



merlin - a multifunctional trilingual learner corpus related to the CEFR

Katrin Wisniewski, Technical University Dresden



Education and Culture DG

Lifelong Learning Programme



1. The lack of empirical illustration of CEFR levels

- overwhelming focus on CEFR **levels** expresses limited understanding of the document
- still, evidence for how levels relate to learner language is urgently needed in language learning, teaching, and testing

(Fulcher 2004, Hulstijn 2007, North 2000, Wisniewski et al. 2013, Wisniewski forthcoming)

2. The Reference Level Descriptions (RLD)

- Council of Europe initiative: *RLD for National and Regional Languages* (CoE, 2005)
- Recently, some RLD adopt a **bottom-up approach** (Green 2013) → based on learner corpora
 - “... electronic collections of texts produced by foreign or second language learners in a variety of language settings“ (Granger et al. 2002: VII),
 - often annotated with the help of a standardized system of error tags (Díaz-Negrillo and Domínguez 2006)
- Examples:
 - *English Profile Project* (Saville/Hawkey 2010); *Norsk Profil* (Carlsen 2013); *Profilo della lingua italiana* (Spinelli/Parizzi 2010)

3. The MERLIN project

MERLIN: „Multilingual Platform for the European Reference Levels: Interlanguage Exploration in Context”: www.merlin-platform.eu

European Union grant (LLP 518989-LLP-1-2011-1-DE-KA2-KA2MP)

01/2012 – 12/2014

Partners: Technische Universität Dresden (DE) (Lead Partner), European Academy Bolzano (IT), Charles University (CZ), telc GmbH (DE), Berufsförderungsinstitut Oberösterreich (AT), Eberhard-Karls-Universität Tübingen (DE), European Center of Modern Languages - Council of Europe (AT) (associated partner)

Main aims:

- develop a freely accessible online platform to **illustrate** the CEFR levels for **German, Italian and Czech**
- contribute to the **validation** of selected CEFR scales

(cf. Wisniewski et al. 2013)

3.1 The MERLIN project: data

- written productions from standardised language tests (telc, Frankfurt & UJOP, Prague) with strict quality controls
- ~200 texts per (available) CEFR test level (N=2470)
- metadata (L1, age, gender ...)
- ~280.000 tokens

3.2 Re-ratings & transcriptions in MERLIN

- **re-ratings** of written parts of the original tests with ...
 - assessor-oriented, CEFR-related analytical rating grid (cf. Table 3, CEFR); rating criteria based on chapter 5 categorization, i.e. grammatical accuracy, vocabulary range & control, sociolinguistic appropriateness, coherence & cohesion
 - assessor-oriented holistic rating scale („General linguistic range“)
- **quality control**: 10% double-rated; CTT & Multi-facet Rasch analyses; problematic Italian ratings → all re-re-rated → fair averages
- **transcription/digitisation** (xml mind©) → guideline-based, reliability checks

3.3 Developing the annotation scheme

Sources of tags in the annotation scheme

1. CEFR (scale operationalization)
2. SLA research
3. users (needs analysis; questionnaires & interviews; textbook analyses; language test analyses)
4. text analyses

(cf. e.g. Sieber 2012, Yang/Sun 2012, Vajjala & Meurers 2012, Paquot/Granger 2012, Bulté/Housen 2012, Granger & Bestgen 2011, Lu 2010, 2011, Mellor 2011, Carlsen 2010, Housen & Kuiken 2009, Bardovi-Harlig 2009, Malvern et al. 2008, Rimrott & Heift 2008, Burger 2007, Stede 2007, Nesselhauf 2005, Read/Nation 2004, Schmitt/McCarthy 2004, Dewaele 2004, Daller/Van Hout/Treffers-Daller 2003, Ortega 2003, Wray 2002, Santipolo 2002, Bachmann 2002, Nation 2001, 2007, Read 2000, Wolfe-Quintero, Inagaki & Kim 1998, Laufer/Nation 1995, O'Loughlin 1995, Halliday/Hasan 1989, Arnaud 1984, Wisniewski forthcoming, 2013)

3.4 Structure of annotations in MERLIN

- 1) **target hypotheses** (cooperation with Falko, e.g. Reznizeck/Walter et al. 2010)
 - 2) **annotation of errors and other learner language features**
- **manual** and **automatic** annotations
 - **pilotings** and **reliability checks**, extensive **documentation**
 - **layered design** (orthographical and grammatical features for whole corpus, vocabulary, sociolinguistic appropriateness, pragmatics for A2 & B2)

3.5 Envisaged features of the MERLIN platform

- free online access to all resources, open-source tools
- full texts (with & without annotations), metadata, and tasks
- detailed task descriptions (ALTE Grid for writing)
- CEFR profiles (grammatical accuracy , vocabulary range/control, coherence/cohesion, sociolinguistic appr., general linguistic range)
- simple & advanced search options
- sort texts: metadata – annotations – task level – CEFR level ...
- create & export word lists
- display statistical measures (e.g. of lexical variety, grammatical complexity)
- cross-language learner language analyses

4. Applications of the MERLIN platform

- second language acquisition research (corpus-based studies often use intuitive distinctions of more or less competent learners)
- natural language processing research for L2
- language learning
- language teaching
- language testing
- ...

4.1 A closer look: MERLIN in Language Testing

Generally...

- Charles Alderson (1996) identified use of L2 corpora \leftrightarrow still underexploited (Ball 2001, Barker 2004, Flowerdew 2012)
- learner corpora contribute to transparency, consistency, comparability of testing \leftrightarrow high costs \rightarrow L2 corpora often not freely available (e.g. Cambridge ESOL, Barker 2010)

Merlin useful for...

- relating language tests to the CEFR
- CEFR-related item writing, development of assessment materials
- empirically-based rating scale construction (Hawkey/Barker 2004, Fulcher et al. 2011)
- CEFR rating scale validation (Wisniewski 2013, forthcoming)

5. Conclusion

- availability
- needs orientation
- learner-centred approach
- multiple uses
- MERLIN bases CEFR use on empirical data → helps to control reification of CEFR scale system (Fulcher/Davidson 2007: 232)
- but corpus is small – the work must go on!

Thank you for your attention!

Katrin.Wisniewski@tu-dresden.de