

# Coherence ratings

Manual coherence ratings were obtained for the texts in the ITACA corpus and recorded in `ratings.csv` (see metadata). In the following we describe the individual items of the coherence rating.

Note that `ratings.csv` can contain various ratings per text. To calculate inter-rater agreement and ensure rating quality, most texts received two ratings. However, some of the texts were rated by only one rater, because one of the raters decided that those texts were too short to be rated. Ten of the texts were rated in a pilot by 6 raters. Finally, one of the texts (CA31TO\_LAR14V) has 4 ratings, because it was one of three (next to ST23TO\_PER21M and GR25ME\_NAR22A) that has been used for the training of the raters.

## Questionnaire

Questions were administered with the following request:

"Indica, per favore, il tuo grado di accordo o disaccordo con ciascuna affermazione."

The items to rate are recoded as:

|                   |                                                                                                                        |
|-------------------|------------------------------------------------------------------------------------------------------------------------|
| <b>likert_001</b> | Il testo è di tipo narrativo o espositivo. [REVERSED ITEM]                                                             |
| <b>likert_002</b> | Il testo rispecchia il genere testuale previsto dalla consegna (lettera).                                              |
| <b>likert_003</b> | Il registro è generalmente adeguato al tipo di task assegnato (lettera formale a un rappresentante delle istituzioni). |
| <b>likert_004</b> | La trattazione del tema è aderente al task assegnato (ovvero non c'è confusione tra DAD e DDI).                        |
| <b>likert_007</b> | Gli incisi sono di norma correttamente segnalati.                                                                      |
| <b>likert_008</b> | Nel testo sono presenti molte virgole splice. [REVERSED ITEM]                                                          |
| <b>likert_005</b> | All'interno del paragrafo la punteggiatura è scarsamente segnalata. [REVERSED ITEM]                                    |
| <b>likert_006</b> | L'uso del punto nel testo è sempre funzionale alla suddivisione degli enunciati.                                       |
| <b>likert_009</b> | Il testo è suddiviso in modo incoerente o non è affatto suddiviso in sezioni. [REVERSED ITEM]                          |
| <b>likert_010</b> | Il testo è suddiviso in paragrafi coerenti con i cambi di argomento (movimenti testuali).                              |
| <b>likert_011</b> | La segmentazione è funzionale all'articolazione del testo.                                                             |
| <b>likert_012</b> | Le devianze nella punteggiatura sono tali da compromettere la coerenza e/o la comprensione del testo. [REVERSED ITEM]  |
| <b>likert_013</b> | Nel testo vengono rispettate le regole di interpunzione.                                                               |
| <b>likert_014</b> | Di norma i connettivi sono impiegati coerentemente con la loro semantica.                                              |
| <b>likert_015</b> | I connettivi sono spesso impiegati in modo sintatticamente improprio. [REVERSED ITEM]                                  |

|                   |                                                                                                                           |
|-------------------|---------------------------------------------------------------------------------------------------------------------------|
| <b>likert_016</b> | Le devianze nell'uso dei connettivi sono rilevanti. [REVERSED ITEM]                                                       |
| <b>likert_017</b> | Il collegamento tra la ripresa e il suo referente è sempre ricostruibile.                                                 |
| <b>likert_018</b> | I riferimenti (incapsulatori) a contenuti già menzionati nel testo a volte sono difficili da ricostruire. [REVERSED ITEM] |
| <b>likert_019</b> | Le parti del testo sono per lo più giustapposte. [REVERSED ITEM]                                                          |
| <b>likert_020</b> | L'uso dei connettivi e delle anafore è funzionale alla coesione del testo.                                                |
| <b>likert_021</b> | Nel testo è riconoscibile un'idea di fondo.                                                                               |
| <b>likert_022</b> | Il testo è strutturato sulla base dell'idea di fondo.                                                                     |
| <b>likert_023</b> | La tesi è riconoscibile.                                                                                                  |
| <b>likert_024</b> | La tesi è sostenuta da argomenti.                                                                                         |
| <b>likert_025</b> | Non tutte le argomentazioni sono sviluppate. [REVERSED ITEM]                                                              |
| <b>likert_026</b> | Le conclusioni derivano dagli argomenti presenti nel testo.                                                               |
| <b>likert_027</b> | Le idee presenti nel testo sono ben collegate tra loro.                                                                   |
| <b>likert_028</b> | Il testo è costruito a elenco, ovvero è privo di una gerarchia riconoscibile. [REVERSED ITEM]                             |
| <b>likert_029</b> | Sono presenti delle devianze nella sequenza logica, tematica o temporale degli argomenti. [REVERSED ITEM]                 |
| <b>likert_030</b> | Gli argomenti introdotti nel testo sono ridondanti. [REVERSED ITEM]                                                       |
| <b>likert_031</b> | Il testo è stato progettato in modo efficace.                                                                             |
| <b>likert_032</b> | Il contenuto è contraddittorio/incoerente. [REVERSED ITEM]                                                                |
| <b>likert_033</b> | Il contenuto è convincente.                                                                                               |
| <b>likert_034</b> | Gli argomenti introdotti nel testo sono funzionali allo scopo.                                                            |
| <b>likert_035</b> | Gli argomenti introdotti nel testo vanno fuori tema (off-topic). [REVERSED ITEM]                                          |
| <b>likert_036</b> | Gli argomenti introdotti nel testo sono generici. [REVERSED ITEM]                                                         |
| <b>likert_037</b> | I termini utilizzati provocano fraintendimenti. [REVERSED ITEM]                                                           |
| <b>likert_038</b> | I termini utilizzati sono coerenti con la loro semantica e il contesto in cui appaiono.                                   |
| <b>likert_039</b> | Nel testo ci sono informazioni implicite che rendono difficoltosa la comprensione. [REVERSED ITEM]                        |
| <b>likert_040</b> | Tutte le informazioni necessarie alla piena comprensione del testo sono ben esplicitate.                                  |
| <b>likert_041</b> | Il lettore deve sforzarsi per comprendere quanto scritto. [REVERSED ITEM]                                                 |
| <b>likert_042</b> | Sono necessarie diverse riletture per comprendere il testo o alcune sue parti. [REVERSED ITEM]                            |
| <b>likert_043</b> | Il testo si comprende facilmente.                                                                                         |
| <b>likert_044</b> | Lo scopo/gli scopi del testo è/sono chiaro/i.                                                                             |

Ratings were done on a Likert scale recoded as:

```
{  
  "3": "Del tutto d'accordo",  
  "2": "Un po' d'accordo",  
  "1": "Un po' in disaccordo",  
  "0": "Del tutto in disaccordo"  
}
```

The rating scale for reversed items is recoded as:

```
{  
  "0": "Del tutto d'accordo",  
  "1": "Un po' d'accordo",  
  "2": "Un po' in disaccordo",  
  "3": "Del tutto in disaccordo"  
}
```