**Deliverable 7.1 Written specification of interface capabilities**

This specification will list desired functionalities for the user interface, in particular it will provide:

* a detailed specification of search options
* types and formats of search results
* differences in demands by target group
* form and technology-related user requirements

*This specification document aims at presenting the desired functionalities as an “ideal case” scenario. It is meant to provide a comprehensive starting point that will be narrowed down to a concrete plan for a feasible implementation of the interface.*

CONTEXT INFORMATION:

**Users  
-** we distinguish four broad types of users: teachers, trainers, testers, linguists

**Working context for using MERLIN data**- using texts for illustrating competence levels (CEFR) (to learners, in the classroom)  
- using texts for preparing teaching materials, defining teaching targets  
- using texts for defining assessment criteria  
  
- teachers are working on these tasks in a team  
- many users work with more than one of the MERLIN languages in parallel

REQUIREMENTS SPECIFICATION:

**Relevant data types:**

- unannotated learner productions (without annotations and ratings added)  
- annotated learner productions (including original, corrections AND decision criteria/rater comments)  
- complete target hypothesis (ZH1) of learner productions  
- test tasks (also together with learner texts)  
- rating scales/criteria (also together with learner texts)  
- groups of texts (according to metadata filters)  
- prototypical learner productions per level  
- plain feature lists (without texts)

**Relevant information on the data:**

- features of learner language

* vocabulary and grammar (++)
* text characteristics (+)
* sociolinguistic criteria
* vocabulary frequencies

- metadata on texts

* author related information (L1, age, gender)
* level of test task (++)
* type of test task (++) (text type)
* ratings
* learning setting of the learner (e.g. classroom, private studying, personal trainer, etc.) before taking the test (‘learning context (+)
* aim of task
* allowed time for completion
* text typology (e.g. genre)
* text content

- information derived from texts

* size of vocabulary (++)
* sentence complexity (++)
* variety and frequency of learner language features (+)
* listing of learner language features and their values for sampled subgroups (e.g. by CEFR level, by L1 of learner, etc., also mix of L1, to counter “over-sensitivity” of rater), for example vocabulary range

**Different access modes to MERLIN data**

- searching:   
---by features/words/lemmas/RegEx  
---for texts with features of different CEFR levels (e.g. voc B2, but cohesion A2)

- browsing:   
---by metadata information  
---by document type (i.e. learner production vs. test task vs. rating scale/CEFR documentation)

- filtering:   
---by metadata information  
---by features

**Display modes**:

- entire text or wider text section (for teachers also separate sentences), - for both modes with the possibility to have a parallel view on the original learner text and the corresponding target hypothesis   
- learner language features inline in text (e.g. highlighting)

**General functions to manage data:**

- attaching notes and comments  
- storing information (e.g. bookmarking, history)   
- ‘team area’, sharing information among users  
- flexible data export (e.g. single texts, groups of texts, single paragraphs/sections)  
- feedback function for users to comment on annotations/rating decisions  
- discussion area (also with MERLIN experts)  
- contextualized help  
- support for regular expressions searches  
- support for XML data handling

**Target group specific preferences**

* Teachers are interested in vocabulary and grammar
* Trainers/testers are interested in text characteristics, formal conventions; feature lists and test tasks
* Linguists are interested in vocabulary frequencies and sociolinguistic aspects; lists of features and metadata
* Importance of metadata and frequency information for: linguists > trainers/testers > teachers
* Teachers require bookmarking and collaborative working
* Linguists, trainers and testers are more in need of exporting than teachers

**Platform structure**

* Keep functionalities for aggregation/quantitative analysis and text selection/text search separate
* Present search functionalities in non-corpus-linguistics style, but task oriented (e.g. sentence complexity instead of sentence length in words)

**Technical requirements**

* Should not require to install plugins etc. (> 65% of users require support for installations)
* 80% Microsoft Windows environments
* Browser support for Internet Explorer, Firefox, Chrome
* Difficulties with .xml files, also .txt and .odt
* Difficulties with reading/understanding XML data

|  |  |
| --- | --- |
| EU_flag_LLP_EN-01 | This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein. |

Please cite as: MERLIN project, Written specification of interface capabilities, 2014,<http://merlin-platform.eu>