



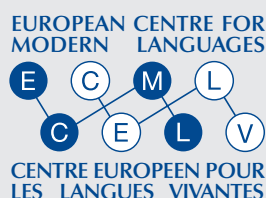
Relating language examinations to the common European reference levels of language proficiency

Frequently Asked Questions

compiled by
the RELANG team,
edited by José Noijons,
RELANG coordinator



EUROPEAN UNION



EUROPEAN CENTRE FOR
MODERN LANGUAGES
CENTRE EUROPEEN POUR
LES LANGUES VIVANTES



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Introduction

Relating language examinations to the common European reference levels of language proficiency” (RELANG) is an initiative carried out within the framework of a cooperation agreement between the European Centre for Modern Languages (ECML) and the European Commission. The initiative offers training and consultancy activities to educational authorities in the Member States of the European Union and the ECML in relating language examinations to the levels of proficiency defined in the Common European Framework of Reference for Languages in a valid and equitable way.

These Frequently Asked Questions (FAQs) have been collected by the RELANG team during their workshops on Relating Tests and Examinations to the CEFR. The workshops have been held in over 15 European countries, with stakeholders such as teacher trainers, test item writers, examiners, curriculum developers and ministry officials. The answers to the FAQs should be regarded as first explorations of topics in aligning tests and examinations to the CEFR. It has been found that in all areas of language education linking examinations to the CEFR is a rather complex issue. These FAQs and the accompanying answers seek to assist stakeholders in identifying issues within their local context.

RELANG team: José Noijons (coordinator), Evelyne Bérard, Jana Bérešová, Gábor Szabó

To find out more about the initiative visit <http://relang.ecml.at>

The opinions expressed in this work are the sole responsibility of the authors and do not necessarily reflect the official policy of the Council of Europe. All rights reserved. No part of this publication may be translated, reproduced or transmitted in any form or by any means, electronic (CD-Rom, Internet, etc.) or mechanical, including photocopying, recording or any information storage or retrieval system, without prior permission in writing from the Directorate of Communication (FR-67075 Strasbourg Cedex or publishing@coe.int).

Council of Europe Publishing
FR-67075 Strasbourg Cedex
<http://book.coe.int>

European Centre for Modern Languages of the Council of Europe
Nikolaipplatz 4
AT-8020 Graz
www.ecml.at

© Council of Europe, December 2015

Frequently asked questions

1. Can tests of different skills be summarized at one CEFR level?
2. Do we need to present authentic real-life texts in reading and listening tests?
3. How can I be sure that a student's performance in speaking or writing is judged at the desired CEFR level?
4. Is it necessary to provide a context for tasks?
5. Once an exam is linked to the CEFR, how can I make sure that new versions of the exam are linked to the CEFR as well?
6. Can we link our tests to the CEFR levels even though we do not have statistics on student performance on our tests, apart from total scores?
7. Can we test at more than one CEFR level in one test?
8. How do we make sure that our pass/fail decisions are related to the CEFR?
9. How do we determine that a student performance in reading or listening is at a specific CEFR level?
10. Can an exam that tests linguistic competence (language use) be linked to the CEFR?
11. How do I know whether a test is at the CEFR level it claims to be at?
12. Do we give the same number of points for each item in reading and listening sub tests?

1. Can tests of different skills be summarized at one CEFR level?

Some language tests claim that they measure a student's language skills at one or more CEFR levels. One must check what the validity of such a claim would be. We cannot simply "average out" performances in different skills. In real life most learners are better at one skill than another, certainly at the lower CEFR levels. Thus, in a test we may be able to average score points, but we cannot average CEFR levels. We may be able to say that the student is at B2 for reading and at A2 for writing. We cannot then say that the student is at B1 for reading and writing combined.

Table 1 in the CEFR (Common Reference Levels: global scale) is often misunderstood as meaning that what is described for a particular level is what is to be expected from a language user at that level, for all language skills. This table however should be interpreted as a description of what a person can do at that level and with particular skills: he or she can function at the given level for reading and listening, but at another level for speaking and writing.

The Council of Europe has advocated the development of profiles, in which the student's proficiencies in the various language skills are described. The European Language Portfolio has also adopted this approach.

- Encourage students to find out about their strengths and weaknesses for each language skill through tests of specific skills.
- Report test results for each language skill separately.

References

- Council of Europe (2009), [Language Education Policy Profiles](#) - A transversal analysis: trends and issues, Strasbourg: Council of Europe
- Council of Europe (1991), [European Language Portfolio](#) (ELP), Strasbourg: Council of Europe

Suggestions for action

- When developing a language test, give separate CEFR level-estimations for each language skill tested.
- Make sure that there is a sufficient number of tasks for each language skill to be able to give a valid estimation of the CEFR level.

2. Do we need to present authentic real-life texts in reading and listening tests?

Texts that are taken from real life and that have a communicative function link up to the CEFR model of language use and would therefore be welcome in CEFR-based language tests. This is not to say that such texts cannot be edited for technical reasons (texts may be too long for inclusion, incidental words may cause undue problems of understanding at the intended CEFR-level). Such editing is permissible, both from a validity point of view and from a legal point of view, as long as certain rules of good practice are observed.

The selection of listening samples may be problematic for various reasons. Authentic materials may be difficult or expensive to obtain, the sound quality may not be acceptable, the cost of producing varied listening samples may be too high. Yet it must be avoided that listening is tested by making use of reading texts that are read out by one or two actors.

Suggestions for action

- **Always** mention the source of a text. Indicate when changes have been made ("Adapted from ...").
- When no author is known or mentioned, it is advisable not to refer to an author in the items. Refer to the text itself, as in: What does the text say about ...?
- When editing a text, make sure that the original message of the text is maintained.
- When editing a text, make sure that there is a clear beginning and ending.
- Select listening samples that are in the public domain, on the Internet (e.g. YouTube).

- In listening tests: if editing existing texts is necessary, use texts that have originally been spoken.
- In recording listening texts: use native speakers to produce oral texts.
- Try and produce semi-authentic recordings, where the speakers do not "read out" texts but create utterances spontaneously on the basis of cues.
- Try to have a varied group of actors with different accents to read out selected texts.
- A certain degree of background noise in recordings is acceptable. In real life most listening activities will be accompanied by background noise.
- Select types of texts that go with specific descriptors at specific CEFR levels.
- And above all, choose texts appropriate for the age of the test takers.

References

- ALTE (2011), **Manual for Language Test Development and Examining**; For use with the CEFR, Strasbourg: Council of Europe. See section 1.1.3 Operationalising the model, section 3.3.2 Commissioning, section 3.4.1 Editing new materials. See also: Appendix 4, Advice for item writers.

3. How can I be sure that a student's performance in speaking or writing is judged at the desired CEFR level?

When judging a student's performance in speaking or writing we can make use of benchmarks: detailed, validated descriptions of specific levels of student performance that is expected at a particular CEFR level. Such benchmarks are often represented by samples of student work (writing) or videos (speaking).

Benchmarks can be produced in the following way. A coordinator selects a number of oral or written student sample performances to introduce a specific CEFR level. For each sample a group of experts then judges and discusses whether the sample does indeed illustrate the level, and why it is not the level above or below this. After reconsideration the group then votes on the level of each performance. During a last phase the individual group members rate the performances once again and compare their scores.

It must be emphasized here that benchmarking is a group process, rather than one expert showing and telling the other experts which performances best illustrate a performance at the desired CEFR level.

Suggestions for action

- Collect a number of relevant student samples.
- Select a panel of experts (around 10).
- Make sure that the experts are familiar with the CEFR and its descriptors through familiarisation activities.
- Collect all materials for the benchmarking session.
- Inform panellists beforehand about the details of the benchmarking procedure.

- Carry out a benchmarking procedure as suggested in the linking Manual.
- Publicize the results of the benchmarking session.

References

- Council of Europe (2009), Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR), [A Manual](#), Strasbourg: Council of Europe. See Chapter 5 Standardisation training and benchmarking.
- Noijons, José & Jana Bérešová, Gilles Breton, Gábor Szabó (2011), Relating language examinations to the Common European Framework for Languages: Learning, Teaching, Assessment (CEFR). [Highlights from the Manual](#), Graz: ECML. See Chapter 5 Standardisation training and benchmarking
- Council of Europe (2008), [Report on the cross linguistic benchmarking seminar to calibrate examples of spoken productions in English, German, French, Italian, Spanish with regard to the six levels of the CEFR](#), Strasbourg: Council of Europe.

4. Is it necessary to provide a context for tasks?

In the case of speaking and writing tasks this question may seem to be rhetorical. In real life, how we speak and write much depends on the circumstances we are placed in. So if speaking and writing tasks are to be authentic, we cannot usually do without providing a context. Generally speaking, most speaking and writing tasks are therefore placed in a context.

In theory it may be possible to ask a student to speak or write about a subject without providing a context. This is often done when students are to give their (personal) opinion about a subject, a phenomenon or an incident. Yet from a CEFR point of view this may be questionable: in our communication with others we need to think about who we are addressing and why. To simply give our/an opinion without thinking of the person we are addressing may be rather counterproductive, it may hurt other people's feelings or it may simply not be understood or indeed be misunderstood. We must also realize that the aim of assessing speaking or writing in a foreign language is not to test a student's ability to express a view or an opinion, but rather to test if the student can express that opinion in the foreign language. In other words: we need to test if the student can express a view or an opinion, but we do not assess the content of the message (e.g. facts, data, etc.).

In the case of reading and listening tests, we often find that students are instructed to read a text or listen to a passage without any context given ("Read the following text and answer the questions."). It cannot be denied that in such cases we may be testing reading or listening. However, from a CEFR point of view we would also expect a context when testing reading and listening: we need to give a reason for reading the text.

Suggestions for action

- In testing spoken and written interaction it is advisable to provide the students with contexts that are realistic and link up with their age and their experience in life. We also need to remember that there are cultural differences between speakers and writers from different backgrounds. Students may or may not be comfortable in saying or writing certain things.
- In testing spoken production it is also advisable to provide a context based on a printed text or on a visualisation (pictures, photos, etc.). This may help students in constructing a view or an opinion. To avoid us testing the students' opinions or views or imagination or cultural knowledge, rather than their ability to express them, it is advisable to provide key words or arguments for and against.
- In tests of reading or listening, by providing students with authentic texts placed in a realistic context, and with the student being given a purpose to read or listen, we may improve the validity of reading and listening tests.
- When providing contexts bear in mind that there are unsuitable contexts such as war, politics, racism (including cultural clichés and stereotyping), sex and sexism (including stereotyping), potentially distressing topics (e.g. death, terminal illness, severe family and social problems, natural disasters, and the object of common phobias such as spiders and snakes – where the treatment might be distasteful), examinations, passing and failing, drugs.

References

- Council of Europe (2001), **Common European Framework of Reference for Languages**: Learning, teaching, assessment, Cambridge: Cambridge University Press. See section 4.1 The context of language use and Table 5, External context of use: descriptive categories.
- ALTE (2011), **Manual for Language Test Development and Examining; For use with the CEFR**, Strasbourg: Council of Europe. See Sections 1.4 Ethics and fairness and 3.4.1 Editing new materials. See also Appendix IV, Advice for item writers.

5. Once an exam is linked to the CEFR, how can I make sure that new versions of the exam are linked to the CEFR as well?

The most obvious answer to this question would be to apply all the steps of linking the new exam to the CEFR: familiarisation, specification, standardization, standard setting and validation. Some steps may be made easier if in a matrix test specifications have been produced: what is tested, how it is tested, the number of items, the item types, the types of texts etc. This is a way to make sure that a test measures the same construct with the same CEFR-related skills as earlier versions.

What is ideally needed is to run a pre-test of the new exam with a representative selection of items of the earlier exam embedded (so-called anchor items). With the help of advanced statistics it would then be possible to set standards comparable to those of the earlier exam. If the earlier standards have been related to the CEFR, one may argue that the new exam is linked to the CEFR. In fact this is part of the linking process of validation.

Suggestions for action

- Produce a test matrix, related to the CEFR, with reference to CEFR descriptors and levels, containing test specifications.
- Collect data in pre-tests, preferably using anchor items to link exams to each other.
- Possibly adapt standards when a new version of the exam proves to be easier or more difficult than the older one.

References

- ALTE (2011), **Manual for Language Test Development and Examining; For use with the CEFR**, Strasbourg: Council of Europe. See Sections 2.4, 2.5 and Appendix
- Council of Europe (2009), Relating language examinations to the Common European framework of reference for languages: learning, teaching, assessment (CEFR), **A Manual**, Strasbourg: Council of Europe. See Chapter 7 Validation.
- Noijons, José & Jana Bérešová, Gilles Breton, Gábor Szabó (2011), Relating language examinations to the Common European Framework for Languages: Learning, Teaching, Assessment (CEFR). **Highlights from the Manual**, Graz: ECML. See Chapter 7 Validation.

6. Can we link our tests to the CEFR levels even though we do not have statistics on student performance on our tests, apart from total scores?

In general the validity and the reliability of a test will be enhanced when on the basis of data of student performance on the test, the test itself and/or the format of the test is adapted (difficulty of tasks, type of tasks, length of the test etc.). The first and most important step in the linking process is to make sure that the test in question is valid (that it indeed tests what it claims to be testing) and reliable (that the testing is consistent). Ideally this is done through pretesting the collection of data on the test. However, there are situations where such a procedure is not possible or too costly, such as in classroom-based testing.

Suggestions for action

If it is not possible to collect evidence that a test is valid and reliable through statistics, we can nevertheless make an attempt to link the test to the CEFR through specification. In fact specification is a phase in the linking process that always needs to be carried out.

The specification phase in the linking process helps to raise awareness among test providers of:

- the importance of good content analysis of language examinations to ensure content validity;
- the CEFR, especially its descriptor scales;
- the rationale for relating language examinations to an international framework like the CEFR; and
- ways in which the CEFR can be exploited in planning and describing language examinations.

There are four steps to be taken in the specification phase:

- Assuring adequate familiarisation with the CEFR by those describing the test;
- Analysing the content of the test in question in relation to the relevant categories of the CEFR (should an area tested not be covered by the CEFR, the user is asked to describe it);
- Profiling the test in relation to the relevant descriptor scales of the CEFR on the basis of this content analysis;
- Making a first claim on the basis of this content analysis that an examination or test in question is related to a particular level of the CEFR.

References

- Council of Europe (2009), Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), **A Manual**, Strasbourg: Council of Europe. Appendix A, sections A2 to A5 provides a helpful set of forms (A1-A24) that can help test developers specify a test in terms of the CEFR.
- ALTE (2001). Principles of good practice for ALTE examinations; Revised draft October 2001, www.alte.org

7. Can we test at more than one CEFR level in one test?

In theory this is possible, in practice this may be challenging. It also depends on the type of test and the skill that is being tested. A test that is (to be) linked to the CEFR is supposed to tap a representative number of descriptors at the desired CEFR level. For each descriptor at each level we would need a sufficient number of items to be able to give a valid judgement on whether the student can do what is described in the descriptor. In practice this may mean that tests of reading and listening would have to be longer than is feasible.

In the case of speaking tests, there are formats such as in the Oral Proficiency Test (OPI), where a trained interlocutor moves from one level to another depending on the proficiency of the candidate. In such tests it would be possible to find out whether the student is able to function at more than one CEFR level. It must be stressed here that interlocutors will need to be thoroughly trained in administering such a test. Generally speaking this would not be within the reach of untrained teachers.

There exist computer-based adaptive tests, in which the students are presented with items at various levels, depending on the responses that they give. In principle a more difficult task (possibly at a higher CEFR level) is presented every time the student gives a correct response. In this way testing time may be reduced considerably. This means that numerous items are needed to create an item bank. However, creating a calibrated item bank is costly, and needs time.

The CEFR is built on the idea that a person who can perform at a given CEFR level can also perform at the level(s) below the given level. A person at B1-level is supposed to be

able to perform at levels A2 and A1 as well. However, this does not mean that we can simply assign levels A2 or A1 to the student when he or she has a low score on a B1-level test, for the reasons outlined above.

Suggestions for action

- When testing receptive skills such as reading and listening, aim at one level only. If students score high on a one-level test, set them a test at a higher level (and the reverse).
- When testing writing, set students a series of tests with tasks that are linked to descriptors at one level only. If students score high on such one-level tests, set them a series of tests at a higher level (and the reverse).
- When testing speaking, either train interlocutors to administer OPI type of tests, or do as with administering writing tests (see above).

References

- ALTE (2011), **Manual for Language Test Development and Examining; For use with the CEFR**, Strasbourg: Council of Europe. See Chapter 5, Marking, grading and reporting of results and Appendix VII.

8. How do we make sure that our pass/fail decisions are related to the CEFR?

In many countries pass/fail scores in exams are laid down in the law or described in the syllabus, without reference to the CEFR. Thus it is possible for students to pass an exam at a given CEFR level without attaining a score that would indicate that students have a proficiency at the desired CEFR level.

For a pass/fail score that is related to the CEFR we need to carry out a standard-setting procedure (for the receptive skills) or a benchmarking procedure (for the productive skills). In these procedures a group of experts determines what minimum score is needed for the students to claim that they have reached the desired level. In the case of speaking or writing tests, performances can be selected by experts that illustrate how students should perform for them to be graded at a specific CEFR level.

It is thus possible for a student to have a score on the exam that indicates two things: (1) the student has or has not passed the exam from a legal perspective and (2) the student has or has not reached the desired CEFR level.

It is often said in syllabuses that the exam is at a given CEFR level. However, if there has been no CEFR-related standard setting or benchmarking, the scores on that exam cannot be said to be related to the CEFR.

Suggestions for action

- Redesign the pass/fail decisions in the syllabus in terms of the CEFR.
- Carry out standard-setting procedures (receptive skills) and benchmarking procedures (productive skills).

- See also FAQs: How can I be sure that a student's performance in speaking or writing is judged at the desired CEFR level? and How do we determine that a student performance in reading or listening is at a specific CEFR level?

References

- Council of Europe (2009), Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), **A Manual**, Strasbourg: Council of Europe. See Chapter 5 Standardisation training and benchmarking and Chapter 6 Standard setting procedures.
- Noijons, José & Jana Bérešová, Gilles Breton, Gábor Szabó (2011), Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR). **Highlights from the Manual**, Graz: ECML. See Chapter 5 Standardisation training and benchmarking and Chapter 6 Standard setting procedures
- Council of Europe (2008), **Report on the cross linguistic benchmarking seminar to calibrate examples of spoken productions in English, German, French, Italian, Spanish with regard to the six levels of the Common European Framework of Reference for Languages** (CEFR), Strasbourg: Council of Europe.

9. How do we determine that a student performance in reading or listening is at a specific CEFR level?

When grading a student's performance in reading or listening we may need to set a performance standard. This is the boundary or cut score between two scores on a performance scale. A cut score of 30, for example, says that a score of 30 or more indicates a performance at a particular level (for example B1) while a lower score indicates that the student has not reached the desired level.

There are various ways to set standards. It has been found that applying two or more of such methods may yield the best results. For all these methods a coordinator needs to gather student scores on a reading or listening tests. A number of such methods are described in the linking Manual (see References below). In cases such as reading or listening tests when numerical scores are given, experts estimate at what CEFR level a test taker can be expected to respond correctly to a set of items.

It must be emphasized here that standard setting is a group process, rather than one expert showing and telling the other experts which score is required to determine if a performance is at the desired CEFR level.

Suggestions for action

- Collect a relevant set of data of student performance on a test.
- Select a panel of experts (around 10).
- Make sure that the experts are familiar with the CEFR and its descriptors through familiarization activities.
- Collect all materials for the standard-setting session.

- Inform panellists beforehand about the details of the standard-setting procedure.
- Carry out standard-setting procedures as suggested in the linking Manual.
- Publicize the results of the standard-setting session.

References

- Council of Europe (2009), Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), [A Manual](#), Strasbourg: Council of Europe. See Chapter 6 Standard setting procedures.
- Noijons, José & Jana Bérešová, Gilles Breton, Gábor Szabó (2011), Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR). [Highlights from the Manual](#), Graz: ECML. See Chapter 6, Standard setting procedures

10. Can an exam that tests linguistic competence (language use) be linked to the CEFR?

Even if the CEFR acknowledges that linguistic competence is an important aspect of language competence, it may be difficult to link sections in an exam that test subskills such as grammar or vocabulary to the CEFR. It must be noted that the formulations of the CEFR descriptors for linguistic competence are rather general and can be interpreted in many ways. For some languages (such as French and German) more detailed descriptors have been developed. However, these have not been scaled in the same way as those in the CEFR itself have been.

The problem is that often exams tend to focus on issues in vocabulary and grammar that learners with a specific first language background find difficult when learning a particular foreign language. Such sections in an exam may focus on the structure of a language rather than on the communicative aspects of it. Such sections do not necessarily focus on grammatical constructions that are typical of written texts produced in various contexts at various levels, and the vocabulary going with those. Such sections are not usually linked to specific linguistic CEFR descriptors at various levels.

From a formative point of view such a linguistic focus is understandable. However, in summative situations, if the curriculum and the syllabus claim that students should be able to function at specific CEFR levels at the end of secondary school, then it is to be wondered if an exam that is to be linked to the CEFR should contain (large) sections on linguistic competence.

It can be argued that when testing reading and listening we also test a student's understanding of the structure and the

vocabulary of a language. We may argue the same for tests of speaking and writing: if assessment criteria such as the use of vocabulary and grammatical structures are applied, then there would seem no need to discretely measure vocabulary and structures.

Suggestions for action

- Adopt the action-oriented approach of language use as expressed in the CEFR.
- When testing reading and listening, focus on understanding the text, its message, its style, its genre, the author's (speaker's) intention with reference to relevant CEFR descriptors .
- When testing speaking and writing, use formal criteria: accuracy, range, fluency, interaction, coherence, cohesion, with reference to CEFR criteria.

References

- Council of Europe (2001), **Common European framework of reference for languages: learning, teaching, assessment**, Cambridge: Cambridge University Press. See section 5.2.1 Linguistic competences; Table 3.
- Council of Europe (2009), **Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), A Manual**, Strasbourg: Council of Europe. See Tables C1-C4.

11. How do I know whether a test is at the CEFR level it claims to be at?

Many testing organisations and publishing companies claim that the tests they administer or publish are at a given CEFR level. The validity of such claims may be very important for test takers. On the basis of their results, they may be admitted to further education or hired for a job. There is also a need for institutes and employers to be able to depend on the validity of claims of links to the CEFR and to specific CEFR levels in particular.

Suggestions for action

It is obvious that without sufficient proof of the validity of claims of links to CEFR levels, such claims cannot be trusted. Ideally such proof should be included in the test materials. However, such test materials may potentially refer to documents that are confidential and thus inaccessible to the general public. Yet some published information on linking should be available. Such information may contain evidence of various types:

- Through field testing the test has been shown to be valid and reliable;
- The test has been specified in terms of the CEFR: it links up with the action-oriented, communicative approach the CEFR takes towards language learning;
- The test has been standardized: scores on performances have been linked to levels in the CEFR through benchmarking and standard –setting;
- There is independent evidence which corroborates the results of the standardization: empirical validity.

It will not always be easy to find enough evidence of the validity of a test's claims to links to the CEFR levels. Very often the

only validation of such claims is through a specification of the content of the test in terms of the CEFR, such as low-stakes classroom-based tests. For some tests, the resulting evidence may be sufficient. However, for high-stakes tests all the types of evidence as mentioned above are needed for the links to the CEFR level to be called valid.

References

- Council of Europe (2009), *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), A Manual*, Strasbourg: Council of Europe. See Figure 2.2 and Chapter 7.
- Noijons, José & Jana Bérešová, Gilles Breton, Gábor Szabó (2011), *Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR). Highlights from the Manual*, Graz: ECML. See Figure 2.2 and Chapter 7
- ALTE (2011), *Manual for Language Test Development and Examining; For use with the CEFR*, Strasbourg: Council of Europe. See Appendices 1 and VII

12. Do we give the same number of points for each item in reading and listening sub tests?

Some items are on occasion given more weight than other items because they are thought to be more difficult than others. If the item consists of a number of operations this is acceptable, if the students know that the item is worth more points. In other cases it is hardly necessary to weight items. We will be able to distinguish good students from less good students because in principle the less good students might give an incorrect response and gain no points on that item.

From a CEFR-point of view there is also an issue in weighting items. If one item is considered to be more difficult than another item, then it must be wondered if that item may be tapping a descriptor at a higher level. As is argued in another FAQ (Can we test at more than one CEFR level in one test?), it is advisable to create homogenous tests aimed at one CEFR-level only.

There is one more issue in weighting items. Item writers or indeed the syllabus itself may claim that certain items are more difficult than other items. Without data on the performance of these items such claims are not valid.

Suggestions for action

- In the case of a complex item, try to rephrase the item and turn it into a set of items.
- When complex items need to be used, always inform the students about the number of points that can be gained for such an item.
- When testing language skills, aim at one CEFR level only. If students score high on a one-level test, set them a test at a higher level (and the reverse).
- When constructing language tests, always try to collect data on item performance, preferably also in a pre-test.

12. Donnons-nous le même nombre de points à chaque item dans les sous tests de compréhension écrite et de compréhension orale ?

- Certains items sont l'occasion de leur donner plus de poids qu'à d'autres car on estime qu'ils sont plus difficiles que d'autres. Si l'item consiste en un nombre d'opérations c'est acceptable si les étudiants savent que l'item vaut plus de points. Dans d'autres occasions, il n'est pas nécessaire d'apprécier les items. Nous allons être capables de distinguer les bons étudiants des étudiants moins bons parce qu'en principe les étudiants moins bons devraient donner une réponse incorrecte et ne pas gagner de points à cet item.
- Du point de vue du CECR il y a aussi un problème quand on mesure les items. Si un item est considéré comme plus difficile qu'un autre, alors il faut se demander si cet item peut être lié à un descripteur à un niveau plus élevé. Comme cela est expliqué dans une autre FAQ (Peut-on tester plus d'un niveau du CECR dans un seul test ?), il est recommandé de créer des tests homogènes reliés seulement à un niveau du CECR.
- Il y a une autre question quand on apprécie les items. Les rédacteurs d'items ou le syllabus même peuvent déclarer que d'autres. Sans données sur la performance de ces items, de telles affirmations ne sont pas valides.
- Quand des items complexes doivent être utilisés, informez toujours les étudiants sur le nombre de points qu'ils peuvent gagner pour un tel item.
- Dans le cas d'un item complexe, essayez de reformuler l'item et de le placer dans un ensemble d'items.
- Quand vous testez les compétences langagières, référez-vous à un seul niveau du CECR. Si les étudiants obtiennent un score élevé à un test sur un niveau, donnez-leur un test à un niveau plus élevé (et l'inverse).
- Quand vous élaborerez des tests de langue, essayez de recueillir des données sur la performance des items, de préférence également avec un pré-test.

Suggestions

- Dans le cas d'un item complexe, essayez de reformuler l'item et de le placer dans un ensemble d'items.
- Quand des items complexes doivent être utilisés, informez toujours les étudiants sur le nombre de points qu'ils peuvent gagner pour un tel item.

11. Comment puis-je savoir si un test est au niveau du CECR auquel il déclare être ?

De nombreuses organisations et maisons d'édition déclarent que les tests qu'ils administrent ou publient sont à un niveau donné du CECR. La validité de telles affirmations peut être très importante pour les candidats. Sur la base de leurs résultats, ils doivent être admis pour suivre des études ou engager pour un travail. C'est aussi une nécessité pour des instituts et des employeurs d'être capables de dépendre de la validité des affirmations de liens avec le CECR et avec les niveaux spécifiques du CECR en particulier.

Suggestions

Il est évident que sans preuve suffisante de la validité des déclarations de liens avec les niveaux du CECR, on ne peut faire confiance à de telles affirmations. Idéalement de telles preuves devraient être incluses dans les matériaux du test. Cependant de tels matériaux de test peuvent potentiellement se référer à des documents qui sont confidentiels et ainsi inaccessibles au grand public. Toutefois certaines informations publiées sur les liens avec le CECR devraient être disponibles. Ces informations doivent contenir des preuves de différents types :

- Par des pré-tests, le test a montré qu'il est valide et relié ;
- Le test a été spécifié dans les termes du CECR : il est articulé avec l'approche actionnelle, l'approche communicative que le CECR suit par rapport à l'apprentissage d'une langue ;
- Le test a été calibré : des scores de performances ont été reliés au CECR à travers le calibrage et la définition de scores de césure ;
- Il y a une preuve indépendante qui corrobore les résultats du calibrage : validité empirique.

Références

Il n'est pas toujours facile de trouver assez de preuves de la validité des déclarations qu'un test est relié aux niveaux du CECR. Souvent la seule validation de telles affirmations se fait à travers la spécification du contenu du test dans les termes du CECR, comme les tests réalisés en classe à faible enjeu. Pour certains tests, la preuve du résultat peut être suffisante. Cependant pour les tests à fort enjeu, tous les types de preuves comme mentionné plus haut sont nécessaires pour que les liens avec le niveau du CECR soient déclarés valides.

- Conseil de l'Europe (2009). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) **Un Manuel** Strasbourg Conseil de l'Europe. Voir Tableau 2.2. et Chapitre 7.
- Nojons, José et Jana Béréšová, Gilles Breton, Gábor Szabó (2011). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) **Les points essentiels du Manuel**. Graz, CELV. Voir Tableau 2.2. et Chapitre 7.
- ALTE (2011). **Manuel pour l'élaboration et la passation des tests et d'examens de langue**. A utiliser en liaison avec le CECR. Strasbourg. Conseil de l'Europe. Voir Annexes 1 et 7.

10. Est-ce qu'un examen qui teste la compétence linguistique (utilisation de la langue) peut être relié au CECR ?

Même si le CECR reconnaît que la compétence linguistique est un aspect important de la compétence langagière, il peut être difficile de relier des parties d'examen qui teste des sous-compétences comme la grammaire et le vocabulaire au CECR. Il faut noter que les formulations des descripteurs du CECR pour la compétence linguistique sont plutôt générales et peuvent être interprétées de plusieurs façons. Pour certaines langues (comme le français et l'allemand) des descripteurs plus détaillés ont été développés. Cependant ils n'ont pas été placés sur une échelle de la même façon que ceux du CECR lui-même l'ont été.

Le problème est que les examens tendent souvent à mettre l'accent sur des problèmes en vocabulaire et en grammaire que les apprenants, avec un contexte spécifique de première langue, trouvent difficiles quand ils apprennent une langue étrangère particulière. De telles parties d'un examen peuvent sans doute mettre l'accent sur la structure de la langue plutôt que sur les aspects communicatifs de celle-ci. De telles parties d'examen ne mettent pas nécessairement l'accent sur des constructions grammaticales qui sont représentatives de textes écrits dans des contextes variés à différents niveaux et sur le vocabulaire qui va avec celles-ci. De telles parties ne sont pas en général reliées à des descripteurs linguistiques spécifiques du CECR à différents niveaux.

D'un point de vue formatif, une telle attention sur le linguistique est compréhensible. Cependant dans des situations d'évaluation sommative, si le curriculum et le syllabus déclarent que les étudiants doivent être capables de fonctionner à des niveaux spécifiques du CECR à la fin de l'école secondaire, il faut alors se demander si un examen qui est relié au CECR devrait contenir de (grandes) sections sur la compétence linguistique.

Suggestions

On pourrait défendre que lorsqu'on teste la compréhension écrite ou la compréhension orale, on teste également la compréhension par l'étudiant de la structure et du vocabulaire d'une langue. On peut défendre la même chose pour la production orale et la production écrite : si les critères d'évaluation comme l'utilisation du vocabulaire et des structures grammaticales sont appliqués, alors il ne paraîtrait pas nécessaire d'évaluer de façon distincte le vocabulaire et les structures.

Références

- Adoptez l'approche actionnelle de l'utilisation de la langue comme indiqué dans le CECR.
- Quand vous testez la compréhension écrite ou la compréhension orale, mettez l'accent sur la compréhension du texte, son message, son style, son genre, l'intention de l'auteur (du locuteur) en référence avec les descripteurs pertinents du CECR.
- Quand vous testez la production orale ou la production écrite, utilisez des critères formels : précision, variété, fluidité, interaction, cohérence, cohésion en référence aux critères du CECR.
- Conseil de l'Europe (2001). **Cadre européen commun de référence pour les langues** : apprendre, enseigner, évaluer. Didier. Voir Section 5.2.1. Compétence linguistique et Tableau 3.
- Conseil de l'Europe (2009). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) **Un Manuel** Strasbourg Conseil de l'Europe. Voir Tableaux C1-C4.

9. Comment déterminons-nous que la performance d'un étudiant en compréhension écrite ou en compréhension orale est à un niveau spécifique du CECR ?

Quand nous classons la performance d'un étudiant en compréhension écrite ou en compréhension orale, il peut être utile de fixer un standard de performance. C'est la limite ou le score de césure entre deux scores sur une échelle de performance. Un score de césure de 30 par exemple signifie qu'un score de 30 ou plus indique une performance à un niveau particulier (par exemple B1) alors qu'un score inférieur indique que l'étudiant n'a pas atteint le niveau souhaité.

Il y a plusieurs façons de fixer les scores de césure. Il a été montré qu'appliquer deux de ces méthodes ou plus pouvait produire les meilleurs résultats. Pour toutes ces méthodes, un coordinateur a besoin de recueillir les scores d'étudiants en compréhension écrite et en compréhension orale. Un certain nombre de ces méthodes sont décrites dans le Manuel Relier (voir les Références ci-dessous). Dans les cas de tests de compréhension écrite ou de compréhension orale quand on donne des scores numériques, les experts estiment à quel niveau du CECR on peut attendre qu'un candidat réponde correctement à un ensemble d'items.

- Recueillir un ensemble pertinent de données de performances d'étudiants à un test.

Suggestions

Il faut insister ici sur le fait que la définition des scores de césure est un processus de groupe, plutôt que le fait d'un expert qui montre et dit aux autres experts quel score est requis pour déterminer si la performance est à un niveau souhaité du CECR.

- Sélectionner un panel d'experts (autour de 10).
- S'assurer que les experts sont familiers du CECR et de ses descripteurs à travers des activités de familiarisation.
- Recueillir tous les matériaux pour une session de définition des scores de césure.
- Informer préalablement les panelistes des détails de la procédure de définition des scores de césure.
- Réaliser les procédures de définition des scores de césure.
- Publier les résultats de la session de définition des scores de césure.

Références

- Conseil de l'Europe (2009). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) *Un Manuel* Strasbourg Conseil de l'Europe. Voir Chapitre 6 Procédures de définition des scores de césure.
- Nojons, José et Jana Běrešová, Gilles Breton, Gábor Szabó (2011). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) *Les points essentiels du Manuel*. Graz, CELV. Voir Chapitre 6 Procédures de définition des scores de césure.

8. Que pouvons-nous faire pour être sûrs que nos décisions de réussite ou d'échec sont reliées au CECR ?

Dans de nombreux pays, les scores de réussite ou d'échec dans les examens sont inscrits dans la loi ou décrits dans le curriculum, sans référence au CECR. Il est donc possible pour des étudiants de réussir un examen à un niveau donné du CECR sans atteindre un score qui indiquerait que ces étudiants ont une compétence à un niveau donné du CECR.

Pour un score de réussite/échec qui est relié à une procédure de définition des scores de réussite (pour les compétences de réception) ou une procédure de calibrage (pour les compétences de production). Lors de ces démarches, un groupe d'experts détermine quel score minimum est nécessaire pour les étudiants pour déclarer qu'ils ont atteint le niveau souhaité, dans le cas de tests de production orale ou de production écrite, les experts peuvent sélectionner des exemples de performances qui illustrent comment les étudiants doivent réaliser celles-ci pour être classés à un niveau spécifique du CECR.

Il est alors possible pour un étudiant d'avoir un score à l'examen qui indique deux choses : (1) l'étudiant a réussi ou non l'examen selon une perspective légale et (2) l'étudiant a ou n'a pas atteint le niveau souhaité du CECR.

Il est souvent dit dans les syllabus que l'examen est à un niveau donné du CECR. Cependant s'il n'y a pas eu de définition des scores de césure et de calibrage reliés au CECR, les scores de cet examen ne peuvent pas être considérés comme reliés au CECR.

Suggestions

- Réaménagez les décisions de succès/échec dans le syllabus dans les termes du CECR.

- Réalisez des procédures de définition des scores de césure (compétences de réception) et de calibrage (compétences de production).
- Regardez aussi les FAQs : Comment puis-je être sûr que la performance d'un étudiant pour la compréhension orale et la compréhension écrite est jugée au niveau souhaité du CECR ? et Comment déterminons-nous que la performance d'un étudiant en compréhension écrite est à un niveau spécifique du CECR ?

Références

- Conseil de l'Europe (2009). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) *Un Manuel* Strasbourg Conseil de l'Europe. Voir Chapitre 5 Formation à la standardisation et au calibrage et Chapitre 6 Procédures de définition des scores de césure.
- Nojons, José et Jana Bérésova, Gilles Breton, Gábor Szabó (2011). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) *Les points essentiels du Manuel*. Graz, CELV. Voir Chapitre 5 Formation à la standardisation et au calibrage et Chapitre 6 Procédures de définition des scores de césure.

- Conseil de l'Europe (2008). *Séminaire interlangues pour le calibrage de productions orales en allemand, anglais, espagnol, français et italien sur les 6 niveaux de l'échelle du CECR*. Strasbourg : Conseil de l'Europe.

7. Pourquoi-nous tester plus d'un niveau du CECR dans un seul test ?

En théorie c'est possible, en pratique, cela peut être compliqué. Cela dépend aussi du type de test et de la compétence qui est testée. Un test qui est relié au CECR est censé comporter un nombre représentatif de descripteurs du niveau du CECR visé. Pour chaque descripteur, à chaque niveau, nous aurions besoin d'un nombre suffisant d'items pour pouvoir porter un jugement valable sur la capacité de l'étudiant à faire ce qui est décrit dans le descripteur. En pratique cela voudrait dire que les tests de compréhension écrite et de compréhension orale seraient plus longs à réaliser.

Dans le cas des tests de production orale, il existe des formats comme Oral Proficiency Test (OPT - Test de Compétence Orale), dans lequel un interlocuteur formé passe d'un niveau à un autre en fonction de la compétence du candidat. Dans de tels tests il serait possible de savoir si l'étudiant est capable de fonctionner à plus d'un niveau du CECR. Il faut souligner ici que les interlocuteurs ont besoin d'être rigoureusement formés pour administrer ce type de test. D'une manière générale cela ne serait pas à la portée d'enseignants non formés.

Il existe des tests adaptatifs informatisés dans lesquels les étudiants sont confrontés à des items de différents niveaux, en fonction des réponses qu'ils donnent. En principe une tâche plus difficile (éventuellement à un niveau plus élevé du CECR) est présentée chaque fois que l'étudiant donne une réponse correcte. De cette façon la durée du test est considérablement réduite. Cela signifie que de nombreux items sont nécessaires pour créer une banque d'items. Toutefois la création d'une banque d'items calibrés est coûteuse et nécessite du temps.

Le CECR est construit sur l'idée qu'une personne qui peut réussir à un niveau donné du CECR peut aussi réussir à un / des niveau(x) en dessous du niveau donné. Une personne qui est au niveau B1 est supposée capable de réussir aussi bien aux niveaux A2 qu'A1. Cependant cela ne signifie pas que nous pouvons attribuer les niveaux A2 et A1 à un étudiant quand il ou elle a un score bas à un test de niveau B1 pour les raisons exposées plus haut.

Suggestions

- Quand vous testez les compétences de réception comme lire et écouter, visiez un seul niveau. Si les étudiants ont un score élevé dans un test à un seul niveau, donnez-leur un test à un niveau plus élevé (et l'inverse).
- Quand vous testez la production écrite, donnez aux étudiants une série de tests avec des tâches reliées aux descripteurs d'un seul niveau. Si les étudiants ont un score élevé à de tels tests d'un niveau, donnez-leur une série de tests à un niveau plus élevé (et l'inverse).
- Quand vous testez la production orale, formez des interlocuteurs à administrer des tests de type OPT ou procédez comme pour les tests de production écrite (voir ci-dessus).

Références

- ALTE (2011). **Manuel pour l'élaboration et la passation des tests et d'examen de langue**. A utiliser en liaison avec le CECR. Strasbourg. Conseil de l'Europe. Voir Chapitre 5. Correction, notation et délivrance des résultats et Annexe 7.

6. Peut-on relier nos tests avec les niveaux du CECR même si nous n'avons pas de statistiques sur les performances de nos étudiants dans nos tests excepté des scores totaux ?

En général la validité et la fiabilité d'un test seront accrues quand, sur la base de données des performances des étudiants pour le test, le test lui-même et/ou le format du test est adapté (difficulté des tâches, type de tâches, longueur du test, etc.). La première étape et la plus importante dans le processus de mise en relation est d'être sûr que le test en question est valide (c'est-à-dire qu'il teste effectivement ce qu'il prétend tester) et fiable (c'est à dire que le test est régulier). Idéalement ceci est fait en pré-testant les données collectées dans le test. Cependant il y a des situations pour lesquelles une telle procédure n'est pas possible ou très coûteuse comme dans les tests réalisés en classe.

Suggestions

Si ce n'est pas possible de recueillir des preuves qu'un test est valide et fiable par le biais des statistiques, nous pouvons néanmoins essayer de relier le test au CECR à travers la spécification. En fait la spécification est une phase dans le processus de mise en relation des examens au CECR qui doit toujours être réalisée.

La phase de spécification dans le processus de mise en relation aide les concepteurs de test à prendre conscience :

- de l'importance d'une bonne analyse de contenu des examens de langue pour garantir la validité du contenu ;
- du CECR et particulièrement de ses échelles de descripteurs ;
- de la logique pour relier les examens de langue à un cadre international comme le CECR et

Références

- des façons d'exploiter le CECR dans la planification et la description des examens de langue.
- Il y a quatre étapes à suivre dans la phase de spécification :
- s'assurer que ceux qui décrivent le test sont familiarisés de manière satisfaisante avec le CECR ;
- analyser le contenu du test en question en relation avec les catégories pertinentes du CECR (si une zone testée est non comprise dans le CECR, on demande à l'utilisateur de la décrire) ;
- profiler le test en relation avec les échelles de descripteurs correspondants du CECR sur la base de l'analyse de contenu ;
- faire une première proposition sur la base de l'analyse de contenu affirmant que l'examen ou le test en question est relié à un niveau particulier du CECR.
- Conseil de l'Europe (2009). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) [Un Manuel](#) Strasbourg Conseil de l'Europe. Annexe A, les sections A2 à A5 fournit un ensemble précieux de formes (A1 – A24) qui peut aider les concepteurs de test à spécifier un test dans les termes du CECR.
- ALTE (2001). Principles of good practice for ALTE examinations. Document de travail revu. Octobre 2001. www.alte.org

5. Quand un examen est relié au CECR, comment puis-je être sûr que les nouvelles versions de l'examen sont également reliées au CECR ?

Références

- ALTE (2011), **Manuel pour l'élaboration et la passation de tests d'examens de langue**. A utiliser en liaison avec le CECR. Strasbourg : Conseil de l'Europe. Voir Section 2.4., 2.5. et Annexe 7.
- Conseil de l'Europe (2009). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) **Un Manuel** Strasbourg Conseil de l'Europe. Voir Chapitre 7 Validation.
- Nojoms, José et Jana Béréšová, Gilles Breton, Gábor Szabó (2011). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) **Les points essentiels du Manuel**. Graz, CELV. Voir Chapitre 7 Validation.

La réponse la plus évidente à cette question serait d'appliquer toutes les étapes pour mettre en relation le nouvel examen avec le CECR : familiarisation, spécification, standardisation, définition des scores de césure et validation. Certaines étapes peuvent être réalisées plus facilement si les caractéristiques ont été répertoriées dans une matrice du test : ce qui est testé, comment cela est testé, le nombre d'items, les types d'items, les types de texte, etc. C'est une manière d'être sûr que le test mesure le même concept avec les mêmes compétences reliées au CECR que les versions précédentes.

Ce qui est idéalement nécessaire, c'est de réaliser un pré-test du nouvel examen avec une sélection représentative d'items intégrés du test initial (ce que l'on appelle des items ancres). Avec l'aide de statistiques avancées, ce serait alors possible de définir des scores de césure comparables à ceux de l'examen précédent. Si les scores de césure précédents ont été reliés au CECR, on pourrait faire valoir que le nouvel examen est relié au CECR. En fait cela fait partie du processus de validation pour relier un examen au CECR.

Suggestions

- Produisez une matrice de test relié au CECR, avec des références aux descripteurs et aux niveaux du CECR, comportant des spécifications.
- Collectez des données dans des pré-tests, en utilisant de préférence des items ancres pour relier les examens entre eux.
- Eventuellement adaptez les scores de césure si une nouvelle version de l'examen prouve qu'elle est plus facile ou plus difficile que l'ancienne.

Références

- Conseil de l'Europe (2001). **Cadre européen commun de référence pour les langues** : apprendre, enseigner, évaluer. Didier. Voir Section 4.1. Le contexte d'utilisation de la langue et Tableau 5 Contexte externe d'usage : catégories situationnelles.
- ALTE (2011). **Manuel pour l'élaboration et la passation des tests d'examens de langue**. A utiliser en liaison avec le CECR. Strasbourg. Conseil de l'Europe. Voir Section 1.4. Ethique et équité et Section 3.4.1. La vérification des nouveaux items. Voir également Annexe 4 Conseil aux rédacteurs d'items.
- Quand vous donnez un contexte, gardez présent à l'esprit qu'il y a des thématiques inappropriées comme la guerre, la politique, le racisme (y compris les clichés culturels et les stéréotypes), le sexe et le sexisme (y compris les stéréotypes), les sujets potentiellement pénibles (par exemple la mort, la maladie en phase terminale, les graves problèmes familiaux ou sociaux, les catastrophes naturelles et l'objet de phobies courantes comme les araignées et les serpents, sujets qui pourraient être désagréables à traiter), les examens, la réussite et l'échec, les drogues.

4. Est-il nécessaire de fournir un contexte pour les tâches ?

Dans le cas des tâches de production orale et de production écrite, la question peut sembler rhétorique. Dans la vie réelle, la manière dont nous parlons et dont nous écrivons dépend beaucoup des circonstances dans lesquelles nous sommes placés. Donc si les tâches de production orale et écrite doivent être authentiques, nous ne pouvons généralement pas le faire sans fournir un contexte. De manière générale la majorité des tâches de production orale et de production écrite sont par conséquent mises en contexte.

En théorie, il est possible de demander à un étudiant de parler ou d'écrire sur un sujet sans lui donner de contexte. C'est souvent le cas lorsque les étudiants doivent donner leur opinion (personnelle) sur un sujet, un phénomène ou un événement. Cependant du point de vue du CECR, cela peut être discutable : dans notre communication avec les autres nous avons besoin de penser à qui nous nous adressons et pourquoi. Donner simplement notre/une opinion sans penser à la personne à qui nous nous adressons peut être plutôt contreproductif, cela peut heurter les sentiments d'une autre personne ou peut simplement ne pas être compris ou en fait être mal compris. Nous devons aussi réaliser que le but de l'évaluation de la production orale ou de la production écrite n'est pas de tester la capacité de l'étudiant à exprimer un point de vue ou une opinion mais plutôt d'évaluer si l'étudiant peut exprimer une opinion dans la langue étrangère. En d'autres termes, nous avons besoin d'évaluer si l'étudiant peut exprimer un point de vue ou une opinion mais le contenu du message (par exemple faits, données, etc.).

Dans le cas de tests de compréhension écrite et de compréhension orale, nous constatons souvent qu'il est demandé aux

étudiants de lire un texte ou d'écouter un extrait sans aucun contexte donné (« Lisez le texte suivant et répondez aux questions »). On ne peut pas nier que dans ces cas on peut tester la lecture ou l'écoute. Cependant du point de vue du CECR, un contexte est attendu quand on teste la compréhension écrite et la compréhension orale : nous avons besoin de donner une raison de lire le texte.

Suggestions

- Pour tester une interaction orale ou écrite, il est conseillé de fournir aux étudiants un contexte qui soit réaliste et en relation avec leur âge et leur expérience dans la vie. Nous devons aussi nous rappeler qu'il y a des différences culturelles entre les locuteurs et les scripteurs qui ont différentes origines et expériences. Les étudiants peuvent être ou non à l'aise pour dire ou écrire certaines choses.
- Pour tester la production orale, il est également conseillé de donner un contexte basé sur un texte écrit ou sur un support visuel (image, photos, etc.). Cela peut aider les étudiants à élaborer un point de vue ou une opinion. Pour éviter de tester les opinions ou les points de vue ou l'imagination ou le bagage culturel des étudiants plutôt que leur capacité à exprimer cela, il est recommandé de leur fournir des mots-clés et des arguments pour ou contre.
- Pour les tests de compréhension écrite ou orale, en fournissant aux étudiants des textes authentiques placés dans un contexte réaliste et en donnant aux étudiants un objectif pour écouter ou lire, nous pouvons améliorer la validité des tests de compréhension écrite et orale.

3. Comment puis-je être sûr que la performance d'un étudiant pour la production orale et la production écrite est jugée au niveau souhaité du CECR ?

Pour évaluer la performance d'un étudiant pour la production orale ou la production écrite nous pouvons utiliser des échantillons témoins : des descriptions détaillées, validées de niveaux particuliers de performances d'un étudiant attendues à un niveau spécifique du CECR. De tels échantillons à calibrer prennent souvent la forme d'extraits du travail d'un étudiant (Ecrire) ou de vidéos (Parler).

Les échantillons à calibrer peuvent être produits de la manière suivante. Un coordinateur sélectionne des échantillons de performances d'étudiants à l'oral ou à l'écrit pour introduire un niveau spécifique du CECR. Pour chaque échantillon un groupe d'experts évalue et débattent pour savoir si l'échantillon illustre vraiment le niveau et pourquoi il ne correspond pas au niveau supérieur ou au niveau inférieur. Après réexamen, le groupe vote alors sur le niveau de chaque performance. Dans la dernière phase, les membres du groupe évaluent à nouveau la performance individuellement et comparent leurs résultats.

Il faut souligner ici que le calibrage est un processus de groupe plutôt que celui d'un expert qui montre et dit aux autres experts quelle sont les performances qui illustrent le mieux la performance à un niveau donné du CECR.

Suggestions

- Collectez un nombre important d'échantillons d'étudiants.
- Sélectionnez un panel d'experts (environ 10).
- Assurez-vous que les experts sont familiarisés avec le CECR et les descripteurs à travers des activités de familiarisation.

Références

- Collectez tous les matériaux pour la session de calibrage.
- Informez préalablement les panelistes des détails de la procédure de calibrage.
- Réalisez le processus de calibrage comme indiqué dans le Manuel Relier
- Publiez les résultats de la session de calibrage.

- Conseil de l'Europe (2009). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) **Un Manuel** Strasbourg Conseil de l'Europe. Voir Chapitre 5 Formation à la standardisation et au calibrage.
- Nojons, José et Jana Béréšová, Gilles Breton, Gábor Szabó (2011). Relier les examens de langue au Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer (CECR) **Les points essentiels du Manuel**. Graz, CELV. Voir Chapitre 5 Formation à la standardisation et au calibrage.
- Conseil de l'Europe (2008). **Séminaire interlangues pour le calibrage de productions orales en allemand, anglais, espagnol, français et italien sur les 6 niveaux de l'échelle du CECR**. Strasbourg : Conseil de l'Europe.

lecture et d'écoute ?

La sélection des échantillons pour l'écoute peut être problématique pour différentes raisons. Les matériaux authentiques peuvent être difficiles à obtenir ou coûteux, la qualité du son peut être inacceptable, le coût de production d'échantillons sonores variés peut être élevé. Il faut cependant éviter que la compréhension orale soit testée en utilisant des textes qui sont lus par un ou deux acteurs.

- Mentionnez toujours la source du texte. Indiquez quand il y a eu des changements (« Adapté de... »).
- Lorsque l'auteur n'est pas connu ou pas mentionné, il est conseillé de ne pas se référer à un auteur dans les items. Référez-vous au texte lui-même comme par exemple : Qu'est-ce que le texte dit à propos de... ?
- En reproduisant un texte, assurez-vous que le message original du texte est conservé.
- En reproduisant un texte, assurez-vous que le début et la fin sont clairs.

- ALTE (2011), **Manuel pour l'élaboration et la passation de tests et d'examens de langue. A utiliser en liaison avec le CECR**, Strasbourg : Conseil de l'Europe. Voir Section 1.1.3. Rendre le modèle opérationnel, Section 3.4.1. La commande, Section 3.4.1. La vérification des nouveaux items, Voir aussi : Annexe 4, Conseil aux rédacteurs d'items.

- Sélectionnez des échantillons pour l'écoute qui sont dans le domaine public, sur Internet (par exemple YouTube).
- Pour les tests d'écoute : si l'enregistrement de textes existants est nécessaire, utilisez des textes qui à l'origine étaient des documents oraux.
- Pour l'enregistrement de textes de compréhension orale : utilisez des locuteurs natifs pour produire les textes oraux.
- Essayez de produire des enregistrements semi-authentiques dans lesquels les locuteurs ne lisent pas les textes mais créent des énoncés spontanément à partir de repères.
- Essayez d'avoir un groupe varié d'acteurs avec différents accents pour produire des textes sélectionnés.
- Un certain degré de bruit de fond est acceptable dans les enregistrements. Dans la réalité, beaucoup d'activités d'écoute sont accompagnées d'un bruit de fond.
- Sélectionnez des types de textes qui cadrent avec les descripteurs spécifiques aux niveaux spécifiques du CECR.
- Et surtout choisissez des textes adaptés à l'âge des candidats.

1. Est-ce que les tests de différentes compétences peuvent être résumés à un niveau du CECR ?

Certains tests de langue déclarent mesurer les compétences d'un étudiant de langue à un ou plusieurs niveaux du CECR. On doit vérifier la validité d'une telle affirmation. Nous ne pouvons pas faire simplement la moyenne des performances dans les différentes compétences. Dans la vie réelle, la majorité des apprenants sont meilleurs dans une compétence plutôt que dans une autre et ceci, plus particulièrement dans les niveaux inférieurs du CECR. Ainsi dans un test nous devons être capables de faire la moyenne des points mais nous ne pouvons pas faire la moyenne des niveaux du CECR. Nous sommes capables de dire qu'un étudiant est au niveau B2 pour la compréhension écrite et au niveau A2 pour la production écrite. Nous ne pouvons donc pas dire que l'étudiant est au niveau B1 pour l'ensemble Lire Ecrire.

Le tableau 1 du CECR (Niveaux communs de référence : échelle globale) est souvent mal compris, comme s'il signifiait que ce qui est décrit pour un niveau en particulier est ce qui est attendu par un utilisateur de la langue pour l'ensemble des compétences langagières. Ce tableau doit toutefois être interprété comme une description de ce qu'une personne peut faire à un niveau donné et pour des compétences particulières : il ou elle peut fonctionner à ce niveau pour Lire et Ecrire et à un autre niveau pour Parler et Ecrire.

Le Conseil de l'Europe a préconisé l'élaboration de profils dans lesquels les compétences de l'étudiant dans les différentes compétences langagières sont décrites. Le Portfolio Langagier Européen a également adopté cette approche.

Références

- Quand vous élaborerez un test de langue, donnez une estimation des niveaux du CECR séparée pour chaque compétence langagière testée.
- Assurez-vous qu'il y a un nombre suffisant de tâches pour chaque compétence langagière pour être capable de faire une estimation valable du niveau du CECR.
- Encouragez les étudiants à découvrir leurs forces et leurs faiblesses pour chaque compétence langagière à travers des tests sur les compétences spécifiques.
- Notez les résultats de test de façon séparée pour chaque compétence.

Suggestions

- Conseil de l'Europe (2009) Analyse transversale des Profils des politiques linguistiques éducatives : tendances et points à considérer, Strasbourg : Conseil de l'Europe.
- Conseil de l'Europe (2000) **Portfolio européen des langues** (PEL), Strasbourg, Conseil de l'Europe.

Foire aux questions

1. Est-ce que les tests de différentes compétences peuvent être résumés à un niveau du CECR ?
2. Avons-nous besoin de proposer des textes authentiques de la vie réelle pour les tests de lecture et d'écoute ?
3. Comment puis-je être sûr que la performance d'un étudiant pour la production orale et la production écrite est jugée au niveau souhaité du CECR ?
4. Est-il nécessaire de fournir un contexte pour les tâches ?
5. Once an exam is linked to the CEFR, how can I make sure that new versions of Quand un examen est relié au CECR, comment puis-je être sûr que les nouvelles versions de l'examen sont également reliées au CECR ?
6. Peut-on relier nos tests avec les niveaux du CECR même si nous n'avons pas de statistiques sur les performances de nos étudiants dans nos tests excepté des scores totaux ?
7. Pouvons-nous tester plus d'un niveau du CECR dans un seul test ?
8. Que pouvons-nous faire pour être sûrs que nos décisions de réussite ou d'échec sont reliées au CECR ?
9. How do we determine that a student performance in reading or listening is at a Comment déterminons-nous que la performance d'un étudiant en compréhension écrite ou en compréhension orale est à un niveau spécifique du CECR ?
10. Est-ce qu'un examen qui teste la compétence linguistique (utilisation de la langue) peut être relié au CECR ?
11. Comment puis-je savoir si un test est au niveau du CECR auquel il déclare être ?
12. Donnons-nous le même nombre de points à chaque item dans les sous tests de compréhension écrite et de compréhension orale ?

Introduction

L'initiative intitulée « Relier les examens de langues aux niveaux européens communs de référence pour les compétences linguistiques » a été réalisée dans le cadre d'un accord de coopération entre le Centre européen pour les langues vivantes (CELV) et la Commission européenne. Elle propose des activités de formation et de conseil aux autorités éducatives dans les Etats membres de l'Union européenne et du CELV, visant à relier, de manière valide et équitable, les examens de langues aux niveaux de compétence définis dans le Cadre européen commun de référence pour les langues.

Ces questions fréquemment posées (FAQ) ont été recueillies par l'équipe d'experts de l'activité RELANG au cours de ses ateliers menés sur le thème « Relier les tests et les examens au CECR ». Les événements ont été organisés dans plus de 15 pays européens avec les parties intéressées telles que des formateurs d'enseignants, des auteurs de tests, des examinateurs, des concepteurs de programmes et des responsables ministériels. Les réponses aux questions fréquemment posées devraient être considérées comme un point de départ pour explorer des thèmes liés à l'alignement de tests et d'examens au CECR. Dans tous les domaines de l'éducation aux langues, il s'est avéré que le fait de relier les examens au CECR est une question relativement complexe. Ces FAQ et les réponses fournies ont été compilées en vue d'aider les parties intéressées à identifier les enjeux dans leur contexte local.

Equipe RELANG: José Nojóns (coordinateur), Evelynne Bérard, Jana Běrešová, Gábor Szabó

Pour en savoir plus sur cette initiative, visitez le site web <http://relang.ecml.at>

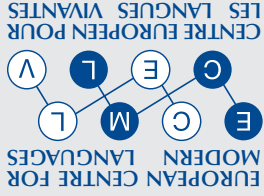
Les vues exprimées dans cet ouvrage sont de la responsabilité de l'auteur et ne reflètent pas nécessairement la ligne officielle du Conseil de l'Europe.

Tous droits réservés. Aucun extrait de cette publication ne peut être traduit, reproduit ou transmis, sous quelque forme et par quelque moyen que ce soit – électronique (CD-ROM, internet, etc.), mécanique, photocopie, enregistrément ou de toute autre manière – sans l'autorisation préalable écrite de la Direction de la communication (FR-67075 Strasbourg Cedex ou publishing@coe.int).

Editions du Conseil de l'Europe
FR-67075 Strasbourg Cedex
<http://book.coe.int>

Centre européen pour les langues vivantes du Conseil de l'Europe
Nikolaiplatz 4
AT-8020 Graz
www.ecml.at

© Conseil de l'Europe, décembre 2015



Compilation établie par
l'équipe RELANG
et éditée par José Noijons,
coordinateur de RELANG

Relier les examens de langues aux niveaux européens communs de référence pour les compétences linguistiques Foire aux questions

