# Dictation

**Dictation** is a widely researched genre of assessing listening comprehension. In a dictation, test-takers hear a passage, typically of 50 to 100 words, recited three times: first, at normal speed; then, with long pauses between phrases or natural word groups, during which time test-takers write down what they have just heard; and finally, at normal speed once more so they can check their work and proofread. Here is a sample dictation at the intermediate level of English.

## Dictation
*First reading (natural speed, no pauses, test-takers listen for gist):*
The state of California has many geographical areas. On the western side is the Pacific Ocean with its beaches and sea life. The central part of the state is a large fertile valley. The southeast has a hot desert, and north and west have beautiful mountains and forests. Southern California is a large urban area populated by millions of people.

The state of California *II* has many geographical areas. *II* On the western side *II* is the Pacific Ocean *II* with its beaches and sea life. *II* The central part of the state *II* is a large fertile valley. *II* The southeast has a hot desert, *II* and north and west *II* have beautiful mountains and forests. *II* Southern California *II* is a large urban area *II* populated by millions of people.

*Third reading (natural speed, test-takers check their work).*

Dictations have been used as assessment tools for decades.
The difficulty of a dictation task can be easily manipulated by the length of the word groups, the length of the pauses, the speed at which the text is read, and the complexity of the discourse, grammar, and vocabulary used in the passage.

**Scoring is another matter.** Depending on your context and purpose in administering a dictation, you will need to decide on scoring criteria for several possible kinds of errors:

- **spelling error only, but the word appears to have been heard correctly**
- **spelling and/or obvious misrepresentation of a word, illegible word**
- **grammatical error (For example, test-taker hears *I can't do it,* writes *I can do it.)***
- **skipped word or phrase**
- **permutation of words**
- **additional words not in the original**
- **replacement of a word with an appropriate synonym**

Dictation seems to provide a reasonably valid method for integrating listening and writing skills and for tapping into the cohesive elements of language implied in short passages.

If the word-groups in a dictation are relatively long (more than five-word segments), this method places a certain amount of load on memory and processing of meaning (Buck, 2001, p. 78).

Finally, one can easily question the authenticity of dictation: it is rare in the real world for people to write down more than a few chunks of information (addresses, phone numbers, grocery lists, directions, for example) at a time.

**Despite these disadvantages, the practicality of the administration of dictations, a moderate degree of reliability in a well-established scoring system, and a strong correspondence to other language abilities speaks well for the inclusion of dictation among the possibilities for assessing extensive (or quasi-extensive) listening comprehension.**

## Communicative Stimulus-Response Tasks

Another - and more authentic- example of extensive listening is found in a popular genre of assessment task in which the test-taker is presented with a stimulus monologue or conversation and then is asked to respond to a set of comprehension questions.
The monologues, lectures, and brief conversations used in such tasks and the subsequent multiple-choice questions don't mirror communicative, real-life situations. But with some care and creativity, one can create reasonably authentic stimuli, and in some rare cases the response mode (as shown in

one example below) actually approaches complete authenticity. Here is a typical example of such a task.

### *Dialogue and multiple-choice comprehension items*

*Test-takers hear:*
**Directions:** Now you will hear a conversation between Lynn and her doctor. You will hear the conversation two times. After you hear the conversation the second time, choose the correct answer for questions 11-15 below. Mark your answers on the answer sheet provided.

Doctor: Good morning, Lynn. What's the problem?

Lynn:   Well, you see, I have a terrible headache, my nose is running, and I'm really dizzy.

Doctor: Okay. Anything else?

Lynn:   I've been coughing, I think I have a fever, and my stomach aches.

Doctor: I see. When did this start?

Lynn:   Well, let's see, I went to the Jake last weekend, and after I returned home I started sneezing.

Doctor: Hmm. You must have the flu. You should get lots of rest, drink hot beverages, and stay warm. Do you follow me?

Lynn: Well, uh, yeah, but ... shouldn't I take some
medicine?
Doctor: Sleep and rest are as good as medicine when
you have the flu. Okay, thanks, Dr. Brown.

*Test-takers read:*

11. What is Lynn/s problem?
 (A) She feels horrible.
 (B) She ran too fast at the lake.
 (C) She/s been drinking too many hot beverages.

12. When did Lynn/s problem start?
 (A) When she saw her doctor.
 (B) Before she went to the lake.
 (C) After she came home from the lake.

13. The doctor said that Lynn _____.
 (A) flew to the lake last weekend
 (B) must not get the flu
 (C) probably has the flu

14. The doctor told Lynn _____.
 (A) to rest
 (B) to follow him
 (C) to take some medicine

15. According to Dr. Brown, sleep and rest are
_____ medicine when you
have the flu.
(A) more effective than
(B) as effective as
(C) less effective than

To compensate for the potential inauthenticity of post-stimulus comprehension questions, you might be able to find contexts where questions that probe understanding are more appropriate.
Consider the following situation:

*Dialogue and authentic questions on details*
*Test-takers hear:*
You will hear a conversation between a detective and a man. The tape will play the conversation twice. After you hear the conversation a second time, choose the correct answers on your test sheet.
Detective:  Where were you last night at eleven P.M., the time of the murder?
Man: Uh, let's see, well, I was just starting to see a movie.
Detective:  Did you go alone?
Man: No, uh, well, I was with my friend, uh, Bill. Yeah, I was with Bill.
Detective:  What did you do after that?

Man: We went out to dinner, then I dropped her off at her place.
Detective:  Then you went home?
Man: Yeah.
Detective: When did you get home?
Man: A little before midnight.

*Test-takers read:*

7.  Where was the man at 11:00 P.M.?
    (A) In a restaurant.
    (B) In -a theater.
    (C) At home.

8.  Was he with someone?
    (A) He was alone.
    (B) He was with his wife.
    (C) He was with a friend.

9.  Then what did he do?
    (A) He ate out.
    (B) He made dinner.
    (C) He went home.

10. When did he get home?
    (A) About 11:00.
    (B) Almost 12:00.
    (C) Right after the movie.

11. The man is probably lying because (name two clues):

    1. _____

    2. _____

In this case, test-takers are brought into a little scene in a crime story. The questions following are plausible questions that might be asked to review fact and fiction in the conversation. Question #11, of course, provides an extra shot of reality: the test-taker must name the probable lies told by the man (he referred to Bill as "her"; he saw a movie and ate dinner in the space of one hour), which requires the process of inference.

## Authentic Listening Tasks

Ideally, the language assessment field would have a stockpile of listening test types that are cognitively demanding, communicative, and authentic, not to mention interactive by means of integration with speaking. However, the nature of a test as a *sample* of performance and a set of tasks with limited time frames implies an equally limited capacity to mirror all the real-world contexts of listening performance.

Beyond the rubrics of intensive, responsive, selective, and quasi-extensive communicative contexts described above, can we assess aural comprehension in a truly communicative context? Can we, at this end of the range of listening tasks, ascertain from test-takers that they have processed the main idea( s) of a lecture, the gist of a story, the pragmatics of a conversation, or the unspoken inferential data present in most authentic aural input? Can we assess a test-taker's comprehension of humor, idiom, and metaphor? The answer is a cautious yes, but not without some concessions to practicality. And the answer is a more certain yes if we take the liberty of stretching the concept of assessment to extend beyond tests and into a broader framework of alternatives. Here are some possibilities.

1. *Note-taking.* In the academic world, classroom lectures by professors are common features of a non-native English-user's experience. One form of a midterm examination at the American Language Institute at San Francisco State University. (Kahn, 2002) uses a 15-minute lecture as a stimulus. One among several response formats includes note-taking by the test-takers. These notes are evaluated by the teacher on a 30-point system, as follows:

### *Scoring system for lecture notes*

**0-15 points**

*Visual representation:* Are your notes clear and easy to read? Can you easily find and retrieve information from them? Do you use the space on the paper to visually represent ideas? Do you use indentation, headers, numbers, etc.?

**0-10 points**

*Accuracy:* Do you accurately indicate main ideas from lectures? Do you note important details and supporting information and examples? Do you leave out unimportant information and tangents?

**0-5 points**

*Symbols and abbreviations:* Do you use symbols and abbreviations as much as possible to save time? Do you avoid writing out whole words, and do you avoid writing down every single word the lecturer says? The process of scoring is time consuming (a loss of practicality), and because of the subjectivity of the point system, it lacks some reliability. But the gain is in offering students an authentic task that mirrors exactly what they have been focusing on in the classroom. The notes become an indirect but arguably valid form of assessing global listening comprehension. The task fulfills the criteria of cognitive demand, communicative language, and authenticity.

**2. Editing.** Another authentic task provides both a written and a spoken stimulus, and requires the test-taker to listen for discrepancies. Scoring achieves relatively high reliability as there are usually a small number of specific differences that must be identified. Here is the way the task proceeds.

### Editing a written version of an aural stimulus

*Test-takers read:* the written stimulus material (a news report, an email from a friend, notes from a lecture, or an editorial in a newspaper).

*Test-takers hear:* a spoken version of the stimulus that deviates, in a finite number of facts or opinions, from the original written form.

*Test-takers mark:* the written stimulus by circling any words, phrases, facts, or opinions that show a discrepancy between the two versions.

One potentially interesting set of stimuli for such a task is the description of a political scandal first from a newspaper with a political bias, and then from a radio broadcast from an "alternative" news station. Test-takers are not only forced to listen carefully to differences but are subtly informed about biases in the news.

**3. Interpretive tasks.** One of the intensive listening tasks described above was paraphrasing a- story or conversation. An interpretive task extends the stimulus material to a longer stretch of discourse and forces the test-taker to infer a response. Potential stimuli include
- **song lyrics,**
- **[recited] poetry;**
- **radio/television news reports, and**
- **an oral account of an experience.**

Test-takers are then directed to interpret the stimulus by answering a few questions (in open-ended form). Questions might be:
- "Why was the singer feeling sad?"
- "What events might have led up to the reciting of this poem?"
- "What do you think the political activists might do next, and why?"
- "What do you think the storyteller felt about the mysterious disappearance of her necklace?"

**4. Retelling.** In a related task, test-takers listen to a story or news event and simply retell it, or summarize it, either orally (on an audiotape) or in writing. In so doing, test-takers must identify the gist, main idea, purpose, supporting points, and/or conclusion to show full comprehension. Scoring is partially predetermined by specifying a

minimum number of elements that must appear in the retelling. Again reliability may suffer, and the time and effort needed to read and evaluate the response lowers practicality. Validity, cognitive processing, communicative ability, and authenticity are all well incorporated into the task.

A fifth category of listening comprehension was hinted at earlier in the chapter: **interactive listening**. Because such interaction presupposes a process of *speaking* in concert with listening, the interactive nature of listening will be addressed in the next chapter. Don't forget that a significant proportion of real world listening performance is interactive. With the exception of media input, speeches, lectures, and eavesdropping, many of our listening efforts are directed toward a two-way process of speaking and listening in face-to-face conversations.

<p align="center">***</p>