





Esta obra está bajo una [licencia](#)  
[Creative Commons Atribución-](#)  
[NoComercial-SinDerivadas 2.5 Perú](#)

Repositorio institucional PIRHUA – Universidad de Piura





























































So, when talking about best practices for language testing in aviation, we have to consider all these characteristics and test providers should also apply them based on good language testing principles and practices. For the case of this ICAO standard, they include the International Language Testing Association (ILTA) Code of Ethics for languages testers, as part of Document 9835 (Appendix D).

It is also advisable for a good practice in language test development to have all stakeholders participating in this process. In this particular case, we are talking about pilots, controllers, administrators, operational trainers, aviation language teachers and qualified linguists.

As it is also mentioned at ICAO Document 9835<sup>8</sup>:

“Other essential elements of the test development process, with input from all stakeholders, include writing test specifications; deciding test method and content; developing test items; trialing the items; analyzing the results; revising test items; re-trialing the test and test items; validating the test; establishing a rating procedure; establishing a rater training process, and establishing record-keeping administrative functions.”

And what is also very important for this kind of testing is the performance of testers and raters. According to this ICAO standard<sup>9</sup>: “Best practice in language proficiency assessment call for at least two trained and calibrated raters, at least one of whom is a language teacher.”

As requested by ICAO in Circular 318<sup>10</sup>, raters should also complete recurrent training, at least once every year, mentioning

---

<sup>8</sup> ICAO Document 9835: *Manual on the Implementation of ICAO Language Proficiency Requirements*, pp. 6-4.

<sup>9</sup> ICAO Document 9835: *Manual on the Implementation of ICAO Language Proficiency Requirements*, pp. 6-5.

<sup>10</sup> *International Civil Aviation Organization* (2008). Circular 318: Language Testing Criteria for Global Harmonization, p. 37.

that it is important because: “Initial and recurrent training aiming to standardize rater behavior is vital to objectivity. As a language testing standard, raters should undergo approximately 40 hours of initial rater training and 24 to 40 hours of recurrent training per year.”

## **2.7. Areas of English competency for aviation personnel**

When talking about safe aeronautical radio communications, we should identify three areas where pilots and ATCOs need to demonstrate competency when using the English language: air traffic control (ATC) phraseology, English for Specific Purposes (ESP), and English for General Purposes (EGP).

Each of them plays a very important role in the performance of these persons and they are also very important variables when selecting training and testing strategies.

### **2.7.1. ATC Phraseology**

The aeronautical phraseology used for radio-communications between pilots and air traffic controllers is taught in flight training schools or civil aviation training centers (CATC) at the basic schooling of these professionals.

It is used as the main way of communication around the world and to be able to use it effectively it is necessary to have a good knowledge of the aeronautical procedures that they refer to because their meanings are not similar to the ones used in general language scenarios.

As it is mentioned by Marjo Mitsutomi<sup>11</sup>, from the University of Redlands, in USA:

“The phrases used in radiotelephony are designed to make the communicative function between the ground and aircraft as

---

<sup>11</sup> Mitsutomi, Marjo (2004): “Fundamental aviation language issues addressed by new proficiency requirements”, *ICAO Journal* 59, Number 1, p. 7.

concise and brief as possible, with the emphasis on accurate content as opposed to linguistic form.”

“Typically, grammatical markers, such as determiners (“the” or “a”) and auxiliary verbs are deleted, a feature making ATC communications markedly different from natural language.”

As an example of the difference between aeronautical English phraseology and general language, we can show the following typical ATC clearance<sup>12</sup>:

“American Airlines 081 cleared for take off, runway 15, climb and maintain flight level 350, cross Lima (LIM) at or above flight level 050. Contact Lima approach control on frequency 121, 5. Squaw code 5711.”

The aeronautical English phraseology normally works well and from my experience I can give testimony of pilots and ATCOs that worked for many years at the air traffic services without having any linguistic problem, just by using the ATC phraseology accurately, because they have memorized it.

But, when communications required working with unusual situations or plain language, they were unable to understand anything and that is why this ICAO standard reinforces the necessity to have a good command of the plain English language, in an aviation context. As it is mentioned by Mathews (2001:26): “the need for closer conformity to standard phraseology and for greater care in communication on the part of native and non-native speakers alike becomes readily apparent.”

It is also important to mention that according to ICAO standards, when pilots and controllers can not communicate using the language of the ground station, they must use the English language, and that is why English is considered the language of aviation worldwide.

---

<sup>12</sup> Example taken from the own experience of the author as an Air Traffic Controller.



To refer to standardized ICAO Aeronautical English Phraseology it is necessary to review the Chapter XII of ICAO Document 4444: Air Traffic Management.<sup>13</sup>

### **2.7.2. English for Specific Purposes in an Aviation Context**

This area refers to the additional knowledge of vocabulary that pilots and controllers should have in order to manage some other non-routine communicative situations.

These expressions and vocabulary are closely related to their daily duties, although they are not official ATC phraseology. They are related to some other aviation fields like meteorology, flight dispatch, ground services, radio navigation aids, urgency and emergency situations, etc.

So, aviation personnel should be also competent in aviation related ESP, which is something that goes beyond routine ATC phraseology.

An example of this aviation ESP could be as follows<sup>14</sup>:

“N1234X the runway is blocked by cows at the moment. It is not advisable to land now. We will try to clear the runway as soon as possible. Confirm intentions.”

It is an unusual situation and it could be a dangerous one if the pilot does not receive this information at the right moment and occasionally there have been cases when the air traffic controllers did not remember the name of the animals crossing or blocking the active runway, even when it is supposed to be a basic part of their English vocabulary.

To avoid a “linguistic stall”, Mitsutomi (1999:351) says: “The pilot-controller dialogue requires language readiness that goes beyond the current assumption that ATC phraseology is sufficient. It is not.”

---

<sup>13</sup> International Civil Aviation Organization (2007). “Document 4444: Air Traffic Management”. Montreal.

<sup>14</sup> Example taken from the own experience of the author as an Air Traffic Controller.

### 2.7.3. English for General Purposes

Now it is when the issue of EGP appears, because when pilots and air traffic controllers have a good command of the general English language, they are able to “negotiate for meaning” at any circumstances and that is another very important goal of the ICAO language proficiency requirements.

The problem is that there is not a worldwide recognition of what is the minimum level of general English language necessary for successful aeronautical radio-communications at all possible scenarios.

As it is mentioned by Mitsutomi in her previously mentioned article<sup>15</sup>:

“The ability to communicate when there is no prescribed script (i.e. ATC phraseology) is critical to safety. In practice this means that pilots and air traffic controllers must be able to achieve mutual understanding through the use of plain or general language to get their messages heard and understood.

It is precisely this issue of plain or general language use that has been problematic in the aviation context. Although strict adherence to phraseologies is always preferred, situations arise for which there is no adequate ATC phrase, or the phrase needs to be expanded with real-time information.”

It is a very important issue nowadays. For example, Bozena Slawinska, who is one of the Vice-Presidents of the International Civil Aviation English Association (ICAEA) wrote on the ICAEA forum, on April 24<sup>th</sup>, 2010: “During the ILTA 32<sup>nd</sup> Language Testing Research Colloquium, Professor Dan Douglas reminded us of the important fact, which some current academic research and indeed some tests available have still not grasped, i.e. that “plain language” of ICAO documents is not general English, but non-formulaic language used in an operational context.”

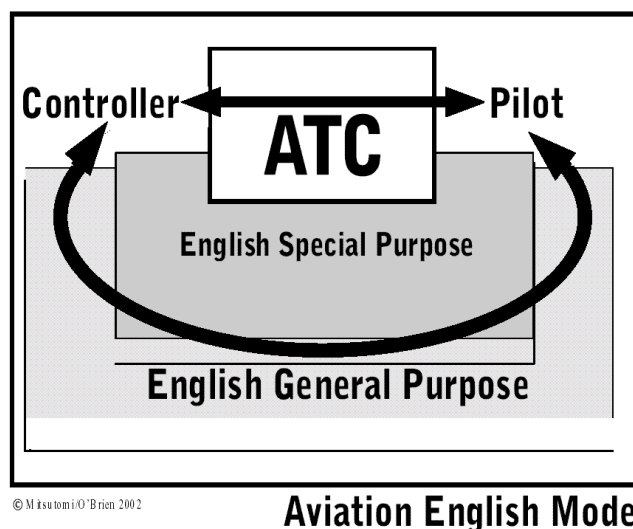
---

<sup>15</sup> Mitsutomi, Marjo (2004): “Fundamental aviation language issues addressed by new proficiency requirements”, *ICAO Journal* 59, Number 1, p. 9.

In my working experience, I prefer to agree with Marjo Mitsutomi<sup>16</sup> (2004:27) when finishing the above mentioned article with the following conclusion:

“Communicative competence in aviation English means that aviation personnel have common and standardized proficiency levels in the critical areas of highly specialized ATC phraseology, English for specific purposes as it applies to aviation, and the foundational general English. These three components together form the linguistic safety cushion that will significantly enhance safe communications in the aviation context worldwide.”

Mitsutomi concludes with the following graphic as to represent the *Aviation English Model* that summarizes what is above mentioned:



**Figure 1:** Aviation English model illustrates critical areas of English competency required for safe communications

(Source: ICAO Journal. Volume 59. Number 1, 2004)

---

<sup>16</sup> Mitsutomi, Marjo (2004): “Fundamental aviation language issues addressed by new proficiency requirements”, *ICAO Journal* 59, Number 1, p. 27.

This is a very important distinctive characteristic of the TOEFA examination, because it judges the competency of pilots and air traffic controllers in these three areas of language proficiency and I strongly think that this approach is more appropriate to strength safety in operational radio-communications.

I have also had the opportunity to follow up the operational performance of ATCOs that were evaluated following TOEFA procedures and they are able to “negotiate for meaning” even at unexpected situations, where General English also appears.

It is also something easy to confirm when supervising the performance of *ab-initio* ATCOs whose general English language proficiency has been a requisite to start their ATC studies, compared with those who only have a good command of ATC phraseology and ESP.

So, it is not only important to use valid and reliable tests, but to have enough evidence that these tests are able to predict future performance (wash forward effect).

## **2.8. Characteristics of the TOEFA Examination**

The main TOEFA characteristics are the following:

### **Objective**

The exam is designed to measure the English language proficiency level, in the abilities of speaking and understanding (at an agreed minimum level), of the aeronautical personnel that take part in the radiotelephony communications, with the purpose of contributing to the safety and regularity of the air traffic control services in the international environment in which they carry out the typical functions of their professions.

For this reason, there is a special emphasis on the use of the foreign language before the operational procedures, according to the holistic descriptors pointed out in the ICAO Rating Scale.

### **Linguistic descriptor 1: Pronunciation**

The evaluation is carried out by means of an interview, with open questions, so that the candidate demonstrates that he is able to make himself understood, with a dialect or accent intelligible to the aeronautical radio-communications. The pronunciation, rhythm and intonation are evaluated, as well as the grade of interference with the ease of understanding.

### **Linguistic descriptor 2: Structure**

Considering that the abilities of the language evaluated refer to the oral production of the English language, the practical application of the grammatical structures is evaluated, when the candidate answers to the questions posed by the language assessor(s), during the interview.

Consequently, the correct use of the pertinent grammatical structures is measured, as well as the structures of the sentences and the appropriate use of the functions of the language, according to the scenario in which they are used.

### **Linguistic descriptor 3: Vocabulary**

This descriptor is also evaluated through the interview. The extent and precision of the vocabulary used by the candidate is judged, with the purpose of communicating efficiently about the variety of familiar and unfamiliar topics that are used during the communication with the language tester. The capacity of the candidate to use the appropriate vocabulary (to manage successfully in unexpected circumstances) is also evaluated.

#### **Linguistic descriptor 4: Fluency**

The fluency of the candidate to communicate is evaluated through the oral interactions with the Interlocutor, as well as the dialogues that are generated after listening to the audio files.

The capacity of the candidate to make himself understood with detail and with natural fluency is also evaluated, as well as the stylistic effects, accent, and conjunctions used to achieve an effective communication, about familiar, unfamiliar or unexpected situations.

#### **Linguistic descriptor 5: Understanding**

The candidates listen to dialogues and real communications of native speakers of the English language, from audio files, and they explain to the language tester (using the foreign language) the scenario or situation that he/she has just listened to; with the purpose of measuring their grade of understanding of the language and the linguistic variants (dialects and accents) or tones that are intelligible for the international community of aeronautical users.

#### **Linguistic descriptor 6: Interactions**

The capacity of the candidate to interact with ease in unexpected situations is evaluated, as well as his or her ability in capturing verbal and non-verbal indications and to respond appropriately to them, by means of immediate, appropriate and informative answers that allow him or her to manage the speaker/receiver relationship efficiently; verifying, confirming or clarifying appropriately, when it is necessary.

The language tester presents situations and unexpected scenarios, through aviation photos, to verify the consistency and coherence of the answers of the candidate, according to the outlined scenario, what also allows the tester to confirm the candidate's capacity to understand and his or her ability to interact appropriately.

The *Tasks and Test Items Format* of the TOEFA examination can be found in **Appendix 3** of this document. Additional information about the test can be seen at its Internet site: [www.toefa.com](http://www.toefa.com) .

For scoring purposes, we should remember what is established in this international standard: “An individual must demonstrate proficiency at least at Level 4 in all categories in order to receive a Level 4 rating”. (Document 9835. Page 2-9. Item 2.8.4.: ICAO Rating Scale).

Regarding the issue of pronunciation, it is important to consider what is established as a satisfactory level (Operational level 4) for the skill of pronunciation at the ICAO rating scale: “Pronunciation, stress, rhythm and intonation are influenced by the first language or regional variation but only sometimes interfere with ease of understanding.”

## **2.9. Linguistic Considerations for TOEFA Construct**

For the construct of the TOEFA test, the following linguistic analysis and principles of language testing have been applied:

### **Competence -vs. - Performance**

The first point of discussion is about competence versus performance. As established by Chomsky: “We thus make a fundamental **distinction between competence** (the speaker-listener's knowledge of his language) **and performance** (the actual use of language in concrete situations). Only under the idealization set forth in the preceding paragraph (...) is performance a direct reflection of competence. In actual fact, it obviously could not directly reflect competence. A record of natural speech will show numerous false starts, deviations from rules, changes of plan in mid-course and so on.” (Chomsky 1965:4).

In this case, the test has been designed to judge the performance of the candidates instead of their competence and that is why it does not have discrete items about the knowledge of the language, because it is not the purpose of the test, as requested by the ICAO language proficiency requirements.

### **Usage – vs. - Use**

In this case, the test uses samples of language use (in an aviation scenario) instead of usage, making it possible to predict the future performance of candidates, according to the test results, what is called the *wash forward effect* of the test. For validation purposes, the test results were compared with the real performance of ATCOs on their job and the results were very good ones.

### **Direct – vs. - Indirect Assessment**

This is a direct test because it uses examples of performance as an indicator of communicative competence.

All the test tasks and items are directly related to aviation or work-related issues (the work of an air traffic controller) which are directly delivered to the candidate by the interlocutors/raters and also by using recordings obtained from real or simulated work situations.

It is also a direct test if we take into account the manner of delivery of the prompts and test items, which is done directly by the interlocutors/ raters. In the case of semi-direct tests, the prompts can be delivered via recordings through phones or computers.

The disadvantage of direct tests (like TOEFA) is their practicality, because it is time-consuming, since it has to be delivered individually to each candidate, in a face-to-face way, while semi-direct examinations can be delivered to groups and even using the internet, although it is not recommended for these kind of high stakes tests.

### **Discrete - vs. - Integrative Assessment**

Since it is a test aimed to judge the communicative abilities of candidates, we are talking about an Integrative Assessment, which uses integrative or global items, like speaking about the last



changes in the work scenario, describe an aviation picture, re-tell an aviation audio, etc.

In this case, the test-takers have to demonstrate their ability to use the language in appropriate real or simulated situations that reflect real scenarios, where they have to put their knowledge in use.

### **Objective - vs. - Subjective Assessment**

This test is subjectively assessed, because the assessors judge the performance of candidates according to their experience and criteria, and also considering the training received in the Initial ICAO Language Proficiency Raters Course and the recurrent versions that they study every year. The ratings are done following the linguistic descriptors of the ICAO Rating Scale.

To diminish this subjectivity, the test is scored by two assessors (when applied for licensing purposes), in order to have a better inter-rater consistency and reliability, as well as the intra-rater reliability that is assured because the two assessors have to comply with a very strict profile of competencies, as to be able to perform these very important functions; so, even when the test is scored at different times or with different raters, the overall results should be the same.

### **Receptive – vs. – Productive skills**

This test judges only the abilities of speaking (productive) and listening (receptive). However, since it is a direct face-to-face test, the listening ability is not scored through objective marking, but by judging the listening comprehension of the candidates through activities where they have to state orally what they have understood.

In the case of the productive skill of speaking, it is also measured through interactive activities. The entire interview is always recorded, allowing interlocutors/raters to listen to the language production of the candidates, as much as they need to listen to















































































































