

# **Interpreting the Presentational Rubric**

## **General guidelines**

**The reader/listener** is assumed to be a sympathetic native speaker who is willing to go the extra mile to try to understand what the writer/speaker is saying. This person is NOT a foreign language teacher who is counting or evaluating the mistakes, lowering the student's grade for every error that is made. She is willing to overlook any error that doesn't leave her wondering what the writer/speaker was trying to say. However, this person does NOT speak any English, so, except for cases where a word is spelled the same (or nearly the same) in both languages, every time she sees or hears an English word inserted into the text it is **completely without meaning** to her - it might as well have been left blank as far as she is concerned, and that is how you should treat English as the grader – as if it were invisible.

Every writing/speaking sample should be **evaluated globally**. This means that even if a response is determined to contain one or more errors that do interfere with comprehensibility, it could still receive a passing score of 2 for *Comprehensibility* if, taken as a whole, a sympathetic reader/listener would feel that **they understood significantly more than they didn't understand**.

Writing/speaking samples should always be evaluated in light of the **performance standards** (see below) which have been established for the **benchmark level** at which the question is rated. (The "L" number in the third slot of the item name indicates the benchmark level; e.g., an item named F-W-L1-002 is a benchmark level one (1) question.)

**A score of "3"** should be reserved for responses which can be considered "above and beyond" the normal expectations for completing a given task **at a given benchmark level**. For instance, simple complete sentences with proper punctuation would be considered "above and beyond" the normal expectations of a benchmark level 1 task, and thus would receive a score of "3" for *Text Type*. However, the same types of sentences would only receive a score of "2" at benchmark level 3, where they would meet the normal expectations of that benchmark level, but would not demonstrate the ability to go "above and beyond" normal expectations.

The **total score** for any writing/speaking sample is determined by computing the average of the individual scores for each category contained in the grading rubric. This means that an item which receives individual scores of 3 for *Text Type*, 1 for *Comprehensibility* and 2 for *Task Completion* would receive a total (average) score of 2.0, which is a passing score. On the other hand, individual scores of 2 for *Text Type*, 1 for *Comprehensibility* and 2 for *Task Completion* would result in a total (average) score of 1.7, which is not a passing score. A score of 0 (zero) in any single category automatically constitutes a failure for the entire item. Averaging is not used in this case.

-----

## Categories of Evaluation

### Text Type:

- \* Evaluate the language that appears on the screen, as it is, regardless of whether or not it meets the demands of the task.
- \* Evaluate what type of text/speech the writer has **demonstrated full control of**. Is she demonstrating a **consistent** ability to construct coherent sentences, or does her writing/speaking consist primarily of words and phrases? Can she organize her sentences into paragraphs, or are they simply strings of disconnected sentences that could be reordered without any real loss of meaning?
- \* *Word Fragments*: word fragments, as such, are rare in European languages. A response like “2 min” where a student is asked to write/express a unit of time, would have to be considered a fragment, since the student should have written or said “deux minutes” in French, or “dos minutos” in Spanish, and has not actually written any complete words in the target language.
- \* *Sentences*: the term *Sentences* (note that it is in the plural) refers to two or more **different** sentences, each containing a subject and a predicate, and each of which is in some way different from the other. For example, the response, “I like pizza. I like Coke. I like ice cream.”, would not qualify as *sentences*, because, even though this sample does technically consist of three distinct clauses, each with a subject and a predicate, it is really only *one* sentence type mechanically repeated three times, so that the text type is actually only a *list of words* presented in a single sentence frame. The problem here is that one formulaic clause repeated over and over again does not give us any indication that the writer has **full control** of the *sentence* text type. It is perfectly acceptable for *sentences* to be formulaic (i.e. canned responses for certain situations), but they should still be *different from one another*. For example, “Good morning. My name is John. I’m pleased to meet you.”, is definitely a memorized, formulaic response that is typically used in replying to a standard greeting, but at least it is not just a repetition of the *same* sentence with only a slight modification each time, as in the previous example. This response would qualify as *sentences*.
- \* *Phrases and Sentences*: If a response is made up of a **mixture** of two or more *sentences* (as defined above), **and** several isolated *phrases* (i.e. incomplete sentences), then it qualifies as *phrases and sentences*. For example, if, when asked about favorite recreational activities, a student responds with, “playing the guitar, Raider football games, I like to go to concerts, I play tennis every Saturday”, this response would qualify as *phrases and sentences*, and would receive a score of 3 at benchmark levels 1 & 2, a score of 2 at benchmark level 3, and a score of 1 benchmark levels 4 & 5.
- \* *Strings of sentences*: The term *strings of sentences/questions* (a score of 2 at benchmark level 4 and a score of 1 at benchmark level 5) refers to a text that is composed entirely (or almost entirely) of complete sentences (as defined above) which are not organized in any particular way and could be **freely reordered without any loss of meaning**.
- \* *Strings of sentences with some paragraph structure*: The term *strings of sentences with some paragraph structure* (a score of 3 at benchmark level 4 and a score of 2 at benchmark level 5) refers to three or more unique and non-formulaic sentences (as defined above), which show evidence of at least some topical or logical organization, or are joined together, even if inconsistently, with appropriate transition words, such as *first, then, however, finally*, etc. **If the sentences in question could be freely reordered without any loss of meaning, then they would not meet this criterion.**
- \* *Complete paragraph structure*: the term *Complete Paragraphs* refers to a structured text consisting completely, or primarily, of complete sentences (as defined above) which are organized into meaningful thematic paragraphs, and which display appropriate indicators of paragraph structure including indentation **and** logical topic shifts.

## Comprehensibility:

\* *Comprehensibility* is NOT primarily about accuracy, in the sense of errors in spelling or grammar, except to the extent that any such errors might render the *meaning* of the text incomprehensible to a sympathetic reader/listener (see above). As long as the meaning is comprehensible, even a response that is riddled with errors can receive a passing score of 2 for *Comprehensibility*.

\* *Comprehensibility* IS primarily about a text being meaningful, logical and plausible as a response to the task at hand. For example, if a student is asked to list five favorite foods, and they list five recreational activities instead, then this is **not** a comprehensible response for this task. In a real-world context, the person asking the question would be no more well-informed about this person's favorite foods than they were before they asked the question. They also wouldn't understand (comprehend) why the person was not answering their question. The response might be written or spoken in perfectly good French, Spanish, German, or Japanese, but it still is **not** a *meaningful* or *reasonable* response to this question.

\* Anything written in English is considered to be invisible. So, when you encounter a response like "trois days", for example, instead of the French "trois jours", you should first consider the English word "days" invisible, and then determine whether the remaining "trois" can be accepted as a *meaningful* (i.e. logical or plausible) response for the task at hand. Clearly, a response like "3 days" does not give **any** indication that the student knows anything at all in the target language, and would therefore have to be considered as having "little or no meaning" and receive a score of "0".

## Task Completion:

\* Evaluate whether or not the question was *fully* answered and/or *all* aspects of the task were completed as per the instructions.

\* Evaluate the degree to which a response meets the demands of the task (i.e., is it "on-task", or is it "off-task"?). If the response seems to meet the requirements in terms of quantity, but does not actually respond to the question which was asked or meet the specific demands of the task, it cannot be considered "complete" because it is not on-task.

\* If a prompt asks the student to write a note or a letter, and then specifies that they should discuss at least three of the topics from a bulleted list of possible topics, then the **minimum** requirement necessary for a passing score of 2 is met if the student discusses any three of the items listed, even if they don't include the normal elements of a letter format. If they do add the appropriate greeting, leave-taking and signature for the type of letter or note they are writing, then this should be considered as having exceeded the minimum requirements, and they should receive a score of 3. Likewise, if they discuss more than three topics from the list (or whatever number is specified in the prompt), they have exceeded the minimum requirements.

---

## ***Benchmark Level Performance Standards***

### **Benchmark Level I (Novice-Low)**

- \* Can reproduce from memory a very limited number of isolated words or familiar phrases, but errors are to be expected.

### **Benchmark Level II (Novice-Mid)**

- \* Able to copy or transcribe familiar words or phrases, and reproduce from memory a modest number of isolated words and phrases in context.
- \* Can supply limited information on simple forms and documents, and other basic biographical information, such as names, numbers, and nationality.
- \* Exhibit a high degree of accuracy when writing on well-practiced, familiar topics using limited formulaic language.
- \* Exhibit a marked decrease in accuracy with less familiar topics. Errors in spelling or in the representation of symbols may be frequent.

### **Benchmark Level III (Novice-High)**

- \* Able to meet limited basic practical writing needs using lists, short messages, postcards, and simple notes, relying mainly on practiced material.
- \* Able to recombine learned vocabulary and structures to create simple non-formulaic sentences on very familiar topics, but may only partially communicate what is intended.
- \* There may be basic errors in grammar, word choice, punctuation, and spelling.

### **Benchmark Level IV (Intermediate-Low)**

- \* Able to formulate non-formulaic sentences based on familiar material. Most sentences are recombinations of learned vocabulary and structures.
- \* Sentences are short and simple, written mostly in present time with occasional and often incorrect use of past or future time.
- \* Vocabulary is limited to common objects and routine activities, and topics are limited to highly predictable content areas and personal information.
- \* There may be basic errors in grammar, word choice, punctuation, and spelling.

### **Benchmark Level V (Intermediate-Mid)**

- \* Able to write short, simple communications, compositions, descriptions, and requests for information in loosely connected texts that are based on personal preferences, daily routines, common events, and other topics related to personal experiences and immediate surroundings.
- \* Most writing is framed in present time, with inconsistent references to other time frames.
- \* Shows evidence of control of syntax in non-complex sentences and in basic verb forms.
- \* This writing is best defined as a collection of discrete sentences and/or questions loosely strung together; there is little evidence of deliberate organization.