# merlin

The CEFR level descriptions and empirical learner data:
MERLIN, a multi-lingual corpus initiative

Katrin Wisniewski, Technical University Dresden

Andrea Abel, European Academy Bozen/Bolzano

# Outline

1) The need for illustration of the CEFR levels

2) The Reference Level Descriptions initiative

3) The MERLIN project

4) Applications of MERLIN in Language Teaching

5) Applications of MERLIN in Language Testing

6) Discussion

7) Conclusion/Outlook

Katrin Wisniewski & Andrea Abel

# 1) The need for illustration of the CEFR levels

- in spite of wide distribution of the CEFR, many practitioners do not know what the levels actually mean

- unclear/vague wording, lack of consistence in CEFR, lack of homogeneity

- CEFR text ←→ CEFR scales

- more empirical evidence needed, without analysis of learner language

- applicability across many languages requires general level descriptions ←→ applicability for single languages enhanced by language-specific illustrations

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

# 2) The Reference Level Descriptions (RLD) initiative

- Council of Europe initative: *RLD for National and Regional Languages* (CoE, 2005)
    - level specifications (→ tradition of *Threshold Level*, 1975)
    - purpose: anchor for the development of language programmes
- mostly built on expert intuitions, **top-down approach**, e.g.: *Profile Deutsch* (Glaboniat et al. 2010), A1-C2
- **bottom-up approach** (Green 2013) i.e. RLD based on learner language that has been compiled to a corpus

# 2) The Reference Level Descriptions (RLD) initiative

Learner corpora

"Computer learner corpora are electronic collections of …. texts produced by foreign or second language learners in a variety of language settings" (Granger et al. 2002:  VII), "systematic computerized collections of texts produced by language learners" (Nesselhauf, 2005)

usually annotated with the help of a standardized system of error tags (Díaz-Negrillo and Domínguez, 2006)

tend to provide meta-information, such as the authors' L1, age, gender, etc. and other valuable information from all relevant levels of linguistic description

usually a combination of automatic linguistic annotation (e.g. part of speech) & manual annotation (e.g. errors or other features)

Katrin Wisniewski & Andrea Abel

# 2) The Reference Level Descriptions (RLD) initiative

examples of RLD based on learner corpora:

- *Norsk Profil* (Carlsen 2013), ASK learner corpus, B1-B2, restricted access

- *English Profile Project* (Saville/Hawkey 2010*), Cambridge English Profile Corpus,* aims at 10 million words, 20% spoken, 80% written data, A1-C2; English Grammar Profile, English Vocabulary Profile, English Functions Profile; corpus itself is not publicly available

- *Profilo della lingua italiana* (Spinelli/Parizzi 2010), learner corpus (CELI  test) A2-B2: 100, corpus itself not publicly available (CD-Rom)

# 3) The MERLIN project

MERLIN: „**M**ultilingual Platform for the **E**uropean **R**eference **L**evels: **In**terlanguage 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)

Katrin Wisniewski & Andrea Abel

# 3) The MERLIN project: data

- written productions from standardised language tests with strict quality controls

- ~200 texts per (available) CEFR level (N=2470)

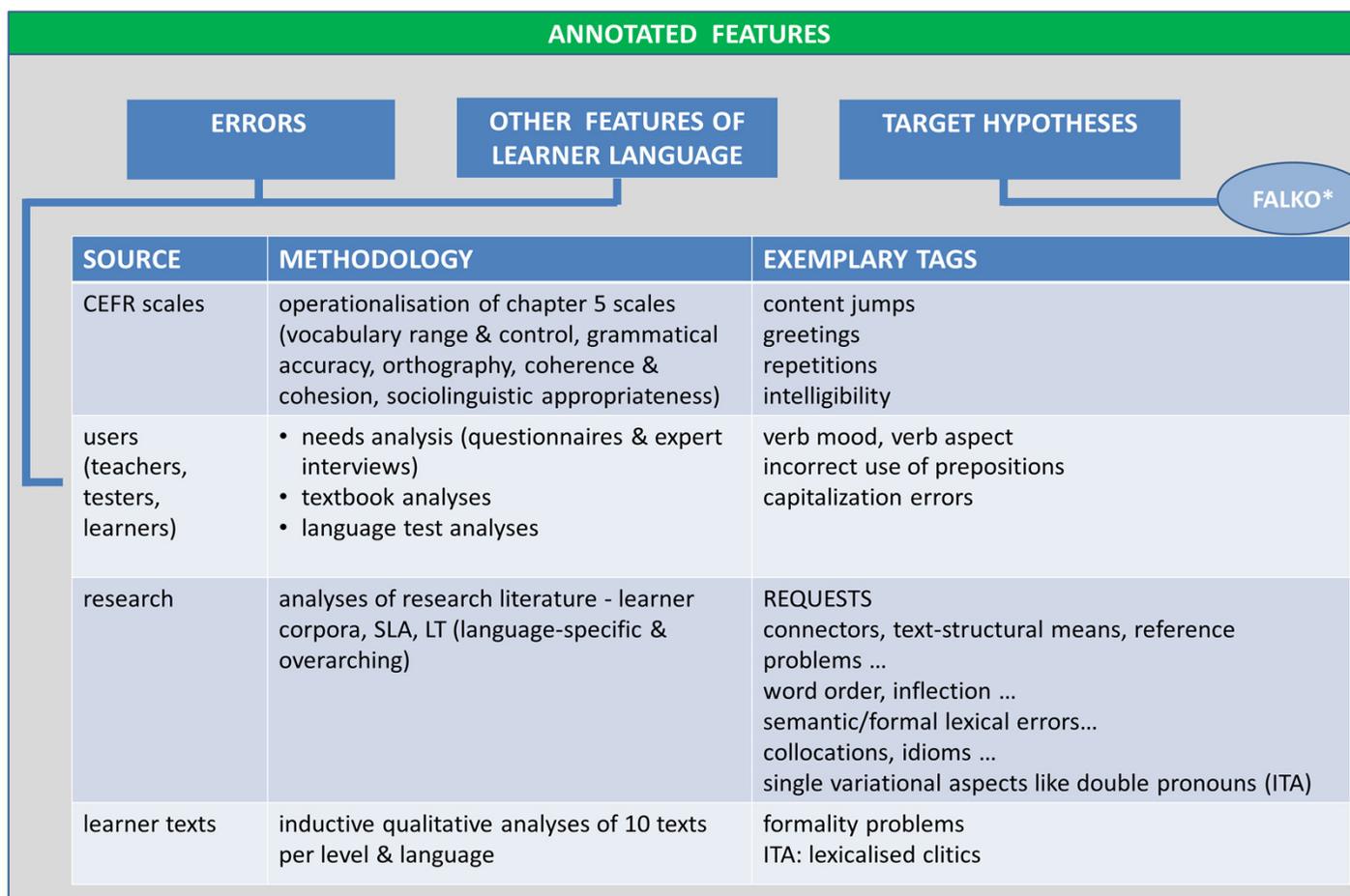- metadata (L1, age, gender ...)

- ~280.000  tokens

|  | German | Italian | Czech | |
|---|---|---|---|---|
| A1 | 229 | 229 | | |
| A2 | 228 | 224 | 131 | |
| B1 | 231 | 223 | 171 | |
| B2 | 225 | 222 | 131 | |
| C1 | 226 | | | |
| **TOTAL** | **1139** | **898** | **433** | **2470** |

Katrin Wisniewski & Andrea Abel

# 3) The MERLIN project: workflow

1) rater training, re-ratings with CEFR-related analytical rating grid (cf. Table 3, CEFR), 10% double-rated → reliability checks (MFR analyses/CTT)

2) transcription/digitisation (xml mind©) → reliability checks

3) annotation: manual & automatic (ongoing) →reliability checks

4) corpus exploration (queries, visualisation) & corpus statistics (complex indicators)

# 3) The MERLIN project: annotations

| ANNOTATED FEATURES | | |
|---|---|---|

**ERRORS** — **OTHER FEATURES OF LEARNER LANGUAGE** — **TARGET HYPOTHESES** — FALKO*

| SOURCE | METHODOLOGY | EXEMPLARY TAGS |
|---|---|---|
| CEFR scales | operationalisation of chapter 5 scales (vocabulary range & control, grammatical accuracy, orthography, coherence & cohesion, sociolinguistic appropriateness) | content jumps greetings repetitions intelligibility |
| users (teachers, testers, learners) | • needs analysis (questionnaires & expert interviews) • textbook analyses • language test analyses | verb mood, verb aspect incorrect use of prepositions capitalization errors |
| research | analyses of research literature - learner corpora, SLA, LT (language-specific & overarching) | REQUESTS connectors, text-structural means, reference problems … word order, inflection … semantic/formal lexical errors… collocations, idioms … single variational aspects like double pronouns (ITA) |
| learner texts | inductive qualitative analyses of 10 texts per level & language | formality problems ITA: lexicalised clitics |

Katrin Wisniewski & Andrea Abel

# 3) The MERLIN project: What the platform will offer

- 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 profile for each text (grammatical accuracy , vocabulary range & control, coherence/cohesion, sociolinguistic appropriateness, general linguistic range)

- simple & advanced search options

- sort texts: metadata (e.g. L1, age) – annotations – task level – CEFR level …

- create & export word lists

- display statistical measures (e.g. of lexical variety, grammatical complexity)

- cross-language learner language analyses

- …

Katrin Wisniewski & Andrea Abel

# 3) The MERLIN project: applications

- validation of CEFR scales

- second language acquisition research

- natural language processing research

- **language teaching**

- **language testing**

Katrin Wisniewski & Andrea Abel

# 4) Applications of corpora in language teaching

- bottom-up approach to language teaching (McEnery/Xiao 2011)

- usefulness of corpora in language teaching widely acknowledged since Johns (1988), growing meaning ← → teaching itself still largely unaffected (Römer 2006, Chambers et al. 2011)

- learner corpora are especially underrepresented in language teaching (Flowerdew 2012, Granger 2009)

- large native speaker (NS) corpora offer information about real-life frequency:  frequency ≠ pedagogical relevance, difficulty (Flowerdew 2012, Meunier 2002) → usefulness of learner corpora

# 4) Applications of corpora in language teaching
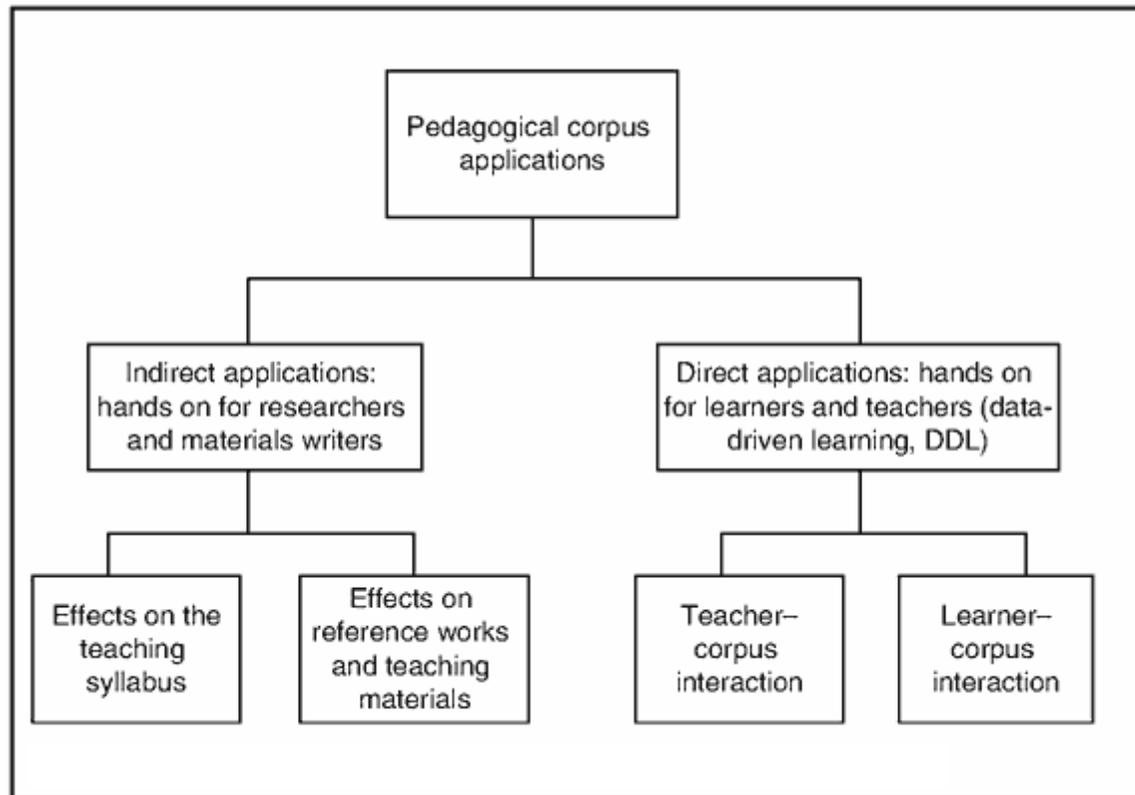


(fig.:  Römer 2010: 19, cf. Aston, 2000, Leech 1997; Nesselhauf 2004, Flowerdew 2009, 2012)

Katrin Wisniewski & Andrea Abel

# 4) MERLIN & syllabus development

- **aim**: help with decisions about selection, presentation, progression of items  (Aston 2000, Granger 2002, Römer 2008)

- **MERLIN** …

    - background: almost all language syllabi are related to the CEFR

    - MERLIN offers empirical data about typical & relevant milestones/errors in learner language with reference to CEFR levels

    - thus helps to ground expectations about learner achievements on learner language that has been reliably rated according to CEFR levels

Katrin Wisniewski & Andrea Abel

# 4) MERLIN & teaching materials development

- corpora can enhance **authenticity** of teaching materials (Römer 2008)
- **MERLIN …**
  - provides information on learners' mastery & variety of use of a multitude of language features on different CEFR levels
  - contributes CEFR-related information to (frequency) (learner) dictionaries, reference works, grammars; e.g., can inform usage notes in learner dictionaries (Granger 2002)
  - provides detailed, CEFR-related information about crucial aspects of language learning, such as learners' use of formulaic sequences/collocations (Wray/Perkins 2000, Pawley/Syder 1983)

# 4) Direct applications of corpora in language teaching

- autonomous, learner-centred learning; Johns 1997: „Every student a Sherlock Holmes"

- exploit MERLIN to teach (Leech 1997)

- ‚data-driven learning' (Johns 1991), discovery learning' (e.g. Aston 2001)

- 'three I's' : Illustration – Interaction – Induction (McEnery/Xiao 2011)

- **direct** NNS corpus applications are **rare**, mostly carried through on institutional level (Flowerdew 2012)

Katrin Wisniewski & Andrea Abel

# 4) Direct applications of MERLIN in language teaching

- possible activities include:

    - peer-group error analyses with relation to the CEFR (Mukherjee/ Rohrbach 2006); self-assessment for learners

    - find out if L1, task type, age… influences usage by sorting texts

    - compare interlanguages (Czech/Italian/ German)

    - find out how learners rated higher/lower master specific language phenomena

    - compare MERLIN and native speaker corpus (Granger 2002: CIA, Contrastive Interlanguage Analysis)

# 5) Applications of MERLIN in Language Testing

- Alderson (1996) first identified potential use ←→still underexploited (Ball 2001, Barker 2004, Flowerdew 2012), but more common than in Language Teaching

- learner corpora contribute to transparency, consistency, comparability of testing ← → compiling learner corpora is very expensive; as a rule, corpora are not freely available (e.g. Cambridge ESOL, Barker 2010)

# 5) Applications of MERLIN in Language Testing

- archiving function

- comparison of test forms (Barker 2006), test revisions

- relate language tests to the CEFR

- CEFR-related item writing, developing assessment materials

- empirically-based rating scale construction (Hawkey/Barker 2004, Fulcher et al. 2011)

- CEFR rating scale validation (Wisniewski 2013, forthcoming)

- future developments include automated essay evaluation; detection of cheating, automatic annotation (NLP) of learner language

Katrin Wisniewski & Andrea Abel

# 6) Discussion

- **generalizability** → MERLIN can only give an **orientation**, not representative

- only a few, mainly qualitative studies about **effectiveness** of corpus use on learning process, generally positive results (Chambers et al. 2011)

- tendency for learner corpora to serve only as **negative evidence** (Leech 1997) in teaching

# 6) Discussion

- **access** → MERLIN is freely **available**

- **needs orientation,** pedagogic mediation at the level of corpus design and  corpus exploitation (Braun 2005) → MERLIN: user needs study regarding annotations & interface design/functionalities

- **corpus technology** as possible obstacle for learners & teachers (Chambers et al., Gilquin et al. 2007) → make benefits clear; graded approach (Chambers et al. 2011) → simple/advanced search options in MERLIN

Katrin Wisniewski & Andrea Abel

# 7) Conclusion/Outlook

- MERLIN bases CEFR use on empirical data, learner-centred approach

- helps to control reification of CEFR scale system (Fulcher/Davidson 2007: 232)

- pilot project design; platform can be extended (→ empirically based „criterial features" for German, Italian, Czech?)

- "(…) corpora will not only revolutionize the teaching of subjects such as grammar (…), they will also fundamentally change the ways we approach language education, including both what is taught and how it is taught." (McEnery/Xiao 2012)

- further uses (e.g., second language acquisition research, natural language processing)

# Thank you for your attention!

Katrin Wisniewski ([katrin.wisniewski@tu-dresden.de](mailto:katrin.wisniewski@tu-dresden.de))

Andrea Abel ([andrea.abel@eurac.edu](mailto:andrea.abel@eurac.edu))

Education and Culture DG

Lifelong Learning Programme

**TECHNISCHE UNIVERSITÄT DRESDEN**

**EURAC** research

# References

Alderson, J.C. (1996). Do corpora have a role in language assessment? In J.A. Thomas & M.H.Short (Eds.), *Using corpora for language research* (pp. 248-59). London: Longman.

ALTE Grid for writing tasks, http://www.coe.int/t/dg4/linguistic/manuel1_en.asp

Aston, G. (2000). Corpora and language teaching. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective: Papers from the third International Conference on Teaching and Language Corpora* (pp. 7-17). Hamburg: Peter Lang.

Ball, F. (2001). Using corpora in language testing. *Research Notes* 6, 6-8, http://www.cambridgeesol.org/rs_notes/rs_nts6.pdf

Barker, F. (2004). Using corpora in language testing. *Modern English Teacher,* 13(2), 63-67.

Barker, F. 2010. 'How can corpora be used in language testing?' in *The Routledge Handbook of Corpus Linguistics*. Abingdon, UK: Routledge, pp. 633-645
   Book editor(s): O'Keeffe, A. and M. McCarthy.

Braun, S. (2005): From pedagogically relevant corpora to authentic language learning contents. In: *ReCALL 17* (1): 47–64

Carlsen, C. 2012. Proficiency level - a fuzzy variable in computer learner Corpora. *Applied Linguistics* 33 (2) 161-183

Chambers, A./Farr, F./O'Riordan, S. (2011): Language teachers with corpora in mind: from starting steps to walking tall. In: *Language Learning Journal* 39 (1) 85-104.

Council of Europe (ed..) (2005): *Reference Level Descriptions  for National and Regional Languages (RLD). Guide for the production of RLD: Version 2*. https://www.coe.int/t/dg4/linguistic/dnr_EN.asp

Egbert, Paulus and Nakamichi 2002: 111

Flowerdew, L. (2012): Corpora and Language Education. Palgrave Macmillan**.**

Flowerdew, L. (2009): Applying corpus linguistics to pedagogy. A critical evaluation. In: International Journal of Corpus Linguistics 14/3:393-417. (download)

Fulcher, G./Davidson, F./Kemp, J. (2011): Effective rating scale development for speaking tests: Performance decision trees. In: *Language Testing* 28 (5) 5-29.

Katrin Wisniewski & Andrea Abel

Gilquin, G./ Granger, S./Paquot, M. (2007): *Learner corpora: The missing link in EAP pedagogy*. In: Journal of English for Academic Purposes 6 (4) 319-335.

Granger, Sylviane (2002), A Bird's-eye View of Learner Corpus Research. In: Granger/Hung/Petch-Tyson 2002, 333.

Granger, S. (2009)*. The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation*. In: Aijmer, K., *Corpora and Language Teaching*, Benjamins: Amsterdam and Philadelphia, 13-32.

Granger, S., Hung, J. & Petch-Tyson, S. (Eds.) (2002). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam: John Benjamins Green, T. (o.J.): The English Profile Network. Report on EPN Methods and their implications for Reference Level Descriptions. http://www.englishprofile.org/index.php/resources/recommendations-report.

Green, T. : English Profile Network Recommendation Report. http://www.englishprofile.org/index.php/resources/recommendations-report

Hawkey, R./Barker, F. (2004): Developing a Common Scale for the Assessment of Writing. In: Assessing Writing 9, 122-159.

Johns, T. (1988):  Whence and whither classroom concordancing? In: Bongaarts,  T./de Haan, P./Lobbe, S./Wekker, H. (eds.): Computer Applications in Language Learning.  Dordrecht: Foris, 9-33.

Johns, T. (1997): Contexts: The Background, Development and Trialling of a Concordance-based CALL Program. In: Wichmann, Anne/Fligelstone, Steven/McEnery, Tony/Knowles, Gerry (eds.) (1997), *Teaching and Language Corpora.* London: Longman, 100-115.

Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1-23)*.* New York: Addison Wesley Longman.

McEnery, T./Xiao, R./Tono, Y. (2006): *Corpus-based language studies. An Advanced Resource Book.* London/New York: Routledge.

Meunier, Fanny (2002), The Pedagogical Value of Native and Learner Corpora in EFL Grammar Teaching. In: Granger/Hung/Petch-Tyson 2002, 119141.

Katrin Wisniewski & Andrea Abel

Mukherjee, Joybrato/Rohrbach, Jan-Marc (2006), Rethinking Applied Corpus Linguistics from a Language-pedagogical Perspective: New Departures in Learner Corpus Research. In: Kettemann, Bernhard/Marko, Georg (eds.), *Planing, Gluing and Painting Corpora. Inside the Applied Corpus Linguist's Workshop.* Frankfurt: Peter Lang, 205232.

Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In J.McH. Sinclair (Ed.), *How to use corpora in language teaching* (pp.125-152). Amsterdam: Benjamins.

Pawley, A./Syder, F.H. (1983): Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In: Richards, J.C./Schmidt, R.W. (eds.): *Language and Communication.* London: Longman, 191-226.

Römer, Ute. 2008. 7. Corpora and language teaching. In: Lüdeling, Anke & Merja Kytö (eds.). *Corpus Linguistics. An International Handbook (volume 1).* [HSK series] Berlin: Mouton de Gruyter. 112-130

Römer. U. (2006): Pedagogical applications of corpora: some reflections on the current scope and a wish list for future developments. In: Zeitschrift für ANglistik und Amerikanistik 54 (2) 121-134.

Römer, U. (2010): Using general and specialized corpora in English language teaching: past, present and future. In: Campoy-Cubillo, M. et al. (eds.): Corpus-based approaches to English Language Teaching. London: Continuum, 18-38.

Saville/Hawkey 2010

Tenfjord, K., Meurer, P. and Hofland, K. 2006: The ASK-corpus - a language learner corpus of Norwegian as a second language. Proceedings from 5th International Conference of Language Resources and Evaluation (LREC), Genova 2006. (http://hnk.ffzg.hr/bibl/lrec2006/pdf/573_pdf.pdf)

Katrin Wisniewski & Andrea Abel

Wisniewski, K. (2013): The empirical validity of the CEFR fluency scale: the A2 level description. In: Galaczi, E.D./Weir, C.J. (eds.): *Exploring Language Frameworks: Proceedings of the ALTE Krakow Conference, July 2011*. Cambridge: Cambridge University Press, 253-272.- Studies in Language Testing, 36.

Wisniewski, K.: (in preparation:)) *Die Validität der Skalen des Gemeinsamen europäischen Referenzrahmens für Sprachen. Eine empirische Untersuchung der Flüssigkeits- und Wortschatzskalen des GeRS am Beispiel des Italienischen und des Deutschen*. Frankfurt u.a.: Peter Lang. – Language Testing and Evaluation Series, hrsg. v. Rüdger Grotjahn und Peter Sigott.

Wisniewski, K./ Schöne, K./ Nicolas, L. /Vettori, C./Boyd, A./Meurers, D./Hana, J./Abel, A. (2013): MERLIN: An online trilingual learner corpus empirically grounding the CEFR Reference Levels in authentic data. In: *ICT for Language Learning, Conference Proceedings 2013*. Firenze: Libreriauniversitaria.

Wray, A./Perkins, M.R. (2000): The functions of formulaic language: an integrated model. In: *Language & Communication* 20, 1-28.

Katrin Wisniewski & Andrea Abel