12. AIR SAFETY, LANGUAGE ASSESSMENT POLICY, AND POLICY IMPLEMENTATION: THE CASE OF AVIATION ENGLISH

J. Charles Alderson

The language of international aviation communication is English, but numerous aviation incidents and accidents have involved miscommunication between pilots and air traffic controllers, many of whom are not native speakers of the language. In 2004 the International Civil Aviation Organization (ICAO) published a set of Language Proficiency Requirements and a Proficiency Rating Scale, and by 5 March 2008, air traffic controllers and pilots were required by the ICAO to have a certificate attesting to their proficiency in the language used for international aeronautical communication. Although some organizations made efforts to produce tests by the deadline, in the event an implementation period was allowed, with a new deadline of March 2011. This article describes a number of surveys of tests of aviation English, the implementation of the ICAO requirements, and the rating scales. It concludes that many of the assessment procedures appear not to meet international professional standards for language tests, the implementation of the language assessment policy is inadequate, and much more careful and close monitoring is needed of the quality of the tests and assessment procedures required by the policy.

Introduction

The language of international aviation communication, especially between pilots and air traffic controllers (ATCs) via radiotelephony, is English. Annex 10 to the Convention on International Civil Aviation, *Aeronautical Telecommunications*, in Volume II, *Communications Procedures*, and Chapter Five, *Aeronautical Mobile Service* (paragraph 5.2.1.1.1), recommends the following:

> In general, the air-ground radiotelephony communications should be conducted in the language normally used by the station on the ground.

> Pending the development and adoption of a more suitable form of speech for universal use in aeronautical radiotelephony

communications, the English language should be used as such and should be available, on request from any aircraft station unable to comply with 5.2.1.1.1, at all stations on the ground serving designated airports and routes used by international air services.

However, the English of international aviation is not English for general purposes or English for international purposes. Aviation English is a language for specific purposes (see Douglas, 2000), but it is even more restricted than that. Much of the English of aviation can be classified as a code that is used in a very restricted context (see Cabre, Freixa, Lorente, & Tebe, cited in Sarmento, 2005, p. 2), known as standard phraseology. It can be seen, following Ragan (2007, p. 54), as a highly restricted register associated with "distinctive probabilities of discourse functions and choice of lexis and grammar."

Written communication typically takes place through maintenance and operations manuals, produced by the airline manufacturers and airline operators. Both types of document are safety-critical, but especially the operations manual's abnormal and emergency checklists, which provide information on how to cope with nonnormal situations (Sarmento, 2005). The language used in maintenance documents is also often known as "Simplified English" (see Shawcross, 1993).

Radiotelephony communication takes place between pilots and air traffic controllers, with standard phraseology at the core, and operational exchanges in plain English when phraseology is inadequate; such radiotelephonic communication is used almost exclusively for air–ground communication, to direct, inform, question, request, and respond, where the air traffic controller directs and controls pilots. The focus of the communication is aircraft takeoff and landing, flight navigation, and so on, and the channel used is spoken, via radiotelephony.

Although the acoustic quality of radiotelephony is often poor, "routine air–ground communications are typically smooth and effortless. Misunderstandings are infrequent and, in general, they are easily resolved" (Mell, n.d., p. 1). This is largely because, over the years, standard phraseology has developed as an internationally recognized code. This has been refined in part as a result of the analysis of aviation incidents involving miscommunication, whereby ambiguities and confusions have been reduced to a minimum. This restricted code is used in highly predictable circumstances, and normal communications follow a prescribed sequence.

Mell (1992) found that about 50% of conversations with air traffic controllers are initiated by pilots and follow a fixed sequence. Mell gave an example that he analyzed as follows:

Paris. Good afternoon. Jetset 762.

Level 370.

On course Deauville

The controller, who is already in possession of the flight plan details of this aircraft, is expecting the call. In addition, the call follows a well-established script for such messages—that is to say, a greeting, followed by the current flight parameters (flight level and route). Each individual message turns up in a predetermined slot at a point in the sequence when the controller is expecting to hear it. In addition, sentences are extremely short (and correspondingly easy to process), while the words used belong to a very limited lexical set jointly determined by prescribed phraseology and operational practice. (p. 2)

Unfortunately, even in such predictable and restricted circumstances, miscommunications can and do occur, as a result of a number of factors. These can include pilots not realizing a communication is intended for them, interference on the radio frequency, overlapping calls, misunderstood flight parameters, incorrect readbacks, inadequate clarification of flight parameters, and so on (see Cushing, 1994). An obvious cause of misunderstanding of phraseology can be miscomprehension of pronunciation by one of the interlocutors, particularly if one or the other or both has a regional or nonnative accent in English.

One well-known incident, involving Air China 9–81 landing at New York's JFK airport (see Harry Chen's blog at http://harry.hchen1.com/2007/10/15/555) showed the Chinese pilot's English to be incomprehensible, and he also failed to understand the native-English-speaking air traffic controller. However, the latter also failed to use standard phraseology to communicate with the pilot and his lack of sensitivity to the Chinese pilot's problems reveals a degree of communicative incompetence.

Research into cognitive factors in air traffic communication, specifically the demands of concurrent tasks in cockpit operations (Farris, Trofimovich, Segalowitz, & Gatbonton, 2008; Loukopoulos, Dismukes, & Barshi, 2003) shows both the complexity of aviation communication settings, and the interaction of second language proficiency levels with workload on the accuracy of repeated messages, even in routine communications. Native speakers of English are not necessarily immune from poor communication, even in routine communication, as research using conversation analysis to examine cockpit conversations has shown (Helmreich, 1994; Helmreich, Klinect, & Wilhelm, 1999; Nevile, 2002, 2004; Nevile & Walker, 2005). We will come back to this issue later.

However, more serious miscommunications can occur in unpredictable situations, particularly in emergencies, and especially where urgent corrective action or essential information is required, and where one of the interlocutors may be under severe emotional stress. Prescribed procedures exist for emergency calls also, such that the first element of the call should be a distress signal, and the message should be composed of a number of specified elements, including the name of the ground station being called, the aircraft identifier, the nature of the problem, the pilot's

intention, and the flight parameters of the aircraft (Mell, n.d., p. 2). Nevertheless, and unsurprisingly, in emergency circumstances participants may ignore or fail to use phraseology and have recourse to what is known as plain (or natural) language.

Where the circumstances are unpredictable, where speed and clarity of communication are of the essence, and where accuracy of comprehension by both parties is crucial, the demands on the language proficiency of aircrew and air traffic controllers can be extreme. This is obviously a concern where one or more of the participants may not be a native speaker of English.

There are a number of well-documented cases where serious fatal accidents have occurred, where a lack of adequate communication between aviation personnel from different linguistic backgrounds has been a significant contributing factor. According to Ripley and Finch (2004), in three accidents alone, 1,006 people died at least in part because of language problems that gave rise to communication problems. In November 1996, a Kazakhstan Airline plane collided midair with a Saudi Arabian Boeing 747 over Charkhi Dadri, New Delhi, India, killing 351 people. The air traffic controller was Indian, and the pilots were Saudi and Russian.

Avianca Flight 052 crashed in April 1991 at Cove Neck, New York, en route from Bogota Colombia to JFK New York. The crew failed to communicate clearly to air traffic control that their aircraft was running out of fuel. The air traffic controller also failed to use communication strategies that might have clarified the actual situation on board the aircraft.

The most serious incident occurred at Tenerife airport in 1977, when a KLM Boeing 747 collided on take-off during foggy conditions with a Pan Am Boeing 747 that was still taxiing up the same runway and had not yet cleared that runway. In this crash, which involved misunderstanding of the phrase "at takeoff," 583 people lost their lives. The failure of the senior Dutch pilot of the KLM to recognize that messages between the English-speaking pilot of the Pan Am aircraft and the Spanish air traffic controller indicated that the runway was not yet clear. (See *YouTube* reenactment at http://www.youtube.com/watch?v=qa4dr6NoWJc and the flight safety database record at http://aviation-safety.net/database/record.php?id=19770327-1)

Recent research focusing on cross-cultural communication and misunderstandings also shows that a lack of linguistic proficiency may combine with cultural differences to cause miscommunication in aviation settings (see Merritt & Ratwatte, 2004; Stratechuk & Beneigh, 2004; Turney, 2004a, 2004b). This may be especially problematic with multicultural cockpit crews, where variables like silence, repair, turn-taking, and overlapping talk, which conversation analysis (Nevile & Walker, 2005) has shown to indicate problems of communication, may be very differently interpreted in different cultures, particularly in combination with differences in status or authority of participants, potentially giving rise to miscommunication. Clearly, the importance of proficiency in English in aviation communication cannot, and should not, be underestimated.

ICAO Language Proficiency Requirements (LPRs)

The International Civil Aviation Organization (ICAO), a branch of the United Nations, regulates aviation internationally. It establishes and reviews "international standards for the licensing of personnel and aircraft operation, and develops principles and techniques of air navigation, including meteorology, radio communication, and rules of the air" (Ragan, 2007, p. 54). As a result of an increasing awareness of the importance of the English language proficiency of pilots and air traffic controllers in the safety of aircraft, crew, and passengers, ICAO has developed a set of language proficiency requirements (LPRs). These set minimum standards for language proficiency for pilots and air traffic controllers and refer to an ICAO language proficiency scale, which is to be used both as the basis for test construction and for the rating of language proficiency. These LPRs set Operational Level 4 on the 6-point scale as the minimum level required for licensure, and at Level 4, license holders have to be retested; a period of 3 years is recommended. Once a candidate has achieved Level 6, that person licensed is to operate for life.

The ICAO Language Proficiency Rating Scale consists of six levels of skill in six areas of language use: pronunciation, structure, vocabulary, fluency, comprehension, and interactions. The definitions of Level 4 in each of these criteria are as follows:

Pronunciation (Assumes a dialect and/or accent intelligible to the aeronautical community)

Pronunciation, stress, rhythm, and intonation are influenced by the first language or regional variation but only sometimes interfere with ease of understanding.

Structure (Relevant grammatical structures and sentence patterns are determined by language functions appropriate to the task.)

Basic grammatical structures and sentence patterns are used creatively and are usually well controlled. Errors may occur, particularly in unusual or unexpected circumstances, but rarely interfere with meaning.

Vocabulary

Vocabulary range and accuracy are usually sufficient to communicate effectively on common, concrete, and work-related topics. Can often paraphrase successfully when lacking vocabulary in unusual or unexpected circumstances.

Fluency

Produces stretches of language at an appropriate tempo. There may be occasional loss of fluency on transition from rehearsed or

formulaic speech to spontaneous interaction, but this does not prevent effective communication. Can make limited use of discourse markers or connectors. Fillers are not distracting.

Comprehension

Comprehension is mostly accurate on common, concrete, and work related-topics when the accent or variety used is sufficiently intelligible for an international community of users. When the speaker is confronted with a linguistic or situational complication or an unexpected turn of events, comprehension may be slower or require clarification strategies.

Interactions

Responses are usually immediate, appropriate, and informative. Initiates and maintains exchanges even when dealing with an unexpected turn of events. Deals adequately with apparent misunderstandings by checking, confirming, or clarifying.

There is some controversy over the quality and empirical basis of these scales (see below), but at the time of this writing (January 2009), these scales and the associated LPRs were mandated for use in the assessment of proficiency in the use of English for aviation purposes, specifically for licensing, although they are also frequently used as the basis of test construction for placement and achievement as well.

By March 5, 2008, air traffic controllers and pilots were required by the ICAO to have a certificate attesting to their proficiency in the language used for aeronautical communication. ICAO Document 9835, "Manual on the Implementation of ICAO Language Proficiency Requirements," lays out the principles according to which the language proficiency requirements should be met. Although several organizations made every effort to produce suitable tests or assessment procedures by the deadline, in the event an implementation period was allowed, with a new deadline of March 2011.

Aviation language tests are obviously extremely high stakes, and it is crucially important that such tests are constructed to the highest possible standards. It is, therefore, a matter of considerable concern that certificates attesting to the attainment of relevant levels of language proficiency will not be generally required of aviation personnel until March 2011.

Aviation English Tests

In 2006, two years before the ICAO LPRs were due to be implemented, the European Organization for the Safety of Air Navigation (Eurocontrol) commissioned the Lancaster Language Testing Research Group to conduct a validation study of the development of a test called ELPAC (English Language Proficiency for Aeronautical

Communication), which was intended to assess the language proficiency of air traffic controllers.

During the 18 months of this study, an Interim Report was produced, which made recommendations for the improvement of the tests and the associated quality control procedures. A Final Report was also produced, which provided a commentary on the quality of the ELPAC test and made a series of recommendations for further quality control measures. The Executive Summary of that Final Report is available at http://www.elpac.info/, accompanied by a commentary on the recommendations by Eurocontrol. As part of the ELPAC Validation Study, Internet searches for evidence of other tests of air traffic control found very little evidence available to attest to the quality of the tests for licensure of either pilots or air traffic controllers. It was therefore decided to conduct an independent survey of all providers of tests intended for air traffic control. Alderson (2008) reported on the methodology and findings of that survey, the results of which are summarized below.

Since the Executive Summary of the ELPAC Validation Study had been framed by the Guidelines for Good Practice of the European Association for Language Testing and Assessment (EALTA) it was decided to base the survey questionnaire on the same Guidelines, which consist of the following headings:

- Test Purpose and Specification
- Test Design and Item Writing
- Test Analysis
- Test Administration and Security
- Test Revision
- Test Washback and Impact

The Guidelines are available in full in 34 European languages at www.ealta.eu.org/guidelines.htm. Alderson and Banerjee (2008) described how the questionnaire was adjusted to meet the needs of the particular aviation context.

The survey, made available on the Internet, consisted of two stages: a filter questionnaire that elicited opinions about the ICAO scales, as well as general information about aviation English testing (for the results of that first stage, see below and Knoch, 2009). The second stage asked detailed follow-up questions and was sent to 74 organizations and individuals whose tests were thought to be used for licensure of pilots and air traffic controllers. Only 22 relevant responses were received, which varied considerably in quantity and quality. While this disappointing response rate may be indicative of the lack of quality of the tests produced and an associated reluctance to admit this in public, nonresponse cannot be taken to indicate lack of quality, although it may well indicate lack of public accountability of such providers.

The tests used for licensure identified in Alderson (2008) are shown in Table 1.

Table 1. Licensure tests

 1 a) Level 6 Proficiency Demonstration b) Formal Language Evaluation
 2 ALITE
 3 Altec Benchmark Evaluation
 4 Aviation English Proficiency Assessment
 5 CXELT (Cathay English Language Test - Pilots)
 6 ELP Test V1 (obsolete); ELP Test V2 (operational); ELP Test V3 (in development)
 7 ELPAC (English Language Proficiency for Aeronautical Communication)
 8 Aviation Language Proficiency Test (available for both English and French)
 9 English for Aviation Language Test (EALT); Expert Level 6 Speaker Assessment (ELSA)
10 English Proficiency Exam for Aviators (Chile)
11 English Proficiency Test for Airline Pilots (in Japanese)
12 English Proficiency Test for Aviation (EPTA)
13 ICAO English Proficiency Exam for Aviators
14 IELTS
15 Ilmailuhallinnon kielitaitotesti (The language proficiency test of the [Finnish] Civil Aviation Authority)
16 LANG TECH Aviation English Oral Competence Assessment
17 RELTA
18 TEA (Test of English for Aviation)
19 TELLCAP® (Test of English Language Level for Controllers and Pilots)
20 Test of English for Aviation Purposes (the name may have changed after the development period)
21 Thai DCA aviation test
22 Versant Aviation English Test (VAET)

The institutions that appear to be involved in developing these are shown in Table 2.

Test validation reports were received as shown in Table 3.

These validation reports were supplied in confidence, and regrettably, they appear not to be easily available on the Internet. Interested readers are invited to search the Internet for further details of these validation studies, as the results are not published in Alderson (2008).

The following list contains some relevant URLs:

http://elpac.info/index.php?option=com_content&task=view&id=50&Itemid=42
http://www.maycoll.co.uk/aviation-english/tea-research.html
http://www.ordinate.com/versant/business/aviation/overview.jsp

Table 2. Institutions developing tests

1   Aviation Services Limited, New Zealand
2   Griffith University, Australia
3   Altec Internationale
4   Berlitz
5   Cathay Pacific China
6   Aerosolutions, Belgium
7   EUROCONTROL
8   Transport Canada—Civil Aviation
9   MLS International, United Kingdom
10   Universidad Técnica Federico Santa María
11   Sophia Linguistics Institute for International Communication, Sophia University
12   G-TELP KOREA
13   Universidad Técnica Federico Santa María
14   IELTS Australia
15   Finnish Civil Aviation Authority
16   Language Technology
17   RMIT University, Australia
18   Mayflower College, United Kingdom
19   Aviation English Training Center "CompLang," Russia
20   Colegio de Pilotos Aviadores (México)
21   Department of Civil Aviation, Thailand
22   Ordinate Corporation, USA

Table 3. Validation reports

○ ELPAC Validation Study Final Report
○ Japanese pilots test: English Proficiency Test for Airline Pilots, Sophia University
   (translated from Japanese)
○ Expert Opinion on external validation of TELLCAP®
○ TEA (Test of English for Aviation) Mayflower College Research notes
○ VAET (Versant Aviation English Test) Versant with Ordinate Technology

Detailed findings of the survey, under six main headings, are available in Alderson (2008), but it is worth quoting the final summary at length:

> In only a minority of cases was there evidence of adequate concern for quality control and public accountability. Too often, confidentiality was claimed as a reason for not providing essential information. Too often, vague answers were provided which failed to answer the questions asked, and too often, certain test developers or organizations simply failed to respond to questions, or even to

understand what was required, or why it is important to be able to supply relevant information.

Nevertheless, it was reassuring to see that at least some test developers took their responsibilities seriously, and made every effort, both to answer our questions fully, and to supply additional information in the form of validation studies. As we agreed to maintain the anonymity of respondents, we cannot identify those who represent good practice in test development, nor those who appear not to understand what good practice entails. We can point out, however, that,

- if test specifications and test content do not match the needs of test-takers,
- if test developers and item writers do not have relevant aviation and language education experience,
- if item writers and test raters are not appropriately trained,
- if tests are not adequately pretested on suitably representative and sizable samples of test-takers,
- if the results of such trials are not suitably analyzed and actions taken to address any evident problems,
- if test-takers are not aware of how and on what they will be tested and how their performance will be rated,
- if the reliability of marking is not monitored, calculated and reported,
- if there is no statistical information available to support the claimed level of the test, the equivalence of different versions from year to year, and the comparability of the results of different tests purporting to measure the same target level of proficiency,
- and, above all, perhaps, if reports providing evidence as to the thoroughly professional quality of tests are not publicly available,

then little or no confidence can be held in the meaningfulness, reliability and validity of several of the aviation language tests currently available. Unfortunately, at present, this appears to be the case for too many of the tests that we have surveyed, and if other tests exist for the aviation context that have not responded to our survey, then it is highly likely that they, too, fail to meet minimal standards of quality. The consequences of inadequate language tests being made available to license pilots, air traffic controllers and other aviation personnel are almost too frightening to contemplate (pp. 14–15).

The main conclusion of the Alderson report is that "monitoring is required of the quality of language tests used in aviation to ensure they follow accepted professional standards for language tests and assessment procedures" (p. 1).

Further Surveys

Survey of National Civil Aviation Authorities' Plans for Implementation of ICAO

 Language Proficiency Requirements

Alderson and Horák (2008) reported on a follow-up survey into the plans for implementation of the Language Proficiency Requirements requested of national civil aviation authorities by the ICAO. The research was carried out in two rounds. Round 1 addressed an e-mail letter to the person named as responsible for implementation of the LPR in each national civil aviation authority. The letter requested the names of the tests that the authority had recognized or approved. Out of the 190 authorities (plus five overseas territories of three of these countries) listed on the ICAO Web site, only 24 replied to the researchers' request.

The second round involved going back to the ICAO Web site and conducting a detailed analysis of all replies posted on that Web site as of October 2008. A total of 53 countries stated that they had complied with regulations requiring the assessment of the English language proficiency of pilots and air traffic control staff. Another 53 had not supplied ICAO with compliance details and the remaining 89 stated noncompliance.

However, of the 53 stating compliance, different countries replied with varying degrees of information, which rarely constituted evidence of compliance. Of those 14 states that provided estimates of the language proficiency levels of pilots and ATCs, it is far from clear how accurate these estimates are, as details of the assessments are scant and not always obviously relevant.

Alderson and Horák concluded that

- the lack of (evidence for) compliance gives cause for concern;
- ongoing detailed monitoring of implementation plans and compliance is essential;
- there is reason to suspect that the ICAO's recommendations with respect to evidence for test quality are not being taken seriously;
- the conclusions are confirmed of Alderson (2008) that we can have little confidence in the quality of several of the aviation language tests and assessments currently available for flight crew and air traffic controller licensure. (p. 1)

Study of Stakeholders' Views of the ICAO Scales

Knoch (2009) reported on a study of two groups of stakeholders: aviation language testing experts and pilots and pilot trainers. An online survey (as reported briefly above) was used with the first group and focus group interviews with the second, based on recordings of speaking performances of test takers taking a variety of aviation English tests. Results from the first group revealed wide criticism of the

ICAO scale and a range of problems was reported. Informants were evenly split between believing that the ICAO scales were more useful to English experts or aviation experts, and whether a Level 4 did or did not represent an adequate level for operational flying.

Results from pilots' use of the scales showed that they paid less attention to some criteria (e.g., structure) than others, and at times judgments of technical knowledge appeared to influence judgments of language ability. The article concludes with recommendations for improvements to the ICAO scales.

Informal Discussions of Issues Surrounding Aviation Language Testing
and the ICAO Scales

There are a number of discussion fora associated with aviation English teaching and testing, and the present author is a member of some. The content of these discussions is assumed to be confidential to members of the fora, but a content analysis of issues recently under discussion is not believed to breach confidentiality or anonymity. No claim is made as to the representativeness of the summaries below, which may, to some extent at least, reflect the bias of this author.

ICAO Scales. ICAO's recent focus on issues of second language proficiency is welcome, and it is acknowledged that ICAO has provided much assistance to enhance the quality of tests and assessment processes, particularly Document 9835, its harmonization document (ICAO, 2008), and test checklists and speech samples. The current situation was felt to be much better than previously and the awareness of language issues is now greater in the aviation community. However, the lack of evidence for empirical validation of the LPR scales was commented on, and there was discussion of, and suggested revisions to, the ICAO Scale. The suitability of the ICAO speech samples was questioned, with often very poor sound quality and little agreement on ICAO levels.

Need for Standardization of Raters and Ratings. There is felt to be a great need for training and monitoring of raters, and plans for the development of a rater accreditation scheme are being discussed. The likelihood of a global scheme was debated, as was the role of the International Civil Aviation English Association (ICAEA) and the ICAO.

Test Quality. Evidence is needed for the reliability and validity of available tests, not merely their commercial success or how many users there are. There is a generally recognized need for independent audits of tests, raters, and rater training, and there was a feeling that inadequate tests should be "named and shamed" (with recognition, and examples, of good practice). Many anecdotes were recounted of problems with particular tests, certificates, countries.

Test Monitoring and Accreditation or Recognition. There is no list of approved tests. ICAO does not know which tests are approved by which authority, or how they were approved. What knowledge is there of the need for quality language

tests? Who is responsible for quality control? According to its mandate, ICAO sets the ultimate regulatory framework in which all its member States operate; application of the regulatory framework is the responsibility of the civil aviation authority of each State, but ICAO does operate a (restricted and infrequent) audit oversight. On the ground, compliance is apparently largely dependent on a lack of proficiency being reported by pilots or controllers in operational situations. This is frequently felt to be unsatisfactory, and regulatory bodies are sometimes felt to have an ostrich-like attitude.

Politics and Economics.   The politics of getting agreement among all 190 Contracting States of ICAO is clearly an issue. Stakeholders need convincing that it is worth spending money on *valid* English tests, and there is a need for a campaign to make stakeholders aware of the importance of second language proficiency in air safety. Airlines are said to prefer low-cost solutions. Part of the solution might be for aviation English specialists to work more closely with operational trainers. However, in current economic circumstances (and falling profit margins), second language proficiency is not recognized as a priority for scarce resources (at least until a major accident occurs which is shown to be due to language problems). English teachers or testers are thought to lack credibility within the aviation community. Moreover, pilot and air traffic controller associations resist testing.

Corpus of Aviation Communication via Radio Telephony.   Communication between pilots and ATCs is said to be often very poor, even for native speakers of English. The need was expressed for evidence of language-related incidents. Research into such incidents would be greatly facilitated by the development of a corpus of pilot–pilot and pilot–ATC radiotelephony communication. Such a corpus could usefully be complemented by rated and validated speech samples based on properly conducted assessment procedures, representatively collected and quality controlled.

Testing Issues.   Considerable discussion was held of a range of testing issues. What is the value of a language test for ensuring flight safety: Is it not more important to observe how language is used under stressful conditions? Are tests (in general, as well as specific tests) an adequate representation of the target language use domain? Should English experts or aviation operations experts judge proficiency levels? What is the relative value of using human raters versus machine scoring? Should computer-based tests or face-to-face OPI-style interviews be used? What are the merits and demerits of each? Some feel that the ICAO scale is inadequate for the assessment of listening during face-to-face speech performances, and that there is a need for a separate measure of listening comprehension. Does it make sense for each country or airline operator to develop its own tests, or is there a need for globally uniform tests? Is a Level 4 on one test equivalent to a Level 4 on a different test? Are the ICAO scales sufficiently explicit and relevant to guarantee that any test constructed on the basis of the ICAO scales will be at the right level? There would appear to be a need for comparative studies.

Native Speaker Issues.   Despite the fact that Level 6 is not defined as native speaker proficiency, which is made clear in Document 9835, at least some civil

aviation authorities interpret Level 6 as meaning a native speaker level. For example, the UK Civil Aviation Authority states that "a scale of 1 to 6 has been devised where native speakers would be assessed at Level 6." The document goes on to state that as Radiotelephony Operating License Examiners, Type Rating, Flight and Class Rating Examiners are not trained in formal language proficiency assessment, the assessments they carry out can only be for proficiency at Level 6.

In other words, native speakers of English can be examined at the highest level—Level 6—by unqualified examiners, and only nonnative speakers have to be examined by qualified language examiners. This is felt by many to be inappropriate, since nonnative speakers are often better at aviation English than native speakers. Native speakers are often insensitive to their own language use (lack of language awareness)—specifically their inappropriate use of idioms and colloquial language, the likely difficulties nonnative speakers experience in communicating with them, and the need to adapt their speech to their audience. Native speakers should not automatically be granted Level 6, but should be tested in the same way and under the same conditions as nonnative speakers.

Conclusion

This article has examined the development of a policy intended to regulate the language proficiency of pilots and air traffic controllers, and has analyzed the state of implementation of that policy to date. Serious inadequacies have been found in the state of preparedness of national civil aviation authorities to comply with the ICAO regulations. Moreover, it is unclear whether the assessment processes proposed or in place meet international standards for high-stakes language testing. Furthermore, there appears to be no reliable mechanism for overseeing the implementation of the policy and ensuring the quality of the assessment procedures and tests. There is a lack of clarity as to where the ultimate responsibility lies for adequate monitoring of the quality of such tests and assessment procedures, and who might be held responsible for any accidents that may be shown to be due, in part or in whole, to inadequate aviation language proficiency.

Clearly, further research is needed, and the issues summarized in the previous section from the discussion board exchanges can be seen as identifying a range of topics for research agendas, as well as matters that need more discussion and attention by applied linguists generally, and by the language testing community specifically. Aspects that seem of particular importance for research include close analysis of the language of nonnormal aviation communication; examining the accident reports of national transport safety bureaus and air accidents investigation units; critiquing and revising the ICAO Language Proficiency Requirements Scale and subjecting them to empirical investigation and validation; critical reviews and comparisons of available aviation English tests; and a close examination of the politics of policy development, implementation, and resistance.

However, even more important, it is argued, is the urgent need for some means of monitoring the quality of the aviation tests and assessment procedures that are, or will be, available before March 2011.

A suitable conclusion for this article, perhaps, is a quotation from Andy Roberts' wishlist for 2009, posted on one of the aviation English discussion lists and reproduced here by permission, which neatly summarizes the key issues needing urgent attention:

> I'd like to put my own wishes forward for 2009.
>
> 1. To continue raising the standards of Aviation English levels of pilots and controllers worldwide.
>
> 2. To have a mechanism to check that countries that post as being ICAO-compliant on the FSIX Web site really are.
>
> 3. To apply the recommended guidelines for test construction to both tests and the rating scale.
>
> 4. To have a mechanism to check that tests and training do what they claim to do.

ANNOTATED REFERENCES

International Civil Aviation Organization (ICAO). (2004). *Manual on the implementation of ICAO language proficiency requirements* (Doc 9835). Montreal: International Civil Aviation Organization.
This is a key document but is far from easy to find. A catalogue is available at http://www.icao.int/icao/en/sales/cat_2008_en.pdf

This manual provides guidance on the implementation of the ICAO Language Proficiency Standards for flight crew and air traffic controllers. It contains comprehensive information on a range of aspects related to language proficiency and training that will assist Contracting States, operators and air traffic services providers in implementing efficient language training and testing. It also contains information that will help the language training and testing industry in providing services that are relevant to the civil aviation context. (p. 44)

The manual contains eight chapters and five appendices. Various chapters address the need for strengthened language proficiency requirements, and for international cooperation, the nature of ICAO phraseology and the use of plain language, and describe the ICAO rating scale. Matters discussed include the need for linguistic awareness, the importance of cross-cultural communication, the nature of radiotelephony communications, relevant language training, compliance with the language proficiency requirements, and aviation language testing. Appendix A Part II includes the ICAO Language Proficiency Rating Scale and a set of

explanations as Additional Information; Appendix B presents a useful analysis of aviation language, including communicative functions, events and domains, lexis and language tasks. Appendix C describes two case studies in aviation language testing–surprisingly, the OPI and PELA (a placement test, not a test for licensure); Appendix D presents standards for language training and testing, including the ILTA Code of Ethics and the Japanese Language Testing Association's Code of Practice; and Appendix E has a useful set of references.

While containing much useful information, the manual is probably simply too specialized for national civil aviation authorities to follow in its entirety, although it is a useful point of reference, and it would be very helpful if (1) a detailed critical review and set of recommendations were to be made with respect to the content of the manual and (2) a brief, informative digest were available for lay users. Perhaps even more important, however, would be easy and free access to what is clearly an important document.

International Civil Aviation Organization (ICAO). Other useful documents are at http://www.icao.int/anb/fls/AUD001/ and include *Checklist on the Development and Implementation of Testing for ICAO Language Proficiency Requirements*. Retrieved February 21, 2009, from http://www.icao.int/anb/ fls/AUD001/checklist.pdf

Rated speech samples. Retrieved February 21, 2009, from http://www.icao.int/icao/ en/cd_pub_list.htm#Language

This CD is intended to facilitate the implementation of the ICAO Language Proficiency Standards and contains speech samples rated at ICAO Language Proficiency Levels 3, 4, and 5. Each of the speech samples is accompanied by a detailed rating form that contains the underlying rationale for the rating. In addition, the CD contains information on the ICAO Language Proficiency Rating Scale and on language proficiency testing. This CD is of interest to Civil Aviation Authorities, air navigation service providers, training institutions, airlines, and institutions imparting Aviation English courses and conducting language proficiency tests and is related to the *Manual on the Implementation of ICAO Language Proficiency Requirements* (Doc 9835).

Implementation checklist for aviation language testing. Retrieved February 21, 2009, from www.icao.int/icao/en/anb/meetings/ials2/Implementation%20Stepsv3- 21April.pdf

Alderson, J. C. (2008). *Final report on a survey of aviation English tests*. Retrieved December 3, 2008, from www.ealta.eu.org/guidelines.htm

The article reports on the results of a survey that used a Web-based questionnaire, based on the Guidelines for Good Practice of the European

Association for Language Testing and Assessment (EALTA). Invitations to contribute to the survey were sent to some 74 individuals and organizations whose tests were thought to be used for licensure of pilots and air traffic controllers.

In all, 22 responses were received, which varied considerably in quantity and quality. This probably reflects a variation in the quality of the tests, the availability of evidence to support claims of quality, and low awareness of appropriate procedures for test development, maintenance and validation. The Survey indicates that it is unclear whether national civil aviation authorities have the knowledge to judge the quality of tests. It is concluded that there can be little confidence in the meaningfulness, reliability, and validity of several of the aviation language tests currently available for licensure. It is recommended that monitoring is required of the quality of language tests used in aviation to ensure they follow accepted professional standards for language tests and assessment procedures.

OTHER REFERENCES

Alderson, J. C. (2008). *Final report on a survey of aviation English tests*. Retrieved December 3, 2008, from www.ealta.eu.org/guidelines.htm

Alderson, J. C., & Banerjee, J. V. (2008). *EALTA's guidelines for good practice: A test of implementation*. Paper presented at the 5th Annual Conference of the European Association for Language Testing and Assessment, Athens, Greece, May. Retrieved December 24, 2008, from http://www.ealta.eu.org/conference/2008/programme.htm

Alderson, J. C., & Horák, T. (2008). *Report on a survey of national civil aviation authorities' plans for implementation of ICAO language proficiency requirements*. Unpublished manuscript.

Cabre, M. T., Freixa, J., Lorente, M., & Tebe, C. (1998). La terminologia hoy: Replanteamento o diversificacion. Terminologia e integracao. *Organon, 12*(26), 33–41, cited in Sarmento (2005, p. 2).

Chen, H. (n.d.). Blog. Retrieved December 22, 2008, from http://harry.hchen1.com/2007/10/15/555

Cushing, S. (1994). *Fatal words: Communication clashes and aircraft crashes*. Chicago: University of Chicago Press.

Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge, England: Cambridge University Press.

Farris, C., Trofimovich, N., Segalowitz, N., & Gatbonton, E. (2008). Air traffic communication in a second language. Implications of cognitive factors for training and assessment. *TESOL Quarterly, 42*(3), 397–410.

Helmreich, R. L. (1994). Anatomy of a system accident: The crash of Avianca flight 052. *International Journal of Aviation Psychology, 4*, 265–284.

Helmreich, R. L., Klinect, J. R., & Wilhelm, J. A. (1999). Models of threat, error and CRM in flight operations. In R. S. Jensen (Ed.), *Proceedings of the 10th*

International Symposium on Aviation Psychology (pp. 677–682). Columbus: Ohio State University (cited in Nevile & Walker, 2005, p. 3).

International Civil Aviation Organization (ICAO). (2008). *Language testing criteria for global harmonization* (Circular 318-An/180). Montreal: Author.

International Civil Aviation Organization (ICAO). (n.d.). *ICAO language proficiency for flight crew*. Retrieved December 30, 2008, from http://www.caa.co.uk/docs/1688/ICAO%20Language%20Proficiency%20for%20Flight%20Crew.pdf

Knoch, U. (2009). *Collaborating with ESP stakeholders in rating scale validation: The case of the ICAO rating scale* (Spaan Fellow Working Papers, 7). English Language Institute, University of Michigan.

Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2003). *Concurrent task demands in the cockpit: Challenges and vulnerabilities in routine flight operations*. Paper presented at the 12th International Symposium on Aviation Psychology, Dayton, Ohio, April 14–17.

Mell, J. (n.d.). *Emergency calls—Messages out of the blue*. Retrieved December 22, 2008, from http://www.icao.int/anb/sg/pricesg/background/OotB.htm

Mell, J. (1992). *Study of verbal communication between pilot and air traffic controller in standard and non-standard situations*. Doctoral thesis. Paris: Ecole Nationale de l'Aviation Civile.

Merritt, A., & Ratwatte, S. (2004). "Who are you calling a safety threat?" A debate on safety in mono-cultural versus multi-cultural cockpits. In M. A. Turney (Ed.), *Tapping diverse talent in aviation: Culture, gender, and diversity* (pp. 173–183). Hampshire, England: Ashgate.

Nevile, M. (2002). Coordinating talk and non-talk activity in the airline cockpit. *Australian Review of Applied Linguistics, 25*(1), 131–146.

Nevile, M. (2004). *Beyond the black box: Talk-in-interaction in the airline cockpit*. Aldershot, England: Ashgate.

Nevile, M., & Walker, M. B. (2005). Analysis of crew conversations provides for accident investigation. *Flight Safety Digest, 24*(11), 1–17.

Ragan, P. H. (2007). Cross-cultural communication in aviation. In K. Ahmad & M. Rogers (Eds.), *Evidence-based LSP: Translation, text and terminology*. Selected papers from LSP2003, the 14th European Symposium on Language for Special Purposes held at the University of Surrey, Guildford, England, in cooperation with the AILA Scientific Commission on Language for Special Purposes (pp. 54–63). Bern, Switzerland: Peter Lang.

Ripley, R. F., & Finch, J. L. (2004). The efficacy of standard aviation English. In M. A. Turney (Ed.), *Tapping diverse talent in aviation: Culture, gender, and diversity* (pp. 99–103). Hampshire, England: Ashgate.

Sarmento, S. (2005). A pragmatic account of aviation manuals. *English for Specific Purposes World, 4*(3). Retrieved December 22, 2008, from www.esp-world.info/Articles_11/apragmaticaccountofaviationmanuals%5B2%5D.htm

Shawcross, P. (1993). *English for aircraft maintenance*. Paris: Belin (cited in Sarmento, 2005, p. 11).

Stratechuk, T., & Beneigh, T. (2004). Reflexive communication in the multi-cultural crew. In M. A. Turney (Ed.), *Tapping diverse talent in aviation: Culture, gender, and diversity* (pp. 105–117). Hampshire, England: Ashgate.

Turney, M. A. (2004a). Values and orientation differ in mixed crews. In M. A. Turney (Ed.), *Tapping diverse talent in aviation: Culture, gender, and diversity* (pp. 11–19). Hampshire, England: Ashgate.

Turney, M. A. (2004b). Protocols, rank and social status influence communication. In M. A. Turney (Ed.), *Tapping diverse talent in aviation: Culture, gender, and diversity* (pp. 161–170.). Hampshire, England: Ashgate.

Useful Web Sites

The ICAO Web site http://www.icao.int/ and specifically the Flight Safety (FLS) Section, with Frequently Asked Questions relevant to personnel licensing: http://www.icao.int/icao/en/trivia/peltrgFAQ.htm

Background to the development of the LPRs http://www.icao.int/anb/sg/ pricesg/background/

ICAO Doc 9835 (2004) Manual on the Implementation of ICAO Language Proficiency Requirements. 1st Edition. International Civil Aviation Organization.

ICAO catalogue is available at http://www.icao.int/icao/en/sales/cat_2008_en.pdf

Checklist on the development and implementation of testing for ICAO Language Proficiency Requirements is at
   http://www.icao.int/anb/fls/AUD001/checklist.pdf

Rated speech samples is at
http://www.icao.int/icao/en/cd_pub_list.htm#Language

International Civil Aviation English Association (ICAEA)
http://www.icaea.pata.pl/

References to accident investigations and databases
- U.S. National Transportation Safety Board Web site http://www.ntsb.gov/
- Aviation Safety searchable database http://aviation-safety.net/database/
- FAA Aerospace Medicine Technical Reports http://www.faa.gov/library/reports/medical/oamtechreports/2000s/
- The United Kingdom's Air Accident Investigation Branch Web site http://www.aaib.dft.gov.uk/publications/index.cfm

Reports of Aviation Incidents

Tenerife, 1977

YouTube reenactment at http://www.youtube.com/watch?v=qa4dr6NoWJc and the flight safety database record at http://aviation-safety.net/database/record.php?id=19770327-1

Avianca Flight 052, Cove Neck
http://aviation-safety.net/database/record.php?id=19900125-0

Kazakhstan Airline aircraft collision with Saudi Arabian Boeing 747, Charkhi Dadri, New Delhi, India

www.airsafe.com/events/models/boeing.htm

Codes of Practice in Language Testing

European Association for Language Testing and Assessment (EALTA) Guidelines for Good Practice in Language Testing and Assessment (available in 34 European languages) is at www.ealta.eu.org/guidelines.htm.

International Language Testing Association (ILTA) Draft Code of Practice
http://www.iltaonline.com/CoP_3.1.htm

Association of Language Testers in Europe Code of Practice
http://www.alte.org/quality_assurance/index.php

Possible locations for validation studies
http://elpac.info/index.php?option=com_content&task=view&id=50&Itemid=42
http://www.maycoll.co.uk/aviation-english/tea-research.html
http://www.ordinate.com/versant/business/aviation/overview.jsp