gap-fill. Longer neutral or informal monologue. Five gaps to fill with one word or a date or number or a time. 150–170 words.
	Part 3 Five 3-option multiple-choice items. 160–180 words.		Part 3 Six 3-option multiple-choice items. Longer informal or neutral dialogue. 190–220 words.		Part 3 Five 3-option multiple-choice items. Longer informal or neutral dialogue. 160–180 words.
	Part 4 Gap-fill. Five gaps to fill with one or more words or a number. 150–170 words.		Part 4 Six discrete 3-option multiple-choice items with written options. Two or three B1 lexical/structural items to be used to test candidate's understanding of the main idea, message, gist or topic. 40–60 words.		Part 4 Five discrete 3-option multiple-choice items with written options. Two or three B1 lexical/structural items to be used to test candidate's understanding of the main idea, message, gist or topic. 40–60 words.
	Part 5 Gap-fill. Five gaps to fill with one or more words or a number. 150–170 words.		Part 5 Gap-fill. Longer neutral or informal monologue. Seven gaps to fill with one or two words or a number. 190–230 words.		Part 5 Longer informal dialogue. Matching task. Five items and eight options. 160–180 words.

The most significant revision made to the test format was to Part 4 of the test, which was changed and trialled as discrete multiple-choice items that are aimed to test a candidate's understanding of the main idea, message, topic or gist, in line with the Cambridge English approach (Geranpayeh and Taylor (Eds) 2013). The addition of this task has allowed the construct of the Key Listening test to be expanded to include listening for gist.

In Phase 1 of the trialling, the number of items in the test was increased from 25 to 30 across the five parts of the test. This was to improve the accuracy and reliability of the test as well as to have a better coverage of the construct of listening comprehension at this level.

The main focus of this first trial was the newly designed Part 4 task – consisting of six discrete 3-option multiple-choice items with written options. There was also a change to the way the Part 3 task works. In the current test format, Part 3 is a cued dialogue which works on a need-to-know basis where one of the speakers cues in the questions and the other gives the key. This was amended so that the need-to-know basis was removed, yet questions are still cued in, but now by either speaker; both speakers now give the keys as well, thus better replicating a real-world dialogue between the two speakers. The range of question types was increased to test a candidate's ability to identify specific information, feelings and opinions.

Key and Key for Schools Listening trialling took place in the first quarter of 2016 in various locations and included the following language groups: Spanish, Portuguese, French, Polish, Russian, Serbian, Ukrainian, Dutch, Urdu and Malay. Several of the trials were carried out with post-assessment focus groups conducted by Assessment Managers linked to the various papers.

Results for the new Part 4 task

New Part 4 task results of the trial-tests can be seen in Table 2.

Table 2: Part 4 acceptance rates

<i>Key</i>	<i>Test 1</i>	<i>Test 2</i>	<i>Test 3</i>	<i>Test 4</i>	<i>Overall</i>
Number of trial-test candidates	140	135	168	194	637
Part 4 acceptance rate	100%	66%	83%	17%	66%

<i>Key for Schools</i>	<i>Test 1</i>	<i>Test 2</i>	<i>Test 3</i>	<i>Test 4</i>	<i>Overall</i>
Number of trial-test candidates	214	192	202	319	927
Part 4 acceptance rate	66%	83%	83%	0%	58%

All tasks were reviewed post-trial by a panel of experts who currently work on Key Listening, with decisions taken on whether the tasks were performing to the right level being based on both statistical information and expert judgement. Overall, 16 out of 24 (66%) for Key and 14 out of 24 (58%) for Key for Schools of the new Part 4 tasks were found to be at the right level and were accepted as suitable tasks for live test use. Tasks that were accepted at post-trial review were generally found to be limited in their reading load on the paper, limited on the complexity of the message itself, and on the complexity of language used within the script. Messages within the script were generally given in a linear manner with no backtracking from the speakers, and grammatical forms used were at the right level for candidates who can perform well at A2 level. Referencing within the texts needed to be clear and limited for the task to be accepted. The following accepted task is an example that is working at the required A2 level.