

PROFESSIONAL EDUCATION

UDC 811.111:371.3(045)

Olena Petrashchuk**AVIATION ENGLISH TEST DEVELOPMENT**

National Aviation University
 Kosmonavta Komarova avenue 1, 03680, Kyiv, Ukraine
 E-mail: aamm@nau.edu.ua

Abstract. *The article highlights the English language test development for the purposes of English language proficiency assessment against the ICAO Rating scale. The procedure of the test construction is described on the example of the Test of English for Aviation Personnel developed in 2007 in Ukraine.*

Key words: Aviation English; ICAO Rating scale; proficiency language test; TEAP.

1. Statement of Purpose

The English language proficiency assessment in aviation is still an issue which requires more details on how this procedure should be provided correctly. It is crucial to develop a test which is a high-stake proficiency test aimed at measurement of English language proficiency in speaking and listening [Manual...2010].

To comply with the ICAO designated critical features of an appropriate testing system in the context of aviation language testing the Test of English for Aviation Personnel (TEAP) has been developed. The test development was planned and conducted in accordance with internationally recognized approaches and best test design practices [Alderson et al. 1995; Bachman, Palmer 1996; Lazarton 2002; Stern 1983].

Therefore the TEAP provides

- a) proficiency tests in Speaking and Listening;
- b) an assessment means based on the ICAO Rating Scale and ICAO Holistic Descriptors of operational language proficiency;
- c) assessments of speaking and listening proficiency in a context appropriate to aviation in both routine and unexpected or complicated situations;
- d) direct assessments of the full range of an individual's speaking and interactive abilities, and language use in a broader context than in the use of ICAO standardized phraseology alone [English...2009].

Thus, the objectives of the TEAP design and development have been defined as follows:

- to develop a valid, reliable and practical tool for assessment of a candidate's ability to operate in English in the context of aeronautical radiotelephony communications in international civil aviation;

- to develop a tool of assessment reflecting actual language use;
- to develop a tool of assessment acceptable to both pilots and controllers;
- to develop a tool of assessment of English language proficiency in the context of aviation in a manner acceptable to the Ukrainian civil aviation authorities, professional registration bodies and employers;
- to provide assessment mechanism appropriate to the high stake testing of English for aviation personnel for licensing purposes;
- to provide testing of English for aviation personnel and quality of testing service delivered to an international standard;
- to provide testing of English for aviation personnel that is fair to all candidates;
- to get a positive impact on aviation operational communications techniques and the content and delivery of Aviation English training courses and/or test preparation programmes [Alderson et al. 1995; Bachman, Palmer 1996; English...2009; Manual...2010; Lazarton 2002].

2. Review of Research Results

The TEAP has been designed and developed in the period of two large-scale training and testing projects carried out by the Language Training/Testing Company AEROLINGUA for the Ukrainian State Air Traffic Service Enterprise (UkSATSE) in 2005-2006 and 2007-2009.

The theoretical approach to the TEAP design and development is based on the ICAO recommendations [English...2009; Manual...2010; Davidson, Fulcher... 2007] as well as on the Bachman's view of test usefulness [Bachman, Palmer 1996].

According to Bachman's approach the process of useful test construction includes activities on test design, development, operationalization and trialling, administration and quality monitoring [Bachman, Palmer 1996]. Therefore during the design stage the TEAP design statement was proposed and further moderated on the panel of item writers, academicians as well as stakeholders.

The main focuses of the design stage were:

a) to decide on general parameters of the future test in accordance with ICAO requirements and best practices of the proficiency test development;

b) to draft test items/tasks and initially pilot them on a small sampling for purposes of further improvement of the test items to become useful for the whole test construction.

To meet the aforementioned objectives a lot of tasks have been accomplished. They were as follows:

At the design stage (2005):

– The test purpose and test constructs were identified as described above [Manual...2010: Ch. 6].

– The Target Language Use (TLU) domain and test item/task types were defined as described above [Manual...2010: Ch. 2, 3].

– The target population for whom the TEAP is intended was studied as described above (Manual... 2010: Ch. 1).

– A plan for evaluating the TEAP qualities (validity, reliability, practicality and impact) was developed [Manual...2010: Ch. 6].

– The instructions for candidates and test team were formulated and piloted for further refining.

– The sources for test materials were found and their suitability evaluated.

– The test team (item writers, test administrators, examiners/interlocutors, invigilators, raters) was selected and initial training of them was carried out

– An inventory of required and available resources needed for the TEAP development and operationalization was listed and appraised for further availability, allocation and administration.

– The test items for Listening component and tasks for Speaking component were constructed; pre-testing/initial piloting/trialing of individual test items/tasks was conducted on small samplings (groups of 10-12 people) for further improvement of the test items/tasks or/and their administration procedure; good test items and tasks were put into the electronic database/test item bank for further use in the TEAP construction.

– The audio materials (authentic radio telephony transmissions or simulated exchanges) were piloted on small samplings (groups of 10-12 people) for further refining of the instructions, general comprehensibility of audio files, evaluating appropriateness of the background 'white' noise, finding proper allotment of time for each transmission and its task, etc.

– The general format of oral interview, interlocutor's questions and number of the interview parts were designed and piloted on small samplings (groups of 7-10 volunteers of typical contingent) for further defining of the proper time allotment, most suitable questions for the warming-up part, deciding on optimum number of tasks, their structure and content, etc. For example, questions 'What do you think about future of air travel in Ukraine?' or 'What may international organizations do for aviation development in Ukraine?' were deleted from the draft list of test questions due to lack of the subject knowledge. During the design stage it became obvious that each category of aviation personnel (e.g., controllers, fixed-wing pilots, helicopter pilots) should be provided with an individual TEAP version to suit a specific subject matter environment.

At the development stage (2006-2007):

The TEAP development was based on initially piloted test items/tasks which were stored in a test item bank. During the development stage the TEAP blueprint was developed which made possible to identify all details of the future test. The TEAP was constructed as a whole test. All components of the test (parts, sections, tasks, items) were incorporated in a single test. Initially the TEAP was developed for air traffic controllers (Tower, Approach and En-Route) and fixed-wing pilots.

In addition many other activities were performed at the stage, namely:

For Listening Test

a) Instructions to the TEAP and individually to the Listening and Speaking Tests were written;

b) Instructions to each task under Listening Test were written;

c) Layout of the paper answer sheet was constructed;

d) Audio materials based on authentic radiotelephony exchanges were designed;

e) Speaker's narrations were recorded;

f) All rubrics, pauses, 'white' noise were incorporated into the Listening Test audio materials.

For Speaking Test

a) Frame of the oral interview was finalized and tried out for further analysis of the oral discourse regarding its measurability against the ICAO Rating Scale;

b) Cue cards for tasks in the interview Part 2 were tried out to refine its structure and content optimal to elicit intended language of the candidate;

c) Photos for visual prompts were found, checked and selected for further use in the interview Part 3.

Chronologically the TEAP design and development process can be divided into 5 phases:

Phase I (January – September 2005) – Starting test development.

At this phase the major activities were aimed at preparing some conditions initially needed for the TEAP development:

- Acknowledgment of the need of Ukrainian aviation industry in assessment of English language proficiency to meet the ICAO new language requirements by March 5, 2008.

- Analysis of test design/research methodologies.
- Conducting a preliminary research on test purpose and specifications, test takers' characteristics, target (aviation) language use, constructs to be measured, analyses of language needs, availability of resources needed for test development, etc.

- Selecting and initial training of item writers.
- Allocating of available resources (financial, material, technical and human).

Phase II (November 2005 – April 2006) – Test design.

In the second phase the major activities were aimed at providing test design, test item/task writing, pre-test piloting of new items and creation of test item/task bank:

- Developing the TEAP blueprint to give details on the whole test specifications and individual test item/task specifications.

- Finding and selecting relevant materials.
- Writing test items/tasks and their pre-test piloting.

- Developing test Listening component on the base of piloted items/tasks.

- Developing test Speaking component on the base of piloted tasks (oral interview scenarios).

- Training of test administrators/examiners/interlocutors/invigilators/markers/raters.

- Identifying approach and methodology of the TEAP quality evaluation.

- Developing plan of activities to evaluate test quality (validity, reliability, practicality and impact).

Phase III (May 2006 – May 2007) – Test quality provision.

In the third phase the major activities were aimed at the TEAP trialling/quality evaluation:

- Preparing test materials and equipment needed for test administration (printing answer sheet booklets, providing audio files on CDs, printing protocols, photos, tape-recorders and digital Dictaphones, pencils, etc.).

- Live small-scale testing for further validation.
- Managing resources available and needed for large-scale test delivery.

- Administrating live field testing (up to 1000 personnel in 7 regions of Ukraine).

- Analyzing of the test results to have evidences of the test validity, reliability, practicality.

- Refining the test items/tasks.

Phase IV (June 2007 – February 2008) – Test launch.

In the fourth phase the major activities were aimed at presenting the TEAP to the SAA and other stakeholders to prove the test suitability for Ukrainian ATCOs and flight crews for licensing purposes to meet new ICAO language requirements:

- Presenting the TEAP for the SAA panel of experts.

- Receiving national endorsement of the TEAP and official authorization to provide testing services for aviation personnel for licensing purposes.

- TEAP launch in October 2007 to assess English language proficiency of air traffic controllers of Ukraine.

- TEAP launch for flight crews of airlines in February 2008.

Phase V (September 2009 – present) – Regular TEAP administration, test quality monitoring and test quality maintenance.

- Providing regular scheduled testing for aviation industry in Ukraine.

- Monitoring of the TEAP quality.

- Conducting research on vital issues of test quality.

- Monitoring quality of test administration and rating procedures.

- Review of the test versions each six months.

- Annual upgrading of the test format on the base of the research results.

- Regular retraining of the test team staff according to schedules.

3. Purpose of the Work

The TEAP aligns to the stated test purpose through:

- a) the test format (Listening and Speaking components);
- b) the test items and tasks (no discrete items but real world interaction) which enable elicitation of the intended language to demonstrate English language performance ratable against both the ICAO Holistic scale and the ICAO Rating Scale;
- c) definition of constructs according to the ICAO Rating Scale language profiles to be measured;
- d) trained item writers, raters, examiners/interlocutors available to provide quality test administration and rating.

The objectives mentioned above have been reached through keeping to the best testing practices used internationally in testing both general English language (e.g., IELTS) and aviation English language (e.g., ELPAC, RELTA).

The main evidence of the TEAP aligning to the test purpose is the fact that since October 2007 the TEAP has been successfully administered for nearly two thousand aviation personnel for licensing purposes including those having been assessed again in 3 years.

Rating process is considered as an important component of the whole testing procedure. It goes along the interview conducting and test tasks as means of elicitation of intended language.

The main principles/provisions of the TEAP rating process are as follows:

- Initial and periodic trainings of raters are mandatory;
- Intra- and inter-reliability of raters is monitored on regular base;
- Blind double rating is applied to all speech samples;
- The speech samples are kept as digital recordings of the interviews;
- Raters do not know candidates the voices they hear;
- The obtained speech samples are assessed immediately after the test during a period from 3 hours up to 2 days depending on a total number of samples under the rating procedure;
- In case of disagreement between two raters the third rater is involved;
- In any case the chief examiner/rater with full responsibility makes the final decision on the level to be awarded.

In some cases (so called ‘bordering levels’) an interlocutor opinion may contribute to the final score.

4. Selection and training of raters

The raters are selected according to requirements to qualification of the raters. Normally the teachers of Aviation English are trained to be a rater. According to ICAO requirements [Manual...2010], all TEAP raters have University degree in English linguistics and experience in Aviation English teaching. In addition to raters’ training they are familiarized with the working places of a controller and a pilot in a real world environment.

Rater’s training program consists of 2 week initial course and 3 day refresher course.

The initial training is aimed at gaining initial knowledge about assessment, scoring, rating, types of scales and developing skills to use ICAO scale for assessment of recorded speech samples at different levels. The training is based on suitable samples of oral language performances at pre-operational level 3, operational level 4 and extended level 5. The recordings of some samples are available on paper as a transcript handout for visual support of the sample analysis.

A special focus of the training is to teach raters to differentiate ‘bordering performances’ which are difficult to rate precisely. The raters are taught to understand how features of a language performance under each individual level are different from another one, what “sometimes”, “often”, “rarely”, etc. in the descriptor correspond to. With this purpose the sample recordings are commented by the chief examiner who explains the reasons of the scores/levels awarded.

During the initial training raters are trained how to document the rating procedure. They practice to fill in a rater protocol with specific features (errors) of the speech sample.

Raters are familiarized with speech samples which have different degree of their measurability against ICAO scale.

The initial training is finalized with an exam: the rater must assess 21 sample (7 samples of each level) previously assessed by two experienced raters. The exam is passed if only one out of each 7 samples has a different score.

The refresher training is provided for raters on a semi-annually base or more often if needed. The aim of the refresher training is calibration of raters’ ratings.

During the refresher course the raters practice on rating of especially collected samples. The samples represent both typical individual levels and specific to some reasons cases. Normally afterwards the details of the cases are analyzed and discussed.

Experienced aviation specialists with English language competence at level 5 or higher are involved into rating process as subject matter experts.

Though they are not available for daily assessment they are provided with randomly selected samples at each level (3, 4, 5) for their feedback, comments and advice. The aviation specialists are supported through initial training on ICAO scale application. They are exposed to some linguistic and psycholinguistic knowledge, terminology and notions, e.g., “maintain interaction”, “strategies for clarification”, “interfere with ease of understanding”, “stretches of language”, “linguistic or situational complication”, etc.

5. Assessment procedure and maintenance of reliable rating

Within the proposed framework, oral tests consist of several different speech interactions, each of which encompasses a variety of contextual variables that affect its output.

These interactions are performed as simulations in order to tap unmonitored language, which approximates the most stable discourse style, namely, the vernacular.

The sample of speech interactions included on an oral test is randomly selected from specifications of those interactions most relevant to the specific context and purpose of the assessment. The language samples are then assessed by raters/assessors with rating scales, whose criteria vary according to the intended use of the test results.

Maintenance of the raters’ intra- and inter-reliability is provided through semi-annual training and assessment workshops.

With a purpose of effective and reliable assessment of oral speech samples a Rator’s Memo has been developed. It includes main features of the language construct specific for levels 3, 4 and 5 which should not be neglected by the raters when filling in a Rator’s Protocol/Assessment Sheet.

6. Rator’s Memo

Construct to be measured at operational language proficiency (Level 4):

- Ability to use plain English in aviation context accurately (pronunciation, vocabulary, grammar) though minor mistakes may occur especially when attempted complex language;

- Language ability to communicate fluently by managing interaction with an interactant (e.g., demonstrate comprehension responding appropriately, immediately, informatively);

- Language ability to produce stretches of language, using appropriate connectors and fillers;

- Language ability to use clarification strategies and paraphrase when needed;

- Language ability to communicate effectively in situations of unexpected change of topic (turn of events), linguistic or situational complication.

Construct to be measured at extended language proficiency (Level 5):

- Ability to use plain English in aviation context accurately (pronunciation, vocabulary, grammar) with almost no influence by native language though minor mistakes may rarely happen when attempted complex language; vocabulary is sometimes idiomatic;

- Language ability to communicate fluently by managing interaction with an interactant (e.g., demonstrate comprehension responding appropriately, immediately, informatively in both familiar and unfamiliar or unexpected situations);

- Language ability to produce stretches of language, using appropriate connectors, fillers, distracters;

- Language ability to use clarification strategies and paraphrase when needed;

- Language ability to communicate effectively in situations of unexpected change of topic (turn of events), linguistic or situational complication.

Construct to be measured at pre-operational language proficiency (Level 3):

- Disability to provide language accuracy for effective communication due to influence by native language leading to weak control of basic grammar structures, limited vocabulary; no attempt to paraphrase when lack of vocabulary;

- Disability to maintain interaction consistently

- Disability to communicate fluently due to unnatural pauses of silence when lack of vocabulary and pauses of hesitation when not sure about choice of the words; attempts to transfer to native language;

- Disability to produce good consistent stretches of language due to inappropriate phrasing and pausing;

- Disability to prevent misunderstanding by using clarification strategies; instead of it attempt to translate in native language to check appropriacy of comprehension;

- Disability to use discourse connectors or fillers;

- Responses often are formulated on the base of a question words;

- Disability to maintain effective interaction in unpredictable situations.

7. Conclusion

The development of the test of English for aviation is a complicated and crucial task which requires professional knowledge of test design and construction. The language test developed as a proficiency high stake test should meet requirements internationally recognized and implemented as good practices. In aviation the test development also should suit specific requirements of the aviation field.

References

Alderson, C.J.; Clapham, C.; Wall, D. 1995. *Language Test Construction and Evaluation*. Cambridge University Press. 310 p.

Bachman, L.; Palmer, A. 1996. *Language Testing in Practice*. Oxford University Press. 360 p.

Davidson, F.; Fulcher, G. 2007. *Language Testing and Assessment: an advanced resource book*. New York, Routledge. 148 p.

English Language Testing Criteria for Global Harmonization. ICAO Cir 318-AN/180: ICAO Second Edition. Montreal, 2009. 96 p.

Lazarton, A. 2002. *A Qualitative Approach to Validation of Oral Language Tests. Studies in Language Testing 14*. UCLES: Cambridge University Press. 192 p.

Manual on the Implementation of ICAO Language Proficiency Requirements. ICAO Doc 9835 AN/453: ICAO Second Edition. Montreal, 2010. 186 p.

Stern, H. 1983. *Fundamental Concepts of Language Teaching*. Oxford, Oxford University Press. 298 p.

Received 8 October 2013.

О.П. Петрашук. Розроблення тесту з авіаційної англійської мови

Національний авіаційний університет, просп. Космонавта Комарова, 1, Київ, Україна, 03680

E-mail: aamm@nau.edu.ua

Висвітлено розроблення тесту з авіаційної англійської мови з метою оцінювання володіння англійською мовою відповідно до шкали ІКАО. Процедура побудови тесту описано на прикладі «Тесту з англійської мови для авіаційного персоналу», розробленого в 2007 р. в Україні. Наведено цілі проектування й розроблення тесту. Процедуру оцінки результатів розглянуто як важливу складову всієї системи тестування. Визначено основні принципи реалізації рейтингового процесу результатів тестування авіаційного персоналу.

Ключові слова: авіаційна англійська мова; визначення рівня володіння англійською мовою авіаційним персоналом; критеріально-рейтингова шкала ІКАО; тест з англійської мови для авіаційного персоналу.

Е.П. Петрашук. Разработка теста по авиационному английскому языку

Национальный авиационный университет, просп. Космонавта Комарова, 1, Киев, Украина, 03680

E-mail: aamm@nau.edu.ua

Рассмотрена проблема разработки теста по авиационному английскому языку для оценки владения английским языком в соответствии со шкалой ИКАО. Процедура построения теста описана на примере «Теста по английскому языку для авиационного персонала», разработанного в 2007 г. на Украине. Приведены цели проектирования и разработки теста. Установлена процедура оценки результатов как важная составляющая всей системы тестирования. Определены основные принципы реализации рейтингового процесса результатов тестирования авиационного персонала.

Ключевые слова: авиационный английский язык; критеріально-рейтинговая шкала ИКАО; определение уровня владения языком авиационным персоналом; тест по английскому языку для авиационного персонала.

Petrashchuk Olena. Doctor of Pedagogical Sciences. Professor.

The Head of Aviation English Department, National Aviation University, Kyiv, Ukraine.

Education: Dnipropetrovsk State University, Dnipropetrovsk, Ukraine (1977).

Research area: problems of teaching English for specific purposes and test control.

Publications: 100.

E-mail: aamm@nau.edu.ua

Web site: www.flyhigh.nau.edu.ua