

Scenario-based tasks for large-scale assessment

Development and validation

Katharina KARGES

Malgorzata BARRAS

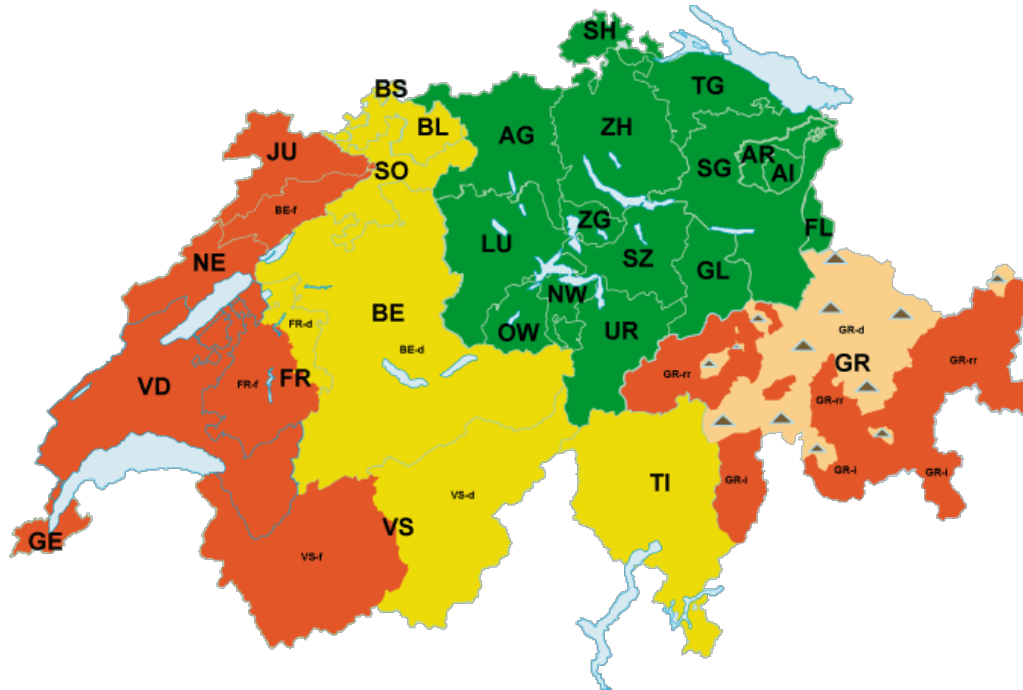
Peter LENZ

Research Centre on Multilingualism
University of Fribourg/Switzerland

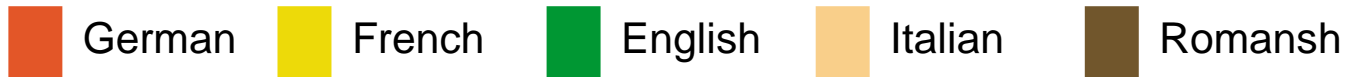
14th EALTA Conference 2017 (Sèvres)



Foreign language teaching in Switzerland

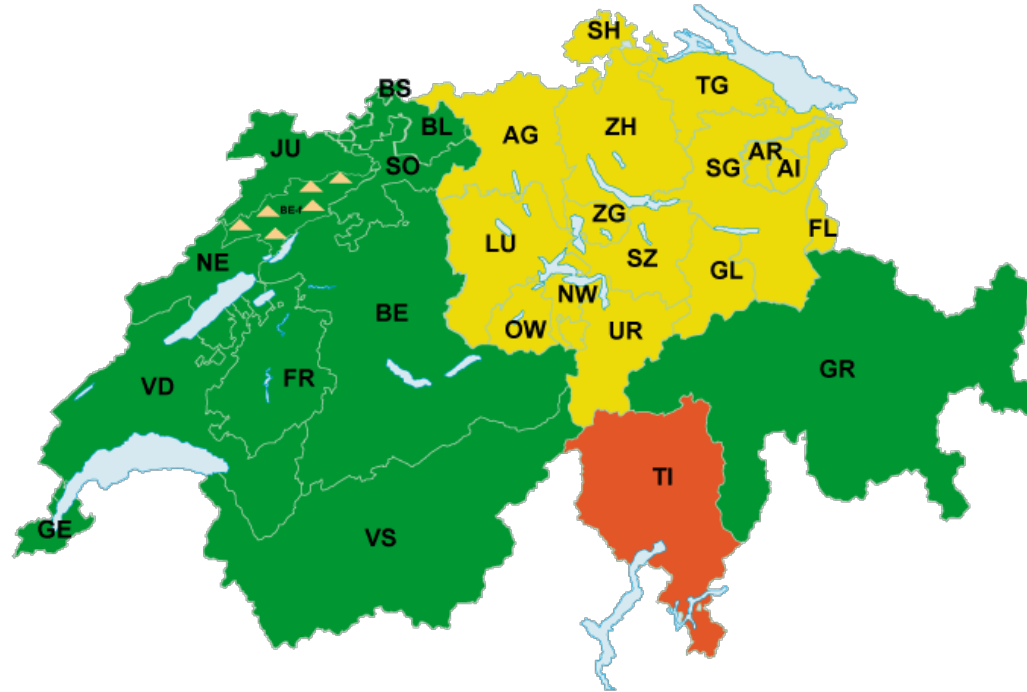


First foreign language (age 8)



Map adapted from: EDK – CDIP – CDPE – CDEP 2016.
[online: <http://www.edk.ch/dyn/15180.php>]

Foreign language teaching in Switzerland



Map adapted from: EDK – CDIP – CDPE – CDEP 2016.
 [online: <http://www.edk.ch/dyn/15180.php>]

Second foreign language (age 10-12)



National standards for foreign languages

- Introduction since 2011
- 2017: first nation-wide evaluation in the first foreign language (receptive competences: reading & listening; computer-based; age 12)
- 2020 (planned): nation-wide evaluation in two foreign languages (as before; age 15)

CEFR levels	A1.1	A1.2	A2.1	A2.2	B1.1	B1.2	B2
Listening							
Reading							
Spoken interaction							
Spoken production							
Writing							

Content

Scenario-based assessment

The project *Innovative Forms of Assessment*

Scenario-based assessment for foreign languages

Examples of scenario-based tasks

Qualitative test task research

Cognitive Validation

Pre-piloting

Future steps

Innovative Forms of Assessment

- Current project of the Swiss Research Centre on Multilingualism
- Based at the Institute of Multilingualism (University of Fribourg, University of Teacher Education Fribourg)
- Leader: Peter Lenz
- Research team: Thomas Aeppli, Nina Bechtiger, Katharina Karges, Joanna Lafine, Peter Lenz
- Associated PhD project: Malgorzata Barras

Innovative Forms of Assessment: Goals

- Development of “innovative” assessment tasks:
 - Scenario-based assessment (SBA)
 - Integrated: reading and listening
 - Computer-based
 - “Modern” text types (ICT)
 - French and English as foreign languages
- Investigation into the tasks’ functioning
 - Cognitive processes
 - Components of reading/listening competence
 - Student characteristics

Innovative Forms of Assessment: Instruments

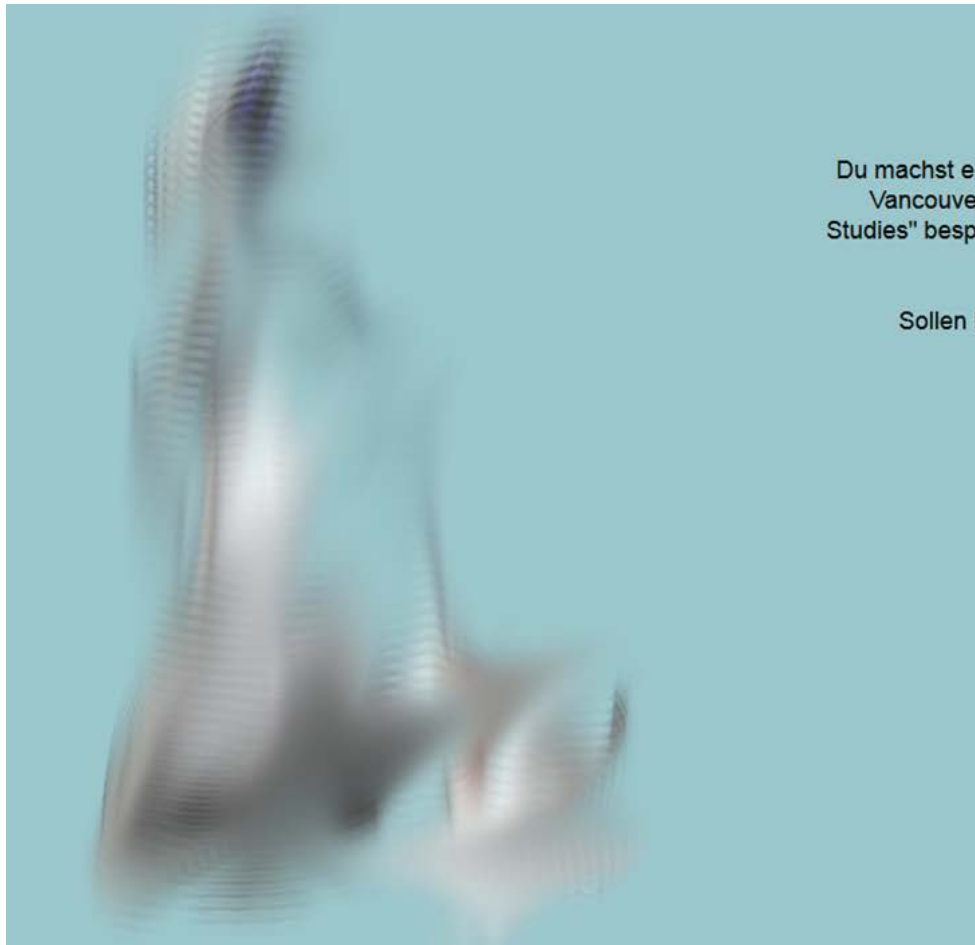
- 6-8 scenarios containing 4 tasks (12-15 items) each
- Component tests
 - Non-verbal intelligence
 - Vocabulary and grammar
 - General reading and listening skills in L2s
 - Questionnaire on metacognitive strategies and individual (linguistic) background
- Individual students: tests in either French or English

Scenario-based assessment

- Originally developed for reading in the language of schooling
- Broader construct of reading:
 - Task-oriented reading
 - Digital reading literacy
 - Integrated reading of related texts (evaluation & synthesis)
- Adaptation for foreign language large-scale assessment:
 - Reading and listening of quasi-authentic texts
 - Scenario descriptions, instructions and items in language of schooling
 - Closed item formats

Sabatini, O'Reilly, Halderman, & Bruce, 2014

Example: a scenario



Du machst ein
Vancouver
Studies" bespr

Sollen R

Klicke auf *Weiter*.

Weiter


You're abroad in Vancouver for a few weeks to learn English. You're going to school there. In "Social Studies" you're talking about a current referendum.

The subject: Should robots be allowed to walk dogs?

Example: a scenario

Task	Skill	Text	Activity	Format
A	reading	Informative text on the city website	Understand relevant details for a presentation in class	3-out-of-5 multiple choice
B	listening	Argument of another student	Understand general argument	1-out-of-3 multiple choice
C	reading	Discussion in a smartphone group chat	Understand instructions	1-out-of-3 multiple choice + justification
D	both	Encyclopedia article + radio interview	Understand relevant details	2-out-of-4 multiple choice

Innovative Forms of Assessment: Research steps

	Scenarios	Component tests	
Pre-piloting	Stimulated recall with individual students	-	
Piloting	Trial run in class Trial run of stimulated recall and think-aloud with individual students	Trial runs in small groups	<input data-bbox="1729 692 1767 749" type="checkbox"/>
Main study	Administration in class Think-aloud and stimulated recall with individual students	Administration in class	<input data-bbox="1729 942 1767 999" type="checkbox"/>

Cognitive validation

- Goal: collect evidence to support the validity of the scenario-based tasks
 - Identify and correct problems in the tasks (usability, content, difficulty, wording)
 - Investigate cognitive processes and test taking strategies used by students
- Method
 - Stimulated recall and observation during pre-piloting, piloting & main study
 - Think-aloud during piloting & main study
 - Quantitative analysis of relation between component skills and task results

Chapelle & Voss 2014; Cohen, 2006, 2014;
Kenyon & MacGregor 2012; Rossa, 2012

Pre-piloting: stimulated recall

- Immediately after each item
- Stimulus: reading or listening text(s) + item + student's answer
- 45 minutes per student, one-on-one with a researcher
- Audio-tapes and written notes

Pre-piloting: results

- Numerous changes in the formats and the appearance of the tasks
- Reading & listening for the test are still prevalent → scenarios & ICT text genres are welcome but motivate students less than expected
- Intelligence, memory, logical thinking, understanding of scenario setting ... seem to be relevant for test success

Piloting & main study: qualitative part

- Goals:
 1. Gain insights into students' test-taking processes and test-taking strategies
 2. Compare informational quality of think-aloud and stimulated recall
- Procedure:
 - 135 minutes (incl. training for think-aloud): think-aloud + stimulated recall
 - 60 minutes traditional listening and reading
 - 40 minutes questionnaire (metacognitive strategies + social and linguistic background)
- 2 students per class (total: \approx 20 students)

Main study: quantitative data

- Interaction of individual knowledge and skills with assessment tasks
- Each student: English OR French
 - 120 minutes scenario-based tasks
 - 60 minutes traditional listening and reading tasks
 - 60 minutes cognitive measures (intelligence, grammar, vocabulary, automaticity)
 - 20 minutes questionnaire
- 35 classes \approx 600 students
- L2 English, L3 French and vice-versa

Preliminary conclusions

- Development of scenario-based tasks for low language proficiency is time-consuming and particularly complex
- Students do not necessarily perceive the difference between SBA and “traditional” assessment tasks
- Verbal reports give valuable insight into reading/listening and task solving processes

June 2017:
Piloting

Autumn 2017:
Main study

2018-2020:
Data analysis

Comments?

Research Centre on Multilingualism

Institute of Multilingualism

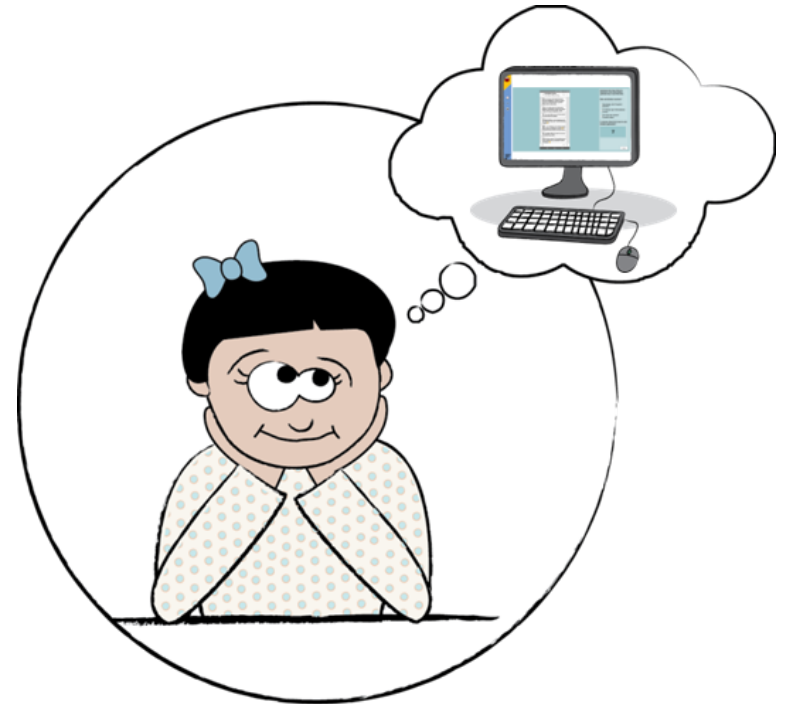
Rue de Morat 24

CH-1700 Fribourg

Switzerland

Mail: katharina.karges@unifr.ch
malgorzata.barras@unifr.ch
peter.lenz@unifr.ch

Web: www.centre-multilingualism.ch



References

- Chapelle, C. A., & Voss, E. (2014). Evaluation of language tests through validation research. In A. J. Kunnan (Ed.), *The companion to language assessment. Volume III: Evaluation, methodology, and interdisciplinary themes* (Vol. 3, pp. 1081–1097). Chichester, West Sussex, Malden: Wiley-Blackwell.
- Cohen, A. D. (2006). The coming of age for research on test-taking strategies. *Language Assessment Quarterly*, 3(4), 307–331.
- Cohen, A. D. (2014). Using test-wiseness strategy research in task development. In A. J. Kunnan (Ed.), *The companion to language assessment. Volume II: Approaches and development* (Vol. 2, pp. 893–905). Chichester, West Sussex, Malden: Wiley-Blackwell.
- Kenyon, D. M., & MacGregor, D. (2012). Pre-operational testing. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 295–306). London, New York: Routledge.
- Rossa, H. (2012). *Mentale Prozesse beim Hörverstehen in der Fremdsprache. Eine Studie zur Validität der Messung sprachlicher Kompetenzen*. Frankfurt am Main: Peter Lang.
- Sabatini, J. P., O'Reilly, T., Halderman, L., & Bruce, K. (2014). Broadening the scope of reading comprehension using scenario-based assessments: Preliminary findings and challenges. *L'Année Psychologique*, 114(4), 693–723.