

This document contains some important explanations and facts regarding the structure of the tasks in this version of the EFCAMDAT.

In the case of conflicts with information listed elsewhere (e.g. in the journal manuscript outlining this version of the EFCAMDAT), the data in this document supersedes everything else. Note that an early version of the EFCAMDAT Cleaned Subcorpus listed the two sub-corpora as having an equal number of tasks per level (8).

- **The first corpus (the “main prompts”) originally contained 8 tasks per EFCAMDAT proficiency level.** This means that when a learner completed a proficiency level using this set of prompts, they completed 8 lessons and 8 corresponding tasks.
- **The second corpus (the “alternative prompts”) originally contained 6 tasks per EFCAMDAT proficiency level.** This means that when a learner completed a proficiency level using this set of prompts, they completed 6 lessons and 6 corresponding tasks.
- **The original numbering of tasks in the second corpus is such that it *generally* contains the first 6 tasks per proficiency level.** For example, when it comes to the second EFCAMDAT proficiency level, tasks #9-#14 are included in the second corpus, while tasks #9-#16 are included in the first corpus. Accordingly, while the tasks in both corpora are on a scale of 1-120 (or 1-118 in the second corpus), in practice, the numbering of the tasks in the second corpus should be on a scale of 1-90.
 - **Exceptions to the above are the following:**
 - Task #51 (third task in level 7) is missing from the second corpus; all the texts in this task were placed in the first corpus based on the classification scheme, even though a manual examination suggests that different prompts were used for this task.
 - In level 1, task #1 is missing from the second corpus instead of task #7.
- **Learners generally only completed tasks from a single corpus, or they started in one corpus and then transitioned to another (usually from the first to the second).** In some cases, there is overlap between the corpora in terms of which tasks a learner completed, but there is variability in terms of which tasks they completed from which corpus.