



ILCEA

Revue de l'Institut des langues et cultures
d'Europe, Amérique, Afrique, Asie et Australie

19 | 2014

**Acceptabilité et transgression en langues et cultures
de spécialité**

The European sTANDEM Project for Certification in Medical English: Standards, Acceptability and Transgression(s)

*Le projet européen de certification en anglais médical sTANDEM : normes,
acceptabilité et transgression(s)*

Jean-Pierre Charpy and Didier Carnet



Electronic version

URL: <http://journals.openedition.org/ilcea/2475>

DOI: 10.4000/ilcea.2475

ISSN: 2101-0609

Publisher

UGA Éditions/Université Grenoble Alpes

Printed version

ISBN: 978-2-84310-280-6

ISSN: 1639-6073

Electronic reference

Jean-Pierre Charpy and Didier Carnet, « The European sTANDEM Project for Certification in Medical English: Standards, Acceptability and Transgression(s) », *ILCEA* [Online], 19 | 2014, Online since 27 June 2014, connection on 21 April 2019. URL : <http://journals.openedition.org/ilcea/2475> ; DOI : 10.4000/ilcea.2475

This text was automatically generated on 21 April 2019.

© ILCEA

The European sTANDEM Project for Certification in Medical English: Standards, Acceptability and Transgression(s)

Le projet européen de certification en anglais médical sTANDEM : normes, acceptabilité et transgression(s)

Jean-Pierre Charpy and Didier Carnet

- 1 Today, certification in languages at university level is based on internationally-recognised proficiency tests. These normalised tests cover general themes of language competence and culture, and thus make it possible to validate linguistic, communicative and cross-cultural skills in situations of communication that are not exclusively professional. All are closely linked to the CEFR (Common European Framework of Reference for Languages), whose normative descriptors are not always appropriate to meet the specific objectives of Languages for Specific Purposes (LSP).
- 2 In the current global professional environment, healthcare professionals from all over the world—whether researchers, hospital practitioners, doctors, pharmacists or nurses—are expected to have an effective command of English for medical purposes. Those involved in research, for example, are called upon to write research articles or give talks in English. Medical students, interns or specialty registrars may decide to do an elective, an internship or a fellowship in an English-speaking country. They may also have to use English in their home countries to interview, examine and treat foreign patients or to dispense drugs to clients, who will use English even though it is not their native language.
- 3 However, certification in medical English is still relatively unexplored despite the growing demand among healthcare professionals. This is why the sTANDEM (sTANdardised language Examination for Medical purposes) certification project came into being in November 2011. Financed by the European Union and piloted by a consortium of European teachers/researchers and international experts, the project is to

be completed by November 2014. The aim of the project is to standardise the requirements concerning skills in Languages for Specific Purposes (LSP) in medical settings all over Europe. Within the context of this EU-funded project, the testing of Languages for Medical Purposes (LMP) is a priority. Although languages such as French, German, Hungarian, Polish, and Romanian will also be represented and tested (at level B1 only), this article focuses solely on English for Medical Purposes (EMP) since English, which has become the *lingua franca* of science in general and medicine in particular, is at the core of the project.

- 4 The aim of this study is to present the evolution of the project with regard to the linguistic, didactic and cultural differences inherent in any work that involves partners of different nationalities, and to address the question of choice and acceptability in language certification in English.
- 5 Since at this stage of the project, actual exam results are not available, the authors have not been able to conduct a scientific analysis of the sTANDEM test or to address key docimological issues. Therefore, this study lies half way between theory and practice and will be based mainly on the groundwork of the sTANDEM partners and on the preliminary results and conclusions of the validation process.
- 6 After a presentation of the sTANDEM descriptors and specifications that distinguish medical certification from the normative context of generalist certification, the authors will address the question of selecting target terms for the various tasks linked to the five skills which are the cornerstone of the CEFR. They will then go on to discuss the difficulties associated with finding a common threshold of acceptability and the reluctance of certain partners to accept linguistic and didactic transgressions in the validation of the test takers' effective receptive and productive skills. Finally, with regard to the cultural dimension of this multinational project, sources of disagreement as well as cross-cultural enrichment will be highlighted.

1. International language proficiency tests: generalist certification vs profession-based certification

- 7 Generalist certification in English for non-native speakers at university level or in the workplace is based on internationally-recognised proficiency tests like the TOEFL (Test of English as a Foreign Language) and the TOEIC (Test of English for International Communication). In these generalist certification tests, the targeted language skills are tested in both general and specific situations and test takers must be able to communicate effectively in public as well as in personal contexts. In a recent article (2010), Powers evoked the objectives of such proficiency tests:

The TOEFL [is a] test, which colleges and universities use to gauge the language skills of prospective international students, and the TOEIC [is a] test, which employers in a variety of industries use to determine employees' readiness to use English in global communication.
- 8 With the exception of the TOEIC and the BULATS (Business Language Testing Service) for business English, the ILEC (International Legal English Certificate) for legal English and the ICFE (International Certificate in Financial English) for the English of finance, generalist tests such as the TOEFL and some of the Cambridge proficiency tests are not professionally oriented.

- 9 Conversely, profession-based certification rests on practices that require specific linguistic and cross-cultural skills. Their high specificity makes these tests much more accurate and reliable in the assessment of language skills for professional purposes. However, what makes them even more specific is the authenticity of the tasks they propose, as pointed out by Douglas (2000, p. 6): “The test task, in other words, must be authentic for it to represent a specific purpose field in any measurable way.”
- 10 With regard to the medical field, to date, certification in English for healthcare professionals is limited to a few national tests such as the OET (Occupational English Test) developed for Australian needs by the University of Melbourne, the CLES 3 (*Certificat de Compétences en Langues de l’Enseignement Supérieur*) developed in France, which may be regarded as a fourth-generation language certification in EMP (Tardieu, 2013), and the Hungarian PROFEX test (Proficiency Examination of English for medical purposes), a language proficiency test for medical students, doctors and other healthcare providers accredited by the Hungarian Board of Language examinations.
- 11 This field is still relatively unexplored on the international level despite the growing demand among healthcare professionals. The Canadian healthcare system, for example, is in great need of professionals, and the Australian healthcare worker market attracts more and more young European doctors. Indeed, professional mobility has become an international reality and language abilities of migrant healthcare professionals need to be assessed. This is all the more urgent as, in the not too distant future, there will inevitably be a shortage of healthcare professionals in Europe leading to the economic migration of such categories of workers, as Kofler (2010) points out:
- Health professions make up one of the largest segments in the overall European labour market, accounting for ten percent of the total workforce. The EU Commission calculates dramatic shortages in healthcare provision in the next decade unless countermeasures are taken now.
- 12 In view of this situation, the ensuing need for standardised EMP testing in European countries becomes obvious, as highlighted in an article accessible on the sTANDEM site:
- According to a recent EU whitepaper entitled AGENDA FOR NEW SKILLS AND JOBS (Strasbourg, 2010), there will be a shortage of more than one million health professionals in Europe by the year 2020. This challenge can only be overcome by promoting economic migration within EU member states and across language barriers. As there are more than 20 different official languages spoken in the European Union alone, enhanced mobility of health professionals will not be possible without taking action to improve key language skills. It is, above all, the ability to communicate professionally in the target language which is necessary to adjust to market needs. The above-mentioned whitepaper demands that “language requirements should be justified and proportionate, in view of the activity that the professional wishes to carry out”. Consequently, the language proficiency required for clinical practice, education and training (particularly in English) needs to be standardised to facilitate better cooperation between professional groups.
- 13 There is no doubt that English is now firmly established as the *lingua franca* of medical studies and practice, and that the command of English has become a key pre-requisite for doing clinical or theoretical research and taking part in international biomedical communication. From a practical viewpoint, bedside teaching and medical practice conducted in English are also essential within the context of international mobility since the quality of interactions in English with colleagues and patients is an obvious factor of successful and effective communication. In the specialised field of medicine, healthcare professionals have interlocutors with varying degrees of specialisation. Therefore,

communication between senior doctors, junior doctors, medical students and patients may fluctuate from EMP to English for General Purposes (EGP) with popularisation in between. Hence the urgent need to develop a widely-accepted proficiency test for healthcare professionals taking into account these discursive and cultural variations within the context of a communicative approach.

2. Certification in medical English: the sTANDEM project

- 14 The sTANDEM project was launched to develop an international examination and certification system suitable for students and healthcare professionals in the medical, pharmaceutical and nursing fields, whose native language is not English. The aim of the test was not to measure medical knowledge, but language skills for medical purposes. As indicated in the introduction to the Manual for sTANDEM Test Assessors: “While in general language tests, content knowledge is widely regarded as a factor distorting test results, in LSP testing it is a pre-requisite for successfully eliciting LSP performance.” That is why the language proficiency required for clinical practice, education, and training, particularly in English, needed to be standardised to facilitate mobility and improve communication and cooperation in the healthcare sector.
- 15 The European Commission granted financial support for the project under the Lifelong Learning programme (Key activity 2 – languages) for the years 2011–2014. The project consortium involves several leading centres for research and innovation in EMP teaching from Austria, France, Hungary, Poland, Romania, and the United Kingdom, as well as international experts from Japan, Malta, and the Netherlands.
- 16 The consortium partners decided to adopt a task-based, action-oriented approach that is based on the role of language users as socio-professional agents and on real-life situations within the context of actual communication (Bachman, 1990, 2007). They also opted for a holistic approach since sTANDEM certification does not depend on the validation of each of the five skills separately. As a matter of fact, a form of compensation was introduced to make sure that test takers are not penalised if they are weaker in any one skill.
- 17 The sTANDEM project comprises three distinct phases. The first one was a research-oriented phase. A detailed language needs analysis for medical purposes was carried out by sTANDEM partners in order to define, in the field of health care, the profiles of future test-takers according to their linguistic needs. More than 300 healthcare professionals, students and experts all over Europe filled in three different questionnaires. The statistical analysis of the various language needs led to the production of a Manual for Test Developers (2012) by the consortium members. This handbook is meant to serve as a standardisation document. It can be regarded as a common frame of reference in the process of test development and administration of sTANDEM tests. Its purpose is to make sure that the sTANDEM certification actually measures what it aims to measure (validity) and improves the consistency and the stability of these measures (reliability).
- 18 The second phase was a development phase in which 24 sets of examination papers were developed by partners working in tandems. These comprised reading papers (Romania + the UK), writing papers (Austria + Poland), listening papers (Austria + Poland) and speaking papers, including oral interaction (Romania + the UK). All these tests are in keeping with the communicative view of language and the model-based approach of the

CEFR. Four examination sets were developed at level B1, ten at level B2, and ten at level C1, and genre-based specifications and detailed descriptors¹ of language competence were defined to suit the specific needs of healthcare professionals. In keeping with the objectives of the European Commission, nineteen examinations sets were developed in English and five in other European languages (one in French, one in German, one in Hungarian, one in Polish and one in Romanian).

- 19 The exam papers drafted by the tandem partners went through a face validity investigation process piloted by the two Hungarian chief examiners before being validated by a Social Validation Board (SVB) piloted by the French partners. Before establishing examination centres in Europe, pre-testing sessions were organised in the consortium centres so as to obtain statistical data to confirm the construct validity of the tests. After validation by the Internal Validation Board (IVB), master copies were formatted, edited, and finalised.
- 20 The final phase was a testing phase during which two handbooks—one for the training of sTANDEM examiners (2013) and one to inform future sTANDEM test takers (2014)—were developed. Examiners are currently being trained, and standardised tests of English for medical purposes ought to be available all over Europe by the end of 2014.
- 21 Thanks to the sTANDEM site, the objectives of the project were made available to both individuals and institutions, and sTANDEM activities were disseminated worldwide. A number of medical schools and institutions have already offered their support and assistance. These include the Centers for Disease Control and Prevention (CDC), the Johns Hopkins School of Medicine, Myeloma UK, The National Institute of Health (NIH), the New England Journal of Medicine (NEJM), and the World Health Organization (WHO).
- 22 In the same way, a certain number of associated partners have agreed to contribute to the dissemination and promotion of the project: among them, GERAS (*Groupe d'Étude et de Recherche en Anglais de Spécialité*) in France, MedPharm Jobs in Poland, International Medical Publications in the Czech republic, the Ioannina Medical School in Greece, and the Virtual Medical Center in Australia.

3. A question of choice: descriptors and specifications for medical English

- 23 The CEFR and its accompanying guide (Milanovic, 2002) has become the gold standard for language certification because of its official acceptance throughout Europe, and even worldwide, since today most international proficiency tests rely on its descriptors and specifications (Bruderman, 2012). However, they apply mainly to general language and, as such, do not meet the needs of specialised language certification.
- 24 The generalist approach advocated by the Council of Europe is clearly worded in an online document entitled *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (2001, p. 177):

The Common European Framework provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc., across Europe. It describes in a comprehensive way what language learners have to learn to do in order to use a language for communication and what knowledge and skills they have to develop so as to be able to act effectively. The description also covers the cultural context in which language is set. The Framework also defines levels of

proficiency which allow learners' progress to be measured at each stage of learning and on a life-long basis.

- 25 The CEFR relies on an action-oriented approach and targets users and learners of English whose objective is to accomplish tasks in specific environments, but within a general field of action, even though it is true that descriptors at level C1 sometimes make reference to professional situations, as exemplified by the global scale of the Common Reference Levels (2001, p. 5): “Can use language flexibly and effectively for social, academic and professional purposes.”
- 26 In the Reading for Information and Argument Section for level C1, a similar reference to professional purposes can be found (2001, p. 11):
- Can understand in detail a wide range of lengthy, complex texts likely to be encountered in social, professional or academic life identifying finer points of detail, including attitudes and implied as well as stated opinions.
- 27 However, in the Self-assessment Grid for reading skills at level C1 (2001, p. 6), several genres are combined, which tends to show that the CEFR approach tends to be a generalist one:
- I can understand long and complex factual and literary texts, appreciating distinction of style. I can understand specialised articles and longer technical instructions, even when they do not relate to my field.
- 28 At levels B1 and B2, the rather vague notion implied by the term *field* is used repeatedly, as exemplified by the following descriptors:
- I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. (2001, Self-assessment Grid, level B2, Spoken Production, p. 6)
- Can write clear, detailed texts on a variety of subjects related to his field of interest, synthesising and evaluating information and arguments within his field of interest. (2001, Self-assessment Grid, level B2, Overall Written Production, p. 23)
- 29 The following descriptor at level B1 (2001, Self-assessment Grid, level B1, Overall Spoken Interaction, p. 12) highlights what can be perceived as a source of ambiguity: “Can communicate with some confidence on familiar routine and non-routine matters related to his/her interests and professional field.”
- 30 A specific descriptor at level B1 that concerns the medical domain (2002, *The European Language Portfolio*, Spoken Interaction, Appendix 16) is an indicator of the limits of this generalist communicative approach: “I can provide concrete information required in an interview/consultation (e.g. describe symptoms to a doctor), but with limited precision.”
- 31 In this example, the focus is clearly on the patient and not the healthcare professional. It is the generalist approach that is brought to the fore, and the specialised domain of the professional is simply overlooked.
- 32 In an article in which she showed that matching specific objectives and specialised tasks to the CEFR standards and descriptors is no easy matter, Marie-Hélène Fries (2009, p. 118) pointed out:
- In practice, however, making the mosaic of the specific aims that characterise English for special purposes coherent with the plethora of CEFR descriptors is not an easy task, especially as these descriptors are often either inexistent, or inappropriate.²
- 33 In view of these shortcomings, from the very beginning of the project, sTANDEM partners decided to re-write some of the CEFR descriptors in order to tailor them to the needs of healthcare professionals, which is precisely what the CEFR invites language professionals

to do (Foreword, 2001): “If you want to describe a specialised area, you may well need to sub-categorise further than the present certification goes.”

- 34 By way of illustration, here is a selection of examples taken from the Descriptors Recommended for sTANDEM at level B1 (sTANDEM site):

- Can understand factual information in familiar³ professional topics (Listening Comprehension Skills).
- Can understand important information in familiar professional topics (Reading Comprehension Skills).
- Can make short announcements concerning everyday professional topics (Speaking Skills).
- Can write short accounts of factual information in standard formats related to a specific genre (Writing Skills).

- 35 More specifically, it can be said that the professionally-oriented specifications that were developed at the sTANDEM initial meeting in Krakow (January 2012), and then re-written in the course of the project, reflect the EMP orientation of the sTANDEM professional proficiency test. The genres and the test types that are presented are all taken from the fields of medicine, pharmacy and nursing.

- 36 Concerning the text genres associated with listening comprehension skills, the text types are either medical narratives (level B1), medical narratives or descriptions (level B2), or medical argumentative texts (level C1). Below are some of the areas that are targeted:

- patient information, health reports, undergraduate lecturers (level B1);
- interviews with healthcare professionals, patients, and caregivers, medical news reports, case reports, graduate lectures (level B2);
- postgraduate medical lectures and talks, research reports, conference presentations, case studies (level C1).

The genres associated with reading skills systematically comprise short and long texts, as exemplified by the following examples:

- textbook extracts, product descriptions, information leaflets (level B1);
- official letters on professional topics, specialised case reports, referral letters, extended abstracts, application letters (level B2);
- research articles, letters to the editor, ethical approvals, editorials, informed consent statements, reading peer reviewers' comments (level C1).

As far as the productive skills are concerned, the targeted genres reflect typical professional settings. To test writing skills, two tasks have been developed. Task one bears on the following genres:

- job applications, invitations, congress applications (level B1);
- cover letters, referral letters, complaint letters (level B2);
- letters to the editor, case reports, answering reviewers' comments (level C1).

- 37 Task two consists in analysing a diagram and giving factual information (level B1), giving factual information and interpreting messages (level B2), giving factual information, interpreting and arguing (level C1).

- 38 To test speaking skills in an action-oriented context, three tasks have been developed. Task 1 is an introductory conversation about the job and/or research field of the candidate (all three levels). Task 2 is a simulation of a doctor-patient conversation focusing on history-taking (levels B1 and B2), and a simulated interview between a healthcare professional and a patient arguing and/or explaining (level C1). Task 3

consists in giving a talk on the basis of a diagram, table or graph (levels B1 and B2), and giving a short presentation on a professional issue (level C1).

4. Acceptability and transgression(s)

- 39 The adjective “transgressive” is defined as “an act involving the violation of moral or social boundaries” (Oxford Dictionaries.com). Thus, the general concept of transgression reflects the action of going beyond or overstepping some boundary, limit or rule. As such, it can be regarded as a form of deviance.
- 40 On the other hand, acceptability is defined as “the fact of being approved of, and considered normal by most people” (Collins Dictionary.com). Another definition of the term, although it refers to engineering, is closer to the linguistic and cultural scope of proficiency tests (Free Online Encyclopaedia.com): “State or condition of meeting minimal standards for use as applied to methods, equipment, or consumable products.”
- 41 Therefore, it can be said that transgression in specialised certification is the violation of linguistic, didactic and cultural standards accepted by the language professionals of the specialised domain. Conversely, acceptability can be regarded as the acceptance of minimal standards for use in professional situations.

4.1. Linguistic issues

- 42 From a lexical viewpoint, test developers, members of the Social Validation Board (SVB) and the two chief examiners sometimes had difficulty agreeing on the key words or phrases to be targeted. One of the bones of contention was the choice between exclusive answers that test developers had in mind when writing the papers and a wider range of acceptable alternative answers showing that test takers had managed to achieve the overall objective of the task. This goes to show that, in the context of medical English as the profession’s *lingua franca*, standards may vary. Choosing key words and phrases in a specialised proficiency test may be problematic, and certain partners were reluctant to accept linguistic transgressions even though such transgressions, in the form of non-standard English, appeared in the original sound track or reading document that test takers were given. This transgressive approach was meant to validate the actual listening or reading comprehension skills of future test takers.
- 43 Partners agreed that the terms and definitions that were to be selected should correspond to medically-based texts only, and that levels of language, the authenticity of the texts and the distinction between British and American English should be taken into account.
- 44 Consequently, in the listening comprehension papers at level C1, the informal term “preemies”, which is mostly used in informal American English, was not heard in the audio file. So, SVB members suggested that “premature babies” should be used instead.
- 45 Likewise, in a short-case clinical viva set in a British context (level C1), the term “doctors” was preferred to the term “physicians”, which is typical of American English. Indeed, in a British context, the term “physicians” would only be used to distinguish between “surgeons” and “physicians”.
- 46 In another listening comprehension paper at level B1, since the primary source was North American, the phrase “anytime soon”, which is typical of informal American English was

accepted by SVB members and not replaced by the more formal phrase “in the near future”, which is commonly used in British English.

- 47 Similarly, in an audio script at level B2, the segment “The way they determine as inappropriate was they said ...”, which may sound ungrammatical, was not modified. SVB members adopted an approach tolerant of transgression and decided that, although the English was very informal, it was authentic, and as such, had to be preserved.
- 48 Concerning variations in the same text (for instance: “Down syndrome”, “Down’s syndrome”, “Down” ...), it was agreed that these variations were acceptable since they came from a primary source. This is in keeping with sTANDEM specifications, which clearly state that authentic documents should not be modified, unless major grammatical or lexical mistakes are found in the text.
- 49 In the reading and listening comprehension papers, it was specified that all acronyms had to be clarified. In accordance with sTANDEM specifications and standardisation efforts, the full form of CDC was used (Centers for Disease Control), and the acronym TB was re-written accordingly (“Analyze the following diagram about tuberculosis (TB) ...”).
- 50 In the listening comprehension papers, levels of language were taken into account. Thus, the phrase “work-shy”, which was thought to be too difficult to understand at level B2, was replaced by “unwilling to work”. In the same way, the phrase “barking cough”, which does not correspond to level B1, was replaced by “bouts of coughing”.
- 51 In the speaking papers, the adjective “excruciating”, which is not in keeping with level B1, was replaced by the adjective “terrible”. Likewise, SVB members rejected the expression “hole in the heart”, which is not medically relevant, and they decided that the phrase “sickle cell” should be completed since “sickle cell anaemia” is the received medical form. On the other hand, the medical experts on the board confirmed that, although “neuron” is the usual spelling, the phrase “neurone disease” is medically and linguistically correct.
- 52 As far as the spelling of key words was concerned, test developers were asked to make sure that it was consistent. In one of the speaking papers at level C1, since the document was North American, American English spelling was used (“hemolysis” instead of “haemolysis”). Conversely, in a document of British origin, British English spelling was adopted (“bacteraemia” instead of “bacteremia” / “septicaemia” instead of “septicemia”).
- 53 In the writing papers, partners decided that, since the diagrams were taken from North American databanks or surveys, the verbs used in the test instructions should be spelled in the following way: “analyze ... and summarize ...” (instead of “analyse ... and summarise ...”). However, although the American spelling “analyze” was used systematically in the instructions, the decision was made not to penalise test takers if they used the British spelling (“analyse”) in their written productions. This is a form of acceptability that certain partners originally rejected.

4.2. Didactic issues

- 54 From a didactic viewpoint, defining a common threshold of acceptability, especially in the fields of listening and reading skills, proved to be a complex undertaking.
- 55 One of the issues raised was that it was the test takers’ language skills that were to be tested and not their medical knowledge. In some reading papers, for instance, potential candidates could very well find the correct answers without reading the texts, on the

basis of their professional knowledge. In order to ensure that reading comprehension and not medical knowledge was tested, the items of several sets of papers had to be re-written.

- 56 Similarly, in some listening comprehension papers, SVB members pointed out to the test developers that, in several questions, all three answers in each item (including the two distractors) were medically correct or plausible. It was agreed that, since medical knowledge was not to be taken into account in the assessment of listening skills, only items that were actually heard in the recordings were to be regarded as correct in the answer justifications. Nevertheless, if three items were medically correct, but if a distractor was not medically plausible (for instance “triglycerides are a blocker of cholesterol”), it could be maintained if a closely-related word or phrase was heard in the recording (“triglycerides are a building block of cholesterol”).
- 57 It was likewise pointed out that several gaps could be filled without listening to the audio file, simply by inferring the missing word(s) from the grammatical context or by using medical knowledge to fill in the blanks. In gap-filling/sentence completion task types, test developers argued that this seemed difficult to avoid, but that, as long as less than 25% of the items could be answered without listening to the audio file, and as long as the targeted word(s) were medically oriented, the task was acceptable. This is another illustration of a form of acceptability.
- 58 Speed of speech in the listening comprehension papers was another subject of debate. In one of the tasks at level B1, for example, SVB members considered that the level of English was fairly simple on the whole, although the speed of speech (200 words per minute) was a bit fast. They went against the chief examiners’ recommendations, and insisted on the fact that the number of words per minute was not the only criterion that should be taken into account, and that clarity of speech and background noise were also key criteria.
- 59 One of the listening comprehension task types, namely MCQs at level C1, seemed to pose a problem since they were more difficult than other task types. The test developers argued that MCQs were often based on information-dense lectures which are, effectively, harder to understand than doctor-patient conversations since doctors tend to adapt their speech when they communicate with non-specialists, a form of popularisation of specialised language. It was agreed that, from a pedagogical viewpoint, it was preferable to ask the test takers to start working on the easier tasks and to do MCQs at the end of the listening comprehension tests and that, after pre-testing, it might be necessary to readjust this task.
- 60 Several SVB members also found multiple matching exercises were too difficult at level C1 due to the presence of two distractors. The test developers answered that these distractors, although they seemed to be a source of confusion, were meant to make sure that listening skills only were assessed. The main argument for maintaining two distractors was that test takers who were expected to match sentence fragments in order to make correct sentences, should not be in a position to do so only because of their good command of syntax.
- 61 Concerning the listening comprehension papers again, one of the SVB members found it strange that a space was provided to enable test takers to take notes while listening. Other SVB members approved of this initiative, arguing that test takers would thus feel more confident when working on the listening tasks. It was agreed that, in order to

perform this task successfully, test takers should try to make selective notes (using the “space for your notes” column) while listening, starting with Part 1 (first listening) before concentrating on Part 2 (second listening). With a view to informing potential test takers, it was also agreed that full tape scripts would be provided with detailed justifications of the answer choices online.

- 62 In the answer key to some of the listening comprehension papers (gap-filling exercises) at level C1, only one correct answer (the one heard in the audio file) was mentioned even though other terms could be accepted as synonyms or equivalents. Although the chief examiners were in favour of a single answer, SVB members believed that all the answers showing that the test taker had understood the meaning of the targeted term should be mentioned as acceptable answers in order to help examiners validate the actual listening skills of the candidates, as the following examples demonstrate:
- Besides “blood thinning agents”, “blood thinners”, “anticoagulants” and “thrombolytics” could be regarded as acceptable answers.
 - Besides “pulmonary embolus”, “pulmonary emboli” and “pulmonary embolism” were also possible answers.
 - In addition to “(an) embolic stroke(s)”, “cerebrovascular accident” and “CVA” could be accepted. However, “(a) stroke(s)”, which was not specific enough, was deemed to be unacceptable.
- 63 Providing the extended list of acceptable answers that test assessors should be able to use was a complex matter and consequently, it was decided that the final list would be determined after pre-testing and might even be adjusted during the first exam sessions.
- 64 Another question raised concerned the systematic use of audio scripts for examiners. Although the chief examiners claimed that item-by-item justifications were not necessary, SVB members confirmed that, for this task type (MCQs), such justifications were acceptable since they enabled test assessors to understand and justify choices that were right and those that were wrong. It was argued that this form of justification could also help potential test takers when doing the online sample tests (accessible since March 2014 on the sTANDEM site), a practice in keeping with the international standards of generalist proficiency tests.
- 65 Concerning the reading comprehension papers, the first pre-testing session showed that the length of the texts was inadequate. For example, at level B1, specifications defined the length of the short texts at 160 words. Partners therefore decided to increase the length of these texts to 270 words. It was also decided that at level C1, longer texts should be used (900–960 words), as opposed to the original length (810–900 words), so as to enable test developers to select longer fragments, even whole paragraphs, for the gap-filling task.

4.3. Cultural and cross-cultural issues

- 66 In a multinational European project, it was inevitable that there would be interferences between several national cultures and hence, naturally, several cross-cultural discrepancies were brought to the fore during the course of the project. However, it was predictable that contact between these cultures would help bridge potential cultural differences between partners.

- 67 In this respect, one of the positive aspects of being part of a European consortium was that partners with different cultural backgrounds, training, education, and research interests were able to communicate and create sets of tests thanks to regular Skype meetings. Although, in the early stages of the project, differences emerged between theoretical and docimological stances, teamwork enabled members of the consortium to reach a consensus, resulting in collaborative work, as the following examples illustrate:
- In one of the listening comprehension papers, the use of the legal phrase “power of attorney” constituted one of the missing fragments in the task. Questioned by some SVB members as not being medically-oriented, it was nevertheless validated by the British medical expert on the Board because it is the kind of legal expressions health professionals need to use in their professional environment in cases of malpractice.
 - In the writing papers, partners had conflicting views concerning the layout of a letter due to different cultural practices. It was eventually agreed that the sender’s address can be printed in the top right-hand corner if the sender is using headed paper, but the date should be on the left if the letter is justified to the left. The British expert also made it clear that, as is usual in letters of referral, the patient’s information should be inserted immediately after the opening phrase “Dear Dr Bell”, which is not always the case in the other countries represented.
 - In several of the reading papers, the original nursing situations were based on medical practice in the test developers’ own countries. For instance, it would appear that in Romania, it is the nurses who are on duty in various departments, such as the Department of Neurology or the Department of Obstetrics and Gynaecology. The test developers were asked to modify the situations as they are very improbable in most of the other countries represented.
- 68 Extending the sTANDEM test to non-European countries such as Australia, Cuba or China, where several institutions have already shown interest in this European project, raises the issue of their respective medical cultures, which will have to be taken into account in the pursuit of the project in an international context.

Conclusion

- 69 In the course of their careers, European healthcare professionals who are not native speakers of English will most probably have the opportunity of using English for specific purposes within the context of international mobility or in the workplace. If their main objective is to communicate in everyday life or even in simple workplace situations, generalist proficiency tests such as the TOEFL, the TOEIC or the CLES (levels B1 and B2) will no doubt serve their purposes. However, if their main goal is to communicate effectively in the international specialised community they belong to, generalist certification may not meet their specific needs.
- 70 In this study, the authors have tried to draw the line between standards for generalist certification and minimal standards for professional certification, hence the initial presentation of specific descriptors and more professionally-oriented specifications. They have presented the successive stages of this project and its evolution in the light of the linguistic, didactic and cross-cultural issues that arose in the course of test development and during the various validation stages.

- 71 After carefully considering the initial objectives and the various stages of the development process, the authors postulate that, like the CLES test, the sTANDEM test can be regarded as a “fourth-generation test” (Tardieu, 2013, p. 246):

To sum up, the CLES test belongs to a fourth generation of tests, which foster a learning mode that is close to real life and require language skills integrated into easily-identifiable cognitive and discursive learning processes likely to be developed before the administration of the test.⁴

- 72 The authors have tried to show that linguistic and didactic transgressions may be necessary, but that cultural transgressions are more difficult to accept. Even though it is not always easy to find a common threshold of acceptability in a project that involves partners of different nationalities, with different cultural and professional backgrounds, the collaboration of the European partners towards creating a common means of assessing EMP has been a source of mutual cross-cultural enrichment.

BIBLIOGRAPHY

BACHMAN Lyle F., *Fundamental Considerations in Language Testing*, Oxford, Oxford University Press, 1990.

—, “What is the construct? The dialectics of abilities and contexts in defining constructs in language assessment”, in J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner, and C. Doe (eds), *Language testing reconsidered*, Ottawa, University of Ottawa Press, 2007, pp. 41-72.

BRUDERMAN Cédric et al., “Le CERCL : un outil pour construire une politique des langues ? Retour d’expérience sur l’évaluation et la certification à l’université UPMC (2009-2011)”, *Les Cahiers de l’APLIUT*, vol. XXXI, no. 3, 2012, pp. 31-41.

COUNCIL OF EUROPE, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, 2001, <www.coe.int/t/dg4/linguistic/Source/Framework_en.pdf> (consulted 7 January 2014).

DOUGLAS Dan, *Assessing Languages for Specific Purposes*, Cambridge, Cambridge University Press, 2000.

FRIES-VERDEIL Marie-Hélène, “Mise en cohérence de l’anglais de spécialité et du CERCL en France : difficultés et enjeux”, *ASp*, no. 56, 2009, pp. 105-25.

KOFLER Birgit, EHFG Press Office, B&K Medien- und Kommunikationsberatung, “Possible shortage of up to two million health care workers by 2020 – EU taking action to prevent impending crisis in providing health care”, <<http://pr.euractiv.com/pr/possible-shortage-two-million-health-care-workers-2020-eu-taking-action-prevent-impending-crisis>> (consulted 12 January 2014).

MILANOVIC Michael, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Language examining and test development*, Strasbourg, Council of Europe Language Policy Division, 2002.

POWERS Donald E., *The Case for a Comprehensive, Four-Skills Assessment of English Language Proficiency (TOEIC Compendium)*, 2010.

TARDIEU Claire, “Testing et certification”, *RCDL*, vol. 10, no 2, 2013, pp. 237–51.

“Why sTANDEM is superior to general academic language proficiency testing”, <<http://www.standem.eu/wp-content/uploads/2013/01/Why-sTANDEM-is-superior-to-general-academic-language-proficiency-testing.pdf>> (consulted 7 January 2014).

Online Resources

Collins American Dictionary.com, <www.collinsdictionary.com> (consulted 11 December 2013).

Free Online Encyclopedia, <www.encyclopedia.com> (consulted 11 December 2013).

Oxford Dictionaries.com, <www.oxforddictionaries.com> (consulted 11 December 2013).

sTANDEM site, <www.standem.eu>.

sTANDEM descriptors, <www.standem.eu/wp-content/uploads/2012/07/sTANDEMdescriptors_03.08.2012.pdf>.

NOTES

1. The complete list of descriptors is available on the sTANDEM site.
2. *“En pratique, cependant, la mise en cohérence de la mosaïque d’objectifs spécifiques qui caractérise l’anglais de spécialité avec la multitude des descripteurs du CECRL n’est pas chose facile, les descripteurs se révélant souvent soit inexistantes, soit inadaptés.”*
3. The adjective “familiar” refers to the three professional environments (medicine, pharmacy, nursing) targeted by the test.
4. *“Pour résumer, le CLES appartient à une quatrième génération de test qui favorise un type d’apprentissage, proche de la vie réelle, sollicitant des compétences langagières intégrées à des processus cognitifs et discursifs repérables et susceptibles d’être développés en amont.”*

ABSTRACTS

In the global world of international communication, employers or institutions that wish to determine the language skills of non-native speakers of English usually turn to generalist proficiency tests such as the TOEFL (Test of English as a Foreign Language) or the slightly more business-oriented TOEIC (Test of English for International Communication). These tests enable test takers to certify that they have reached a certain level in skills needed in everyday life and, sometimes, workplace situations, but they do not assess language skills for specific purposes.

Aware of this vacuum, a consortium of European teachers/researchers and international experts in the domain of medical English undertook to explore the largely uncharted territory of certification in EMP (English for Medical Purposes). They launched the EU-funded sTANDEM (sTANDARDised language Examination for Medical purposes) project in 2011 with a view to evaluating and validating the specialised language skills of healthcare professionals in the fields of medicine, pharmacy and nursing. In order to provide a test tailored to the professional needs of the specialised community, consortium members revisited the normative descriptors of the

CEFR (Common European Framework of Reference for Languages) and developed specifications based on specialised genres and texts.

The aim of this article is to present the evolution of the sTANDEM project in the light of the standardised requirements of language certification. It addresses the question of linguistic, didactic and cultural choices and poses the problem of a common threshold of acceptability and of possible transgressions of conventional rules within the context of this European partnership.

Dans le monde de la communication globale internationale, les employeurs ou les institutions qui souhaitent connaître le niveau de compétence linguistique en anglais de locuteurs non natifs ont généralement recours à des tests de compétence généralistes tel que le TOEFL (Test of English as a Foreign Language) ou un test à orientation plus sensiblement commerciale comme le TOEIC (Test of English for International Communication). Ces tests permettent d'attester que les candidats ont atteint un certain niveau dans des compétences requises dans la vie de tous les jours et, parfois, sur leur lieu de travail, mais ils n'évaluent pas les compétences propres aux langues de spécialité.

Conscients de cette lacune, les membres d'un consortium d'enseignants-chercheurs européens et d'experts internationaux dans le domaine de l'anglais médical ont entrepris de défricher le territoire encore largement inexploré de la certification en anglais médical. Ils ont lancé en 2011 le projet sTANDEM (sTANDARDISED language Examination for Medical purposes), financé par l'Union européenne, dont l'objectif est d'évaluer et de valider les compétences langagières spécialisées des professionnels de santé dans les domaines de la médecine, de la pharmacie et des soins infirmiers. Afin de mettre à la disposition des professionnels de la communauté spécialisée un test sur mesure répondant à leurs besoins, les membres du consortium ont révisé les descripteurs normatifs du CECRL (Cadre européen commun de référence pour les langues) et ont défini des spécifications à partir de genres et de textes spécialisés.

Le but de cet article est de présenter l'évolution du projet sTANDEM à la lumière des conditions normatives requises caractérisant la certification en langue vivante. Il aborde la question des choix linguistiques, didactiques et culturels et pose le problème d'un seuil commun d'acceptabilité et de la transgression éventuelle des conventions dans le contexte de ce partenariat européen.

INDEX

Mots-clés: anglais de spécialité, anglais médical, évaluation de la langue, certification, CECRL (Cadre européen commun de référence pour les langues), standards, acceptabilité, transgression (s)

Keywords: English for Specific Purposes (ESP), English for Medical Purposes (EMP), language assessment, certification, Common European Framework of Reference for Languages (CEFR), standards, acceptability, transgression(s)

AUTHORS

JEAN-PIERRE CHARPY

University of Burgundy (France)

DIDIER CARNET

University of Burgundy (France)