

Тестирование в авиационном контексте

ISBN 978-5-6049111-0-5



Мельниченко Н.П.
Мельниченко С.А.

Теория и практика создания и применения тестирования для определения уровня владения английским языком по Шкале ИКАО

Содержание

1. Вступление
2. Тестирование в области специальных языков
 - 2.1. Авиационный английский язык. Аэронавигационная радиосвязь
 - 2.2. Анализ ситуации использования языка
 - 2.3. Требования, предъявляемые к квалификационному тестированию
 - 2.3.1. Валидность
 - 2.3.2. Надежность
 - 2.3.3. Практичность
3. Разработка тестирующей системы
 - 3.1. Паспорт тестирования
 - 3.2. Планирование этапов разработки тестирования
 - 3.3. Критерии отбора разработчиков тестирования
 - 3.4. Анализ потребностей. Спецификация теста
 - 3.5. Описание конструктора
 - 3.6. Формат теста. Типы тестовых заданий. Продолжительность теста. Временные пределы тестовых заданий
 - 3.7. Разработка банка заданий. Спецификации тестовых заданий
 - 3.8. Испытание бета-версии, сбор данных и их анализ
4. Подготовка экзаменаторов и рецензентов
 - 4.1. Инструкции для экзаменаторов по получению пригодного образца речи
 - 4.2. Инструкции для рецензентов по оценке образца речи
5. Мероприятия по обеспечению валидности и надежности теста
6. Заключение

1. Вступление

В марте 2008 года требования ИКАО к владению английским языком для ведения радиосвязи стали квалификационными. Это означает, что пилоты, диспетчеры, операторы аэронавигационных станций и удаленные пилоты должны не только обладать профессиональной подготовкой и практическими навыками, а также соответствовать медицинским требованиям, но и владеть английским языком как минимум на четвертом (рабочем) уровне по Шкале ИКАО. До принятия этого требования существовала лишь рекомендация в отношении того, что английский язык должен обеспечиваться в качестве средства общения по запросу [*ICAO Doc 4444, 13th Edition*].

Справедливости ради следует отметить, что серьезные инциденты, вызванные языковыми проблемами, происходили довольно редко. Raimund Fridrich [*Skyguide Magazine, 2006*] сообщает, что на тот исторический период по причинам проблем с ведением радиосвязи приходился всего 1% от общего количества известных (заявленных) инцидентов. В то же время, по данным Евроконтроля, эти инциденты составляли 23% от общего количества инцидентов, связанных с УВД, а около 39% всех проблем в общении диспетчеров и пилотов относились к человеческому фактору [*Eurocontrol Safety Study, 2004*].

Целый ряд громких катастроф (столкновение на ВПП рейсов KLM 4805 и Pan Am 1736 в аэропорту Лос-Родос /о. Тенерифе, Испания/, 1977 год; полная выработка топлива у рейса Avianca 052 в районе Нью-Йорка, 1990 год; столкновение с горами рейса American 965 в Колумбии, 1995 год; столкновение в воздухе рейсов Saudi Arabian 763 и Kazakhstan 1907 над Индией, 1996 год; и т.д.) привели к тому, что Совет ИКАО принял решение о создании Исследовательской группы по определению уровня владения языком (PRICE SG – Proficiency Requirements in Common English Study Group).

В работе группы участвовали филологи, пилоты, диспетчеры, преподаватели, служащие сертифицирующих органов полномочных учреждений гражданской авиации и ученые-лингвисты, представлявшие различные страны и международные авиационные организации.

Цели PRICE SG

Целью работы группы [*ICAO, PRICESG, 2000*] стала разработка фундаментальных стандартов владения английским языком для принятия их в качестве квалификационных требований и внесения соответствующих поправок в Приложение 1 – Выдача свидетельств авиационному персоналу, Приложение 10 – Авиационная электросвязь, том 2, а также в ряд других документов Международной организации гражданской авиации.

Как и планировалось, группа изложила требования к владению английским языком диспетчерами и пилотами в «описательных характеристиках», а не в «предписывающих нормативных условиях». Цели работы группы включали разработку обобщающей, но вместе с тем конкретной шкалы уровней владения английским языком, описанной с точки зрения языкового поведения. Такая задача была выполнена, и теперь в гражданской авиации эта шкала известна как Шкала ИКАО для определения уровня владения языком.

Группа в течение своей трехлетней работы пришла к выводу, что государствам следует предоставить возможности для использования «чужих» или разработки «своих» тестов, которые бы наилучшим образом подходили для их собственных обстоятельств. Исследовательская группа PRICE SG также включила в разработанный позднее инструктивный материал ссылки на большое количество статей и исследований в области обучения и тестирования.

С первого рабочего совещания, которое состоялось в Монреале 14-17 ноября 2000 года, было определено, что цели работы группы должны включать, во-первых, положения о владении общим английским языком, и, во-вторых, о владении специальным авиационным английским языком. Было признано, что последний элемент выходит далеко за рамки «законсервированных» фраз из Приложения 10 и Документа 4444, и включает в себя общение пилотов между собой и с диспетчерами при возникновении срочных и аварийных ситуаций.

Группа обратила внимание на то, что для поддержания достигнутого уровня владения языком необходимо регулярное обучение. Что касается разработки тестов, то ИКАО подготовило и опубликовало инструктивный материал в отношении методов построения теста, включавших собеседование (или устное интервью), а также задания на аудирование и говорение. Весь методический материал вошел в отдельный Документ ИКАО 9835 – *Руководство по внедрению требований ИКАО к владению языком*.

Успешной работе группы способствовало то, что различные организации согласились предоставлять материалы и оказывать другую помощь: это Федеральное управление гражданской авиации США, Центр английского языка Лингвистического института министерства обороны США, ЕВРОКОНТРОЛЬ, НАВКАНАДА, ИФАЛПА, Федеральная служба воздушного транспорта России и Государственное предприятие обслуживания воздушного движения Украины.

Благодаря общей позиции членов группы было принято решение разрешить государствам обеспечивать максимальную гибкость в установлении собственных порядков обучения и тестирования. Также было установлено, что ИКАО не будет разрабатывать какую-либо конкретную образовательную программу или систему оценки.

Объем работы

В мире к тому времени уже были проведены серьезные исследования с целью создания тестов для проверки навыков устной речи и аудирования. Однако в отношении чтения и говорения таких исследований не проводилось, и это особо касалось новых на то время технологий, таких как, например, канал передачи данных CPDLC. Поэтому было принято решение, что эти два элемента следует рассматривать отдельно, чтобы не задерживать работу с уже достаточно исследованными устной речью и аудированием.

Необходимо отметить, что вся работа группы PRICE SG была сосредоточена на английском языке, а не на процедурах, используемых в воздушном пространстве, включая фразеологию радиообмена. В то же время участники признавали абсолютную необходимость соблюдения фразеологии ИКАО как основы обеспечения безопасности полетов при ведении радиотелефонной связи.

К сожалению, это требование, содержащееся в Приложении 10, том 2, не соблюдается и после внедрения требований ИКАО к владению языком. В 2011 году ИАТА, ИФАЛПА и ИФАТКА провели совместное исследование [*IATA et al., 2000*] использования стандартной фразеологии пилотами и диспетчерами в различных странах мира. На вопрос: «Существуют ли аэропорты, где стандартная фразеология ИКАО не используется?», почти половина (48%) опрошенных пилотов ответили однозначно «да». Среди наиболее проблемных в этом отношении аэропортов с большим отрывом идут Париж/Шарль де-Голль, Нью-Йорк/Дж.Ф.К. и Мадрид/Барахас.

Обоснование

Принимая во внимание, что главная задача работы группы заключалась в обеспечении безопасности полетов, был разработан единый минимальный стандарт для аудирования и говорения для всей целевой группы (т. е., пилотов, диспетчеров и операторов аэронавигационных станций). Было признано, что авиационный английский язык может быть успешно интегрирован в общий английский язык и, следовательно, стандарт должен относиться к использованию общего английского языка.

Такая стратегия послужила тому, чтобы подчеркнуть необходимость уверенного владения общим английским языком до того, как начнет изучаться какой-либо из аспектов авиационного английского языка.

При достижении минимально допустимого уровня владения общим языком, можно успешно подключить обучение специальному авиационному контенту – как для приобретения знаний, так и практическому его применению, включая словарный запас и использование фразеологии, такой как фразеология, предписанная ИКАО.

Стратегия

Признавая наличие различий среди государств между существовавшими на тот момент уровнями владения английским языком и имеющимися предложениями по их повышению, государствам было предложено разработать поэтапную программу практической реализации квалификационных требований ИКАО. Способы, которыми это должно было осуществляться, включали требование, во-первых, чтобы сначала новому квалификационному стандарту соответствовали пилоты и диспетчеры, а позже – остальные авиационные специалисты. При определении даты соответствия учитывались как непосредственная потребность в безопасности, так и разумный период времени, в течение которого государства смогли бы соответствовать новым языковым стандартам. Предполагалось, что такой подход будет учитывать существовавшие реалии и различия между государствами, процессы составления бюджета и т. д. Стратегия внедрения требований должна была включать возможные временные ограничения, связанные с количеством диспетчеров и пилотов. Она также учитывала некоторые сложности с подтверждением квалификации для пожилых сотрудников.

Принимая во внимание необходимость в том, чтобы пилоты и диспетчеры поддерживали достигнутые уровни владения языком, была изучена и одобрена необходимость периодического квалификационного подтверждения уровня владения языком. Вместе с тем было установлено, что за обучающие программы продолжают нести ответственность сами государства.

2. Тестирование в области специальных языков

2.1. Авиационный английский язык и аэронавигационная радиосвязь

Чтобы начать серьезно говорить о тестировании уровня владения общим и авиационным языком по Шкале ИКАО, сначала необходимо определить место авиационного английского языка в овладении английским языком в качестве второго или третьего языка. Шкала ИКАО может быть использована и для определения уровня владения любым языком, но мы сосредоточим все внимание на английском языке. Для это есть несколько причин. Английский язык давно является языком межнационального общения: начиналось это еще со времен колоний Великобритании. На территориях многочисленных британских колоний сначала использовались так называемые пиджины (Pidgin-English), которые состояли из смеси местного(-ых) языка(-ов) и упрощенных форм английского языка. Окончательно английский язык в качестве *lingua franca* утвердился после Второй мировой войны. США не только меньше всех пострадали от последствий войны, но и сумели занять лидирующие позиции в экономике, науке и технике. «За американским долларом стоял английский язык» (D. Cristal, 1997). Однако в настоящее время для 70% людей, говорящих на английском языке, он не является родным.

В радиосвязи на заре развития авиации использовался сначала французский язык, о чем свидетельствует некоторое количество слов французского происхождения во фразеологии

радиообмена. Однако в 30-е и 40-е годы, после Второй мировой войны особенно, в авиации в качестве языка международного общения утвердился английский язык. На земном шаре еще остались регионы, где языками международного общения являются испанский и китайский языки, и они используются в радиообмене. К сожалению, русский язык после крушения Советского Союза утратил свои позиции в этом качестве, уступив их английскому языку. Если в начале 2000-х годов речь еще заходила о создании тестирующей системы по определению уровня владения русским языком по Шкале ИКАО, то теперь в этом уже нет никакой необходимости.

Авиационный английский язык – это язык для специальных целей, или профильно-ориентированный язык. О языках для специальных целей (English for specific purposes - ESP) серьезно заговорили в 60-х годах прошлого века.

К концу 80-х были сформулированы основные понятия, что такое языки для специальных целей. Одно из определений было дано в 1987 году: язык для специальных целей – это «скорее подход, чем продукт» (*Hutchinson and Waters, 1987*). По мнению авторов, основой языка для специальных целей является простой вопрос: для чего учащемуся необходимо изучать иностранный язык? В какой области, при каких обстоятельствах учащийся собирается использовать язык?

Позже были даны и другие определения языка для специальных целей, мы приведем одно из последних, наиболее полное.

Абсолютные характеристики:

- Английский язык для специальных целей предназначен для удовлетворения конкретных потребностей учащегося;
- Английский язык для специальных целей использует методологию и виды деятельности, лежащие в основе дисциплины, которую он обслуживает. Английский язык для специальных целей сосредоточен на языке (грамматика, лексика, регистр), навыках, дискурсе и стиле, присущим этому роду деятельности.

Переменные характеристики:

- Английский язык может относиться или быть разработан для конкретной дисциплины;
- Английский язык для специальных целей может использовать в конкретной ситуации обучения методы, отличные от методов обучения общему английскому языку;
- Английский язык для специальных целей разрабатывается скорее всего для взрослых учащихся, как для высших учебных заведений, так и для обучения на рабочем месте. Однако он может быть применен и для учащихся средней школы;
- Английский язык для специальных целей разрабатывается большей частью для студентов среднего и продвинутого уровней. Многие курсы английского языка для специальных целей предполагают наличие основных знаний языковой системы, но он может быть использован и с начинающими. (*Dudley-Evans and St John, 2006*).

Авиационный английский язык соответствует всем перечисленным характеристикам. Авиационным английским языком пользуются диспетчеры и пилоты для ведения радиосвязи, бортпроводники для общения с пассажирами во время полета, производители воздушных судов для написания различных руководств по эксплуатации воздушных судов, составители сборников аэронавигационной информации, разработчики оборудования для организации и управления воздушным движением, метеорологи и т. д.

2.2. Анализ ситуации использования языка

Ситуации использования языка во всех этих случаях сильно отличаются друг от друга. Так как Шкала ИКАО предназначена для определения уровня владения языком только членами экипажей воздушных судов и диспетчерами, мы остановимся подробно на ситуации использования языка (*target language use*) в радиообмене. В процессе изучения ситуации использования языка в рамках дискурсивного анализа (*discourse analysis*) можно провести анализ жанра, или функционального стиля (*genre, or register analysis*) и анализ потребностей (*needs analysis*). В ходе изучения ситуации использования языка мы увидим, что тип коммуникации в радиосвязи отличается лексико-грамматическими, семантико-прагматическими и дискурсивными ресурсами от других ситуаций использования авиационного английского. Тип коммуникации является одной из черт дискурсивного сообщества (*discourse community*). Определение дискурсивного сообщества было дано в 1990 году (*J.M. Swales, 1990*):

1. Дискурсивное сообщество имеет широко согласованный набор общих целей.
2. В дискурсивном сообществе существуют механизмы коммуникации между его членами.
3. Дискурсивное сообщество использует свой общий механизм в первую очередь для передачи информации и обратной связи.
4. Дискурсивное сообщество использует и, следовательно, владеет одним или более функциональными стилями для коммуникативного продвижения своих целей.
5. В дополнение к владению функциональными стилями дискурсивное сообщество использует некоторое количество специальной лексики.
6. Дискурсивное сообщество имеет пороговый уровень количества членов с достаточной степенью предметной и речевой компетенции.

Позже J.M. Swales разделил все дискурсивные сообщества на три группы: локальные (*local d.c.*), целевые (*focal d.c.*) и локально-целевые (*folocal*). Целевые он разделил на две подгруппы: профессиональные (*professional d.c.*) и досуговые (*recreational d.c.*).

Приведем несколько примеров дискурсивных сообществ:

1. Жители поселка аэропорта Внуково знают, что место встречи «пяточок» — это пересечение Центральной улицы и 2-ой Рейсовой, но для людей, не живущих или не работающих во Внукове, это понятие ничего не значит (*локальное дискурсивное сообщество*).
2. Филологи сразу поймут, о чем идет речь, если услышат про сингармонизм гласных в тюркских языках или о первом передвижении согласных, а словосочетание «назначение крейсерских эшелонов полета» понятно только членам авиационного сообщества (*целевые профессиональные дискурсивные сообщества*).
3. Рыбаки знают, что «вibroхвост» — это приманка в виде рыбки, изготовленная из силикона, однако вряд ли значение этого слова понятно тем людям, которые не увлекаются рыбалкой (*целевое досуговое дискурсивное сообщество*).
4. «KBOS 230954Z 23010KT 10SM FEW250 12/01 A3007»
Передача данных о видимости в сводке погоды в сухопутных милях и данных для установки высотомера в дюймах ртутного столба применяется в США и Канаде и отличается от общепринятой практики (*локально-целевое профессиональное дискурсивное сообщество*).

J.M. Swales в описании всех типов дискурсивных сообществ обратил внимание на то, что особый интерес для лингвистов представляют все-таки профессиональные дискурсивные

сообщества. Предметом нашего изучения будет дискурсивное сообщество, в которое входят пилоты и диспетчеры.

Каждый человек является одновременно членом нескольких дискурсивных сообществ, но в профессиональной сфере использование языка может отличаться от повседневной жизни. Хотя пилоты и диспетчеры используют язык и вне рамок своей профессиональной деятельности, только в области радиообмена выбор лингвистических средств определяется специфическими требованиями жанра, или функционального стиля. «Название «функциональный стиль» представляется очень удачным, потому что специфика каждого стиля вытекает из особенностей функций языка в данной сфере общения» (И. В. Арнольд, 2006).

Функциональный стиль является механизмом коммуникации в радиообмене. Он используется для передачи информации, диспетчерских указаний и обратной связи. В радиообмене используется большое количество технической и псевдо-технической лексики. К ведению связи допускаются только специалисты, имеющие рабочий уровень владения языком по Шкале ИКАО и владеющие фразеологией радиообмена.

Почему ИКАО определила 4 уровень владения языком как рабочий уровень? Этот выбор определил анализ ситуации использования языка. Ситуация использования языка, или радиообмен, существенно отличается от использования языка в других сферах.

Широко известны особенности радиообмена:

- отсутствует визуальный контакт между говорящим и слушателем, речь не поддерживается мимикой и жестами; голос передается при помощи радиооборудования и подвергается при этом искажению; коммуникация носит преимущественно устный характер;
- используется английский язык для специальных целей.

Отличия радиосвязи определены требованиями, которые предъявляются к ней: радиообмен должен быть кратким и однозначным (*ICAO Annex 10, Volume II*). Мы имеем дело с контекстом, который очень сильно ограничен как в сфере языка, так и ситуационно.

Для стандартных ситуаций разработана фразеология ИКАО, ее применение обязательно во всех ситуациях, для которых она предписана. Однако изучение ситуации использования языка показывает, что этой фразеологии недостаточно, когда ситуация выходит за рамки стандартной. В таких случаях мы наблюдаем переключение кода (*code switching*), т. е. переход от стандартной фразеологии к использованию общего языка. Использование общего языка должно учитывать требования краткости и однозначности, т. е. соответствовать «ожиданиям жанра», или функционального стиля. «Функциональный стиль – это не «специальный язык», это подстраивание использования языка под требования ситуации» (А. Vatnsdal, 1987). Это должно учитываться как при обучении, так и при тестировании. Однако «то, что на первый взгляд кажется простым, неполным, или даже грамматически неправильным, может быть достаточно сложным «в исполнении» (А. Vatnsdal, 1987).

Итак, приведем основные черты радиообмена:

- В радиообмене отсутствуют такие формы вежливости, как “Would you please...”, “Nice to hear you” и т. п. Различные формы приветствий (Good morning, Good day) допускаются, но не приветствуются правилами ведения радиосвязи.

- Для экономии времени применяются коды и специальные лексические единицы (*QFE, QNH, QDM, MAYDAY, PAN PAN, Wilco*).
- Часто используются добавочные коды (*additional attention markers: Caution, Attention, Be advised, Be informed*).
- Для радиосвязи характерно употребление интенсификаторов (*markers of urgency: Immediately, Now*).
- Использование инверсии или препозиции (*Inversion or fronting process*) позволяет диспетчеру обосновать команду до того, как он ее озвучит, и дает дополнительное время пилоту на реакцию на указание (*For identification turn left 30 degrees*).
- В радиообмене используются простые предложения (*Read you five*).
- Для диспетчерских указаний употребляется только повелительное наклонение (*Climb to flight level 260*).
- В авиационном английском и в радиообмене, в частности, из синонимического ряда выбирается обычно лексическая единица с пометкой 'formal': *assistance, until, revert, observe, request, expedite, exceed, resume, require, commence* и т. п.
- В радиосвязи отсутствуют такие средства организации речи, как вводные слова, союзы. Организации речи служит очередность элементов служит.
- Вопросительные предложения заменяются конструкциями с *report, request, confirm*.

В радиосвязи используется грамматика для специальных целей, или эллиптическая грамматика (*a special-purpose grammar, or elliptical grammar*), для которой типичны различные формы речевой компрессии:

- отсутствие артиклей, когда они не несут смысловой нагрузки (*Delay not determined*);
- пропуск вспомогательных глаголов, когда они не несут смысловой нагрузки (*Climbing to flight level 230. Cleared via TELIN 2 Departure*);
- отсутствие многих предлогов (*Contact Radar 126,0, Remain this frequency*);
- употребление левых определений (*Departure airdrome, nose wheel, weather report, airdrome vehicle, surface wind*);
- отсутствие личных и притяжательных местоимений, что частично объясняется стремлением радиообмена как функционального стиля к дистанцированию и обезличенности (*Report position. Turning final*).

Почему возможно использовать эллиптические грамматические формы, опускать местоимения, предлоги?

Это возможно из-за очень высокой контекстуальной чувствительности радиообмена, все недостающие элементы могут быть восстановлены из контекста. В нашем случае контекстом служит ситуация.

Стереотипность ситуаций в радиообмене и предыдущий опыт участников радиообмена позволяют прогнозировать вероятность наступления тех или иных событий.

Стереотипность ситуаций позволяет обходиться меньшим по объему словарем, пользоваться клише, а в синтаксисе широко использовать эллиптические предложения. Перечисленные выше лингвистические особенности ситуации использования языка при радиообмене позволяют сделать его кратким. Как избежать неоднозначности? Неоднозначность может быть лексического и структурного/грамматического свойства (*ambiguities of lexical and structural nature*) (*S. Cushing, 1995*).

При ведении радиосвязи к выбору лексических единиц надо подходить с большой аккуратностью. Известно, что большинство слов в языке полисемантичны, например, слово *right* имеет несколько значений, среди них есть значения "true/correct" и "not left". В

радиообмене слово *right* употребляется только в значении “not left”, а в подтверждении правильности полученной информации обязательно должно прозвучать “correct” или “that is correct”. Это правило звучит так: одно слово – одно значение.

Однозначность достигается и правильным выбором лингвистических средств для выражения речевой интенции, или коммуникативной задачи. В качестве примера рассмотрим такую коммуникативную задачу, как побуждение к действию (*request for an action*). Для выражения этой интенции есть разные способы: повелительное наклонение, модальный вопрос с *could*, вопрос с конструкцией *would you mind + Ving* и другие способы.

Просьба повторить информацию, которую адресат сообщения не понял, может быть выражена в разном виде:

1. Say again assigned flight level.
2. Could you repeat assigned flight level please?
3. Would you mind repeating assigned flight level?
4. You are coming broken. I'm not understanding you!

Каждое высказывание имеет три значения:

What did you say? - буквальное значение (*literal meaning*).

What did you want to say? – интенциональное значение (*intended meaning*).

What did you mean by that? - прагматическое значение (*pragmatic meaning*) (*James E. Purpura, 2005*)

Буквальное значение складывается из значения слов и грамматической формы высказывания. Интенциональное значение выводится из речевой интенции, или коммуникативной задачи, высказывания. В дополнение к буквальному и интенциональному значению высказывание одновременно может шифровать прагматическое значение данного контекста (*What was meant without actually being said*), то есть то, что имеется в виду, не будучи озвучено. Прагматическое значение, другими словами, подтекст, выводится главным образом из контекста, и оно может быть преднамеренным или непреднамеренным.

Буквальное значение и речевая интенция, выраженная в повелительном наклонении, совпадают с интенциональным значением в первом примере. Отношения этих значений прямые и явные. В первом примере сказано именно то, что имелось в виду, без какого-либо подтекста.

Во втором примере коммуникативная задача та же самая, что и в первом примере, но выражена посредством модального вопроса с *could*. Гипотетически буквальное значение этого высказывания может служить для выражения запроса информации или побуждения к действию. Несмотря на то, что буквальное значение высказывания может служить для выражения нескольких коммуникативных задач, как в этом случае, говорящий имеет в виду только одно интенциональное значение, когда высказывание является в контексте. Прагматическое значение в этом примере – это выражение вежливости. Побуждение к действию становится не таким явным.

То же самое и с примером 3, когда для побуждения к действию выбрана другая конструкция, вопрос с *would you mind + Ving*. Как и во втором примере, вне контекста высказывание может быть интерпретировано и как запрос информации, и как побуждение к действию. Прагматическое значение – выражение вежливости.

В этих трех примерах взаимоотношения между грамматической формой и значениями, которые они выражают, между буквальным значением и речевой интенцией могут быть определены по нарастающей как менее прямые и менее явные.

В четвертом примере эти отношения еще менее явные. В четвертом примере высказывание вне контекста может выражать описание фактической информации, а в контексте это все-таки побуждение к действию. Прагматически высказывание несет выражение сиюминутного раздражения. Это высказывание очень сильно зависит от контекста.

Возвращаясь к радиообмену как функциональному стилю, определим, какая языковая форма должна применяться для выражения речевой интенции *request for an action*. Помним, что нам надо ответить не только на вопросы “What did you say? What did you want to say? What did you mean by that?”, но и на вопрос “Why did you say it that way in this context?”, то есть, насколько уместны языковые формы в данном контексте. В Документе ИКАО 9835 написано: “Be specific, explicit, and direct... Utterances should explicitly state the function of the communication...” Уверенно можем сказать, что только первый пример соответствует этим критериям и уместен в радиообмене.

Из нескольких возможных вариантов необходимо выбрать тот, который будет соответствовать «ожиданиям жанра» (*genre expectations*), или требованиям, предъявляемым к радиообмену – краткости и однозначности. На индивидуальном уровне говорящий как член дискурсивного сообщества конструирует свое сообщение, представляя нужды воображаемого слушателя.

Как только участники радиообмена об этом забывают, происходит сбой в коммуникации, например, на вопрос диспетчера, для которого английский не является родным языком: “What level are you passing now?” - пилот (носитель языка) реагирует фразой: “Say again”. Трудно предположить, что пилот не понял диспетчера, акцент которого был абсолютно понятен. Просто в рамках жанра он ожидал услышать более краткое и однозначное: “Report flight level”.

Другой пример абсолютного несоблюдения требований жанра и нарушения правил ведения радиосвязи. Пилот бизнес-джета, несмотря на метеоусловия выше минимума аэродрома, не смог выполнить посадку (не смог пробить облачность) на трех аэродромах Московской воздушной зоны. На предложение диспетчера следовать на запасной аэродром пилот бизнес-джета отвечает: “Yes, if you give me the weather, we need to go somewhere where we can land, we don't have a lot of time, we just came from Japan.” Очень многословный ответ. Сообщение о том, что у них мало времени, что они летят из Японии, — это намек на то, что у них мало топлива, и запасной он просит и поближе, и с метеоусловиями, которые позволят выполнить сразу посадку. Отношения между буквальным и интенциональным значениями не совсем очевидны, лингвистические средства для выражения речевой интенции в радиообмене выбраны неправильно. Реплика пилота: “Request weather at alternate. Be informed we are low on fuel.” – была бы более уместна в этой ситуации. Мало того, что пилот не был готов воспринимать данные о видимости и нижней границы облачности в единицах измерения, отличных от единиц измерения, принятых в США (диспетчер передавал эти данные в метрах, а пилот воспринимал их в футах), он еще нарушил и правила ведения радиосвязи (стандартную фразеологию подменил лексикой и предложными схемами из общего языка). Весь радиообмен во время четырех его попыток сесть на любом из аэродромов состоял из подобных нарушений.

Что же касается прагматических значений (выражение недовольства, разочарования, сарказма и пр.), их в радиообмене не должно быть вообще. Мы уже писали, что радиообмен стремится к обезличенности, а возмущение, упреки, нравуочения, оскорбления в эфире свидетельствуют об эмоциональной незрелости говорящего, что вряд ли сочетается с той ответственностью, которая на нем лежит (*P.E. Illman, J. Pouzar, 1988*)

2.3. Требования, предъявляемые к квалификационному тестированию

Проведя анализ ситуации использования языка и описав подробно особенности радиообмена, мы можем сделать вывод, что и тестирование в этой сфере должно отличаться от тестирования в общем языке и других специальных языках.

«Тест по языку для специальных целей – это тест, методы и содержание которого вытекают из анализа определенной цели ситуации использования языка, таким образом, задания и содержание теста являются аутентичным отражением задач в реальной ситуации и делают возможным взаимодействие между языковой компетенцией и знанием сути конкретных задач с одной стороны и заданий теста с другой. Такой тест позволяет нам сделать предположения о способности экзаменуемого использовать язык в сфере определенной цели» (*D. Douglas, 2002*).

Языковая компетенция — это не только знания о языке, но и способность их применять. Мы уже говорили о зависимости друг от друга языка и ситуации его применения. Речь здесь идет о том, что в языках для специальных целей языковая компетенция всегда связана с предметными знаниями (*topical or subject knowledge*) и с предыдущим опытом в профессиональной сфере. Если задания теста не отражают специфики сферы деятельности экзаменуемого, то сделать предположение о его способности решать коммуникативные задачи на рабочем месте не представляется возможным.

Нельзя, однако, специфику понимать примитивно, то есть, недостаточно посадить пилота в кабину экипажа, а диспетчера перед экраном локатора и поставить им их повседневные задачи. «Без управления самолетом язык перестает нести какое-либо информативное значение и становится подобием фразеологии вне всякого контекста» (*Teasdale, 1996*). Почему? Во-первых, невозможно предусмотреть все ситуации, с которыми экзаменуемые могут столкнуться в реальной жизни. Во-вторых, сразу возникает проблема: учитывать или нет действия пилота или диспетчера, ведь рейтер способен оценить только языковую составляющую, а на результат коммуникации влияют и действия, например, по предотвращению столкновения в воздухе. При оценке такого элемента Шкалы ИКАО как взаимодействие рейтер должен оценить результат коммуникации, - всегда ли речевое поведение экзаменуемого было адекватным. Ситуация тестового задания должна быть такой, чтобы было очевидно, что результат коммуникации отражает уровень языковой компетенции, а не профессиональные знания пилота или диспетчера, т. е. если реплики тестируемого были неадекватными, то это показатель недостаточной языковой компетентности, а не отсутствие предметных знаний или профессионального опыта.

Поясним примером не из авиации. Мигрант из Средней Азии, желая работать таксистом в нашей стране, должен получить российские водительские права, сдав экзамен на русском языке в ГИБДД. Результат может зависеть как от знания правил дорожного движения, так и от владения русским языком. Если не сдал экзамен, причина может быть неочевидна - то ли вопросы не понял, то ли ПДД не знает.

Несмотря на то, что предметные знания и языковая компетенция взаимосвязаны, тест по определению уровня владения языком не должен оценивать профессиональные навыки и предметные знания.

Тест должен быть аутентичным, ситуации тестовых заданий должны отражать специфику профессиональной деятельности, но не должны воспроизводить ситуацию из радиообмена. Позже мы еще вернемся к аутентичности ситуаций тестовых заданий.

Итак, тест по определению уровня владения языком членами летных экипажей и диспетчерами УВД должен дать ответ на вопрос: соответствует ли языковая компетенция экзаменуемых рабочему уровню по Шкале ИКАО?

Тестирование по Шкале ИКАО – это квалификационное тестирование, которое позволяет сделать вывод о том, может ли пилот или диспетчер выполнять свои профессиональные обязанности. Определение уровня владения английским языком пилотами и диспетчерами называется «тестированием высокой ставки», потому что от результата теста зависит не только судьба конкретного пилота или диспетчера, но и безопасность пассажиров и грузов.

Если речь идет о квалификационном тестировании, то тест должен соответствовать таким критериям как валидность, надежность, практичность.

2.3.1. Валидность

Валидность – это адекватность, пригодность, обоснованность. Это – комплексная характеристика теста, отражающая обоснованность, значимость его результатов, адекватность теста целям измерения.

Главный вопрос в языковом тестировании: оценивает ли тест то, что является целью тестирования? Если тест не оценивает то, для чего он предназначен, результаты теста не могут считаться значимыми. Разные тесты могут быть валидными для разных целей. Например, разработчики теста заявляют, что тест разработан для оценки уровня владения языком по Шкале ИКАО с 1 по 5 уровень, но в тесте нет ни одного задания, которое бы позволило экзаменуемому продемонстрировать 5 уровень. Конечно, валидность такого теста подвергается сомнению: безусловно, тест валиден при определении 1 – 4 уровней, но заявленная цель тестирования шире, значит, тест не является валидным.

В исследованиях описаны различные типы валидности, можно придерживаться этого порядка, так как это упрощает объяснение, но обязательно следует подчеркнуть, что «типы» валидности на самом деле являются разными «методами» оценки валидности (*L. Bachman, 1990*). Чем больше типов валидности будет установлено, чем больше свидетельств валидности будет собрано, тем лучше. Валидность теста должна подтверждаться не раз и навсегда. Это должен быть «динамический процесс сбора различных типов свидетельств, которые складываются, как мозаика» (*D. Douglas, 2002*).

Начнем с двух терминов: внутренняя валидность (*internal validity*) и внешняя валидность (*external validity*). Разберем их более подробно.

Внутренняя валидность сводится обычно к изучению содержания и его обратного эффекта.

Внутренняя валидация, т. е. оценка внутренней валидности, может быть проведена несколькими способами, опишем некоторые из них.

Face validity – номинальная, или очевидная валидность (восприятие теста не экспертами, а теми, кто его сдает). Иногда разработчики тестирования ограничиваются одним вопросом: понравился или не понравился тест? Ответ всегда ожидаем: нравится тест, который легче всего сдать. Увы, такой вопрос не имеет отношения к валидации вообще. Для валидации нужно создать опросник для экзаменуемых со следующими вопросами. Все ли инструкции к заданиям были понятны? Хватило ли времени на выполнение того или иного задания? Все ли задания отражают специфику его профессиональной деятельности? Вызывали ли какие-нибудь вопросы или задания внутреннее раздражение своей неуместностью? Мешало ли это выполнению задания? И т. д. Статистически значимое количество полученных ответов и их анализ являются оценкой номинальной валидности.

Пример очевидной невалидности теста: экзаменуемый видит перед собой задание уточнить время задержки перед пересечением ВПП с подсказкой использовать слово *delay*. Такое задание вызовет недоумение у экзаменуемого, ведь о задержке (*delay*) диспетчер предупреждает до запуска двигателей. Обсуждать время ожидания во время руления перед пересечением ВПП или РД пилот не будет, у диспетчера нет времени на пустые разговоры. Как только освободится ВПП, он даст разрешение ее пересечь. Ситуацию нельзя назвать аутентичной. Такого рода задания вызывают недоумение, замешательство у тестируемого, влияют на качество ответа и свидетельствуют об отсутствии предметных знаний у разработчиков теста.

Содержательная валидность (*content validity*). В отличие от номинальной валидности оценка содержательной валидности проводится экспертами, чье мнение уважаемо большинством экспертного сообщества. Как получить мнение экспертов? Опишем один из возможных способов. Анализ потребностей определил лексические домены, коммуникативные функции и навыки, характерные для определенной ситуации использования языка. Конкретные домены, функции и навыки включаются в спецификацию каждого задания. Пусть эксперт определит в образце речи, какой навык или навыки оценивает то или иное задание теста, отражает ли это конструкт теста, совпадает ли это со спецификацией, адекватно ли это задание.

Например, задание в спецификации описано следующим образом: Интервью на авиационную тематику – тестовое задание большей или меньшей управляемости для проверки навыков говорения и понимания на слух, которое предполагает развернутые ответы (*extended response*) и спонтанную речь. Эксперт же, прослушав образец речи экзаменуемого, делает вывод, что это задание в оцениваемом тесте не что иное, как набор случайных вопросов, не связанных друг с другом, в образце речи вместо развернутых ответов и спонтанной речи заученные наизусть определения авиационных понятий.

Эксперт должен обратить внимание и на то, насколько аутентичны задания, отражают ли они специфику профессиональной деятельности (см. пример с *delay*). Для этого можно привлечь пилотов и диспетчеров, обладающих большим объемом предметных знаний (*subject matter experts*).

Перейдем к внешней валидации. Наиболее частый способ внешней валидации – это оценка параллельной валидности (*concurrent validity*) путем сравнения результата теста с результатами других измерений, полученных в тот же период времени, например, с итоговой оценкой в конце курса или результатом другого валидного теста. Метод вычисления коэффициента корреляции приведено у J. Charles Alderson (2006). В случае

параллельной валидности необходимо быть уверенным в валидности и надежности результатов тестов или других оценок, с которыми вы сравниваете свой тест. Параллельная валидность используется и при внутренней валидации вариантов заданий теста и параллельных версий теста.

Конструктивная валидность (*construct validity*) считается многими лингвистами более высокой формой валидности. Конструкт (*construct*) – достаточно широкое понятие, но в случае языкового тестирования это – характеристика теста, определяющая на основе анализа содержание теста. Это – описание умений и навыков, которые проверяет тест.

Чтобы описать конструкт теста, придется воспользоваться результатами анализа потребностей. Мы уже писали, что во время дискурсивного анализа, изучая ситуацию использования языка, мы можем определить те лингвистические средства, которые необходимы именно в данной ситуации использования языка. Мы еще вернемся к описанию конструкта, когда перейдем к практическим аспектам создания теста для определения уровня владения языком членами летных экипажей и диспетчерами УВД.

2.3.2. Надежность

Надежность теста – характеристика качества теста, отражающая степень постоянства, устойчивости результатов тестирования. Если тест надежен, результат не зависит ни от варианта, ни от рейтера.

Здесь следует отметить, что, кажущиеся на первый взгляд простыми, взаимоотношения между валидностью и надежностью на практике достаточно сложные и для многих непонятные.

«Тест не может быть валидным, если он не надежен» (*J. Charles Alderson, 2006*). Если тест не измеряет стабильно, следовательно, он не может всегда измерять точно. Но, как ни странно, тест может быть надежным, но не валидным. Тест может измерять стабильно и точно то, что не является целью измерений. Надежность без валидности ничего не значит.

Самые надежные тесты – это тесты с множественным выбором (*multiple-choice tests*), но множественный выбор никак не подходит для оценки говорения и интерактивного понимания на слух. В таком случае придется делать выбор в пользу валидности за счет надежности. На практике ни валидность, ни надежность не бывают абсолютны. Всегда приходится выбирать компромисс, и выбор будет зависеть от цели тестирования.

Надежность теста достигается разными способами. Чтобы тестовые задания и параллельные варианты были одинаковыми, необходимо на стадии разработки теста написать четкие и понятные для сценаристов, составителей заданий (*item writers*), спецификации всех заданий, а на стадии внутренней валидации проверить как можно больше заданий и вариантов на предмет параллельной валидности.

Другой путь достижения надежности теста – качественное обучение экзаменаторов и рейтеров, постоянный мониторинг их работы и мероприятия по обеспечению стабильности их работы.

2.3.3. Практичность

Практичность теста – это простота организации теста. Многие думают, что чем сложнее организация, тем тест надежнее. Вовсе нет.

Конечно, создание теста требует и много времени, и достаточных финансовых вложений. Лучше потратить деньги на то, чтобы нанять квалифицированный персонал для разработки и применения теста, на обеспечение сохранности материалов тестирования, чем сделать тест технически сложным в эксплуатации. Практичность теста, на наш взгляд, не связана с валидностью и надежностью теста.

Можно потратить много денег на обустройство сервера где-нибудь в горах Норвегии, о местоположении которого будет знать ограниченный круг лиц, но в какой-то момент не очень хорошо владеющий английским языком диспетчер в Македонии обнаружит в себе задатки хакера и вдруг получит 5 уровень. С другой стороны, тест, для проведения которого требуется только ноутбук, флэшка, две гарнитуры, шторка и хорошо обученный репитер/экзаменатор, может быть и валидным, и надежным, и практичным.

3. Разработка тестирующей системы

Перед тем, как приступить к разработке тестирующей системы, а не просто теста (кроме самого теста необходимо разработать механизм его применения и поддержания), нужно составить план с подробным описанием всех этапов разработки и перечень всех документов, которые потребуются для одобрения/лицензирования в органах гражданской авиации, и для принятия решения о применении тестирующей системы той или иной авиакомпанией или провайдером аэронавигационных услуг.

3.1. Паспорт тестирования

Пожалуй, самый главный документ для этого – это паспорт тестирования. В Паспорте тестирования должны быть даны ответы на следующие вопросы:

- Какая информация о тестирующей системе находится в открытом доступе?
- Имеется ли демонстрационная версия теста в открытом доступе?
- Какое оборудование необходимо для проведения теста?
- Как принимается заявка от экзаменуемого?
- Кто назначает экзаменатора/репитера?
- Как и кем выбирается вариант теста?
- Как проводится идентификация личности тестируемого?
- Что включено в брифинг перед экзаменом?
- Сколько репитеров оценивают образец речи?
- Как и в каком виде репитер получает образец речи для оценки?
- Что происходит, если оценки репитеров не совпадают?
- Возможна ли апелляция со стороны экзаменуемого?
- Возможна ли повторная сдача теста?
- Как собирается информация об отношении тестируемых к тесту и проводится ли ее анализ для подтверждения номинальной валидности?
- Как подтверждается валидность и надежность тестирования?
- Как обеспечивается сохранность материалов тестирования?
- Подписывают ли экзаменаторы/репитеры соглашение о нераспространении материалов тестирования?
- Как часто обновляется банк тестовых заданий?

3.2. Планирование этапов разработки тестирования

В паспорте тестирования обязательно описывается план, подготовительные этапы и подробная история разработки теста, указываются конкретные даты и продолжительность

каждого этапа: подготовительный этап, отбор персонала (сценаристов, репитеров-экзаменаторов, администраторов, технического персонала), создание бета-версии, пробное тестирование бета-версии на контрольной группе, сбор и обработка полученной информации, внесение изменений в тест (если необходимо), повторное испытание измененных версий, разработка банка заданий и дата начала использования теста. Эта информация важна, потому что процесс работы над тестом сказывается на его качестве. Если не документировать процесс создания, мутность происхождения теста повлияет на его репутацию.

3.3. Критерии отбора разработчиков тестирующей системы

В паспорт следует включить состав разработчиков (возможно, анонимный) с указанием образования, учебного заведения, опыта преподавания общего и авиационного английского языка, радиообмена или практического применения языка в ситуации использования языка, например, в качестве переводчика или языкового инструктора. Уровень владения языком должен быть не ниже 5 уровня по Шкале ИКАО. Как показывает опыт, лингвисты, не имеющие опыта в ситуации использования языка, оказываются бесполезными при работе над квалификационным тестированием, или требуется время для их обучения. В команду разработчиков необходимо включить пилотов и диспетчеров, обладающих большим объемом предметных знаний.

Обязательным условием для всех разработчиков является высшее лингвистическое образование, а для пилотов и диспетчеров высшее специальное образование и достаточный опыт работы, для пилотов-инструкторов желательно в крупной авиакомпании, для диспетчеров-инструкторов в органах УВД. Если привлечь пилотов-любителей или диспетчеров на пенсии, то тест может не отвечать такому критерию как аутентичность, а это непременно скажется на валидности теста.

3.4. Анализ потребностей. Спецификация теста

Анализ ситуации использования языка и анализ потребностей дают ответы на вопрос, каким должен быть валидный тест. Они определяют содержание и конструкт теста. В п. 3.7 мы приводим лишь небольшой фрагмент анализа ситуации использования языка из более чем 120 часов аудиозаписей радиообмена в нестандартных и аварийных ситуациях. Мы проводили анализ этих записей в 2003-2004 годах, еще до выхода Документа ИКАО 9835, не зная, что подобный анализ был уже выполнен Ecole Nationale de l'Aviation Civile в Тулузе. Теперь при разработке тестирования можно воспользоваться результатами их анализа из Документа ИКАО 9835. Мы публикуем фрагмент наших исследований, чтобы была понятна связь анализа ситуации использования языка, анализа потребностей и содержания тестирования.

Ответив на следующие вопросы, мы получим подробную характеристику теста:

- Какова цель тестирования?
- Какие навыки будет оценивать тест?
- Кто входит в целевую популяцию тестирования? Будет ли это тест для пилотов, для диспетчеров или для тех и других одновременно?
- Если задание в тесте для диспетчеров, то должен ли он иметь разные версии для диспетчеров Вышки, Руления, Подхода, РДЦ? Если задание в тесте для пилотов, должны ли быть разные версии для коммерческих и частных пилотов, пилотов самолетов и вертолетов, удаленных пилотов?
- Будет ли тест включать радиообмен? В каком виде и объеме?

- Какова продолжительность и качество аудиоматериалов, которые мы планируем использовать?
- Столкнется ли экзаменуемый с различными акцентами в тесте?
- Будет ли тест определять 6 уровень по Шкале ИКАО?
- Заинтересованы ли разработчики в *interactional* или *transactional language*?
- Проверку каких лексических кластеров и каких коммуникативных функций включить в тестовые задания?
- Какой образец речи считать пригодным для оценки?

Тест должен оценивать уровень владения языком по Шкале ИКАО членами летных экипажей и диспетчерами УВД. Анализ потребностей определил 4 уровень по Шкале ИКАО как рабочий уровень. Это значит, что только после подтверждения рабочего уровня владения языком по Шкале ИКАО пилот или диспетчер могут получить свидетельства своей профессиональной пригодности. Для этой цели подойдет только такая категория теста как квалификационный тест. Только квалификационное тестирование позволяет сделать вывод о том, может ли пилот или диспетчер выполнять свои профессиональные обязанности.

Итак, тест согласно требованиям ИКАО должен оценивать навыки говорения и понимания на слух, для этого нельзя задействовать навыки чтения и письма. Контроль понимания на слух должен быть интерактивным.

Описание коммуникативных функций пилотов и диспетчеров приведено в Документе ИКАО 9835 в Приложении В.

Общие коммуникативные задачи (*macro-skills*) являются общими и для пилотов, и для диспетчеров: *request action, suggest a solution to a problem, undertake to contact/relay/report, describe an action in progress, announce a change, state necessity, ask about intentions, warn, read back* – и т. п.

Конкретные коммуникативные задачи (*micro-skills*) могут относиться только к диспетчерам (*to brief pilot on weather data in specified format, give holding instructions, give advice, inform of airspace status change* и т. п.), а какие-то – только к пилотам (*to receive information about runway infringement and runway confusion, request assistance of appropriate ramp vehicles, explain reason and extent of deviation off route due to operational reasons* и т. п.). Однако очень часто конкретные коммуникативные задачи пилотов являются зеркальным отражением задач диспетчера и наоборот. Например: *request present whether information* (пилот)/*relay present whether information* (диспетчер). Это означает, что большинство тестовых заданий может быть адресовано как пилотам, так и диспетчерам.

Нужно ли разрабатывать разные версии для диспетчеров Вышки, АДЦ, РДЦ и т. д.? А для пилотов с разными пилотскими свидетельствами? Мы считаем, что не нужно, потому что в Документе ИКАО 4444 сказано, что «пользователи должны быть знакомы и применять при необходимости фразеологию из других разделов, не относящихся конкретно к типу предоставляемого обслуживания воздушного движения». Хотя речь здесь идет только о фразеологии, с уверенностью можем утверждать, что это относится и к общему, и к авиационному языку.

При анализе использования языка мы столкнулись с ситуациями, подтверждающими это. Приведем в пример следующую ситуацию: борт компании Lufthansa, направляясь в Японию, в Московской воздушной зоне запросил срочную посадку в Домодедово из-за больного пассажира. Пилот несколько раз запрашивал у диспетчера РДЦ, потом у

диспетчеров Подхода и Круга наличие в Домодедово подъемника для лежачих больных (*ambulift*), используя все возможности перефразирования, но ответа до совершения посадки так и не получил. Причина: диспетчеры РДЦ и АДЦ не были знакомы с названиями аэродромной техники, считая, что это должны знать только диспетчеры Вышки и Руления.

Что касается пилотов, независимо от того, каким воздушным судном они управляют, ситуационная осведомленность, т. е. понимание того, что звучит в радиоэфире, только способствует безопасности полетов. Содержание теста не должно зависеть от рабочего места диспетчера или типа воздушного судна, которым управляет пилот.

Существуют тесты, в которых оценивается владение фразеологией радиообмена. Однако знание фразеологии ИКАО и способность ее применять всегда было частью профессиональной подготовки пилотов и диспетчеров. И это было требованием задолго до введения требований владения общим и авиационным языком. Радиообмен можно использовать в качестве побудительного стимула для проверки навыков говорения и понимания на слух.

Если же тестовое задание состоит из имитации радиообмена, то оценить образец речи выше третьего уровня нельзя. Описание 4 и 5 уровней содержат дескрипторы, которые невозможно применить к радиообмену (употребление вводных слов для организации речи (4 и 5), употребление идиом (5)). Употребление вводных слов и идиом в радиообмене невозможно без нарушения правил ведения радиосвязи. Как оценивается грамматика? Если тестируемый в имитации радиообмена опускает артикли или формальные признаки грамматических конструкций, что можно и нужно делать в радиосвязи, но не является нормой в общем языке, как это оценивать? Задания теста не должны имитировать радиосвязь. Радиообмен может служить только побудительным стимулом для высказывания или контроля понимания.

Понимание различных акцентов — это требование базируется на Шкале ИКАО, в которой сказано, что понимание должно быть точным в большинстве случаев, когда акцент достаточно понятен международному сообществу пользователей. На пятом уровне оценивается понимание не только акцентов, но и диалектов. Для более точного соответствия Шкале ИКАО в заданиях для аудирования нужно использовать не только акценты носителей языка, но акценты людей, для которых английский язык не является родным языком. При этом описание произношения неносителей языка должно соответствовать описанию 4 уровня, скорость речи не должна превышать 170 слов в минуту.

Оценить 6 уровень по Шкале ИКАО не представляется возможным по следующим причинам. Понимание на 6 уровне описывается как «полное понимание почти во всех контекстах, включающее понимание лингвистических и культурных нюансов», что сделать в рамках теста практически невозможно.

Описание взаимодействия тоже должно учитываться разработчиками теста. Для соответствия 6 уровню тестируемый должен «чувствовать вербальные и невербальные намеки и отвечать на них соответственно». Включить невербальные намеки в задания теста невозможно.

Неспособность разработать задания на понимание и взаимодействие на 6 уровне по Шкале ИКАО не должна расстраивать разработчиков теста, это никак не противоречит политике ИКАО в языковом тестировании. В Документе ИКАО 9835 (6.2.8.10) говорится, что

дескрипторы 6 уровня по Шкале ИКАО содержат описание особенностей использования языка, которые выходят за рамки ситуации использования языка в радиотелефонной связи. Шестой уровень — это уровень владения языком образованным носителем языка или двуязычными гражданами, которые считают английский язык одним из родных. Не каждый носитель языка владеет языком на этом уровне, а неносители языка могут владеть языком на высшем уровне только, если они жили в стране языка долгое время, например, ходили в школу и получали высшее образование, получали ученую степень.

Шестой уровень не обязательно подтверждать путем специализированного языкового теста. Высший уровень владения языком обычно подтверждается беседой и документами, свидетельствующими о праве по рождению (*grandfather right*) или о длительном пребывании в стране языка, получении образования и использовании неродного языка в повседневной жизни.

Если мы говорим о языковом взаимодействии, необходимо различать два вида использования языка: *transactional language* и *interactional language* (*G. Buck, 2001*). В первом случае взаимодействие имеет определенную цель: передачу или получение информации, запрос какого-либо действия. Понимание услышанного здесь очень важно. Во втором случае использование языка служит поддержанию социальных связей, содержание сказанного не имеет важного значения, важен сам факт, что что-то говорится. Примером могут служить так называемые *small talks* о погоде, о самочувствии и т. п. В таких разговорах собеседник всегда доброжелателен, большей частью демонстрирует согласие и понимание независимо от того, согласен он или нет, понял или нет.

Обычно в коммуникации можно встретить оба вида использования языка, но в зависимости от ситуации использования языка один вид будет всегда доминантным. В радиосвязи, безусловно, мы имеем дело с *transactional language*. Во всех тестовых заданиях, кроме первого (*warmer*), мы должны оценивать *transactional language*, т. е. способность о чем-то поболтать, а способность передавать и понимать информацию, запросы, указания и команды.

Итак, какие умения и навыки должен проверять квалификационный тест в нашем случае? Ответ содержится в Приказе Росавиации № 353-П (2018), Документе ИКАО 9835 и Циркуляре ИКАО 318. Тест должен проверять способность пилотов и диспетчеров решать коммуникативные задачи при ведении радиосвязи в ситуациях, когда недостаточно предписанной стандартной фразеологии. В задания теста должны быть включены коммуникативные функции и лексические домены, перечисленные в Документе ИКАО 9835, Приложение В, части I – IV.

Задача квалификационного теста – получение образца речи, пригодного для оценки (*ratable speech sample*). Что значит «пригодный образец речи»? Достаточный по времени, он должен содержать спонтанную, а не заученную или шаблонную речь. В тесте должны быть задания, позволяющие оценить все уровни, которые оценивает данный тест. Образец речи не должен подвергаться редактированию.

3.5. Описание конструкта теста

Опираясь на все вышесказанное, попробуем описать возможный конструкт теста:

Тест оценивает способность пилотов и диспетчеров решать задачи, которые возникают в радиообмене, когда недостаточно стандартной фразеологии ИКАО.

Тест базируется на подходе ИКАО к тестированию, определенном в разделе 6.2.8.9 Документа ИКАО 9832 – Руководство по внедрению требований ИКАО к владению языком, который подразумевает «широкую интерпретацию холистических дескрипторов (описательных характеристик) и Шкалы оценок», и направлен на выявление «общего языка по различным темам, которые относятся к радиообмену и авиационной деятельности, без специальной имитации радиообмена». Итак, тест не оценивает владение стандартной фразеологией.

Тест оценивает неподготовленную, а не шаблонную или заученную наизусть речь.

Тест определяет способность экзаменуемого разрешать лингвистические и ситуационные затруднения, которые могут возникнуть в радиосвязи в нестандартных и аварийных ситуациях. Он оценивает способность экзаменуемого разрешать затруднения, используя уточняющие стратегии.

Тест оценивает только речевые навыки, используемые в радиообмене: говорение и понимание на слух. Коммуникативные навыки оцениваются посредством коммуникативных задач. При выполнении тестовых заданий экзаменуемый не должен демонстрировать другие навыки (чтение и письмо) или умение печатать.

Тест оценивает языковые компетенции, соответствующие как минимум четвертому (рабочему) уровню владения языком по Шкале ИКАО.

Тест оценивает все элементы Шкалы ИКАО: произношение, грамматику, словарный запас, беглость речи, понимание на слух и взаимодействие.

Произношение влияет на понимание речи говорящего собеседником.

Грамматическая правильность обеспечивает передачу речевой интенции.

Тест оценивает как словарный запас общего языка, так и знание технической и псевдо-технической лексики.

Беглость речи оценивается в таких параметрах как плавность речи, скорость речи, расстановка пауз, способность продуцировать связные высказывания.

Понимание оценивается во всех заданиях, но в тест включаются аудиозадания с записями акцентов носителей и неносителей языка, акцент которых понятен сообществу пользователей авиационного английского языка.

Понимание аудиозаданий оценивается при помощи устных вопросов контроля понимания, так что эти задания тоже интерактивны.

Тест не оценивает профессиональные знания экзаменуемого и стандартную фразеологию. Задания, которые базируются на анализе ситуации использования языка, требуют общих для пилотов (независимо от типа самолета, которым он управляет) и диспетчеров (независимо от его рабочего места) предметных знаний.

Тест оценивает способность экзаменуемого адекватно и эффективно взаимодействовать при ведении радиосвязи.

3.6. Формат теста. Типы тестовых заданий. Продолжительность теста. Временные пределы тестовых заданий

Определив конструкт теста, самое время подумать о формате теста, количестве заданий, их форме, очередности, продолжительности. Тест оценивает говорение и понимание на слух. Для оценки используется шкала с описательными характеристиками каждого уровня владения языком (*holistic descriptors*). Рейтер, оценивая уровень владения языком, будет сравнивать образец речи экзаменуемого с описанием каждого уровня.

Среди множества тестовых заданий разработчики теста должны выбрать такие, в результате выполнения которых будет получен образец речи. Все задания закрытой формы (*multiple choice, gap filling, matching etc.*) должны быть отклонены. Придется выбирать из заданий открытой формы: дискуссия, интервью, ролевая игра и т. п. Главная цель всех заданий открытого типа – побудить экзаменуемого использовать язык для демонстрации своих навыков. Это может быть описание, рассказ, инструктаж, сравнение, объяснение, оправдание и т. д. В Oral Proficiency Interview (*Proficiency Standards Division, 1999*) экзаменатор просит экзаменуемого описать что-то, например, дом любимой тети, которую экзаменуемый упомянул в своем рассказе. Критерием оценки описания будет то, как легко может слушатель представить этот дом, или, если стимулом служит конкретное фото или картина, то насколько описание соответствует действительности. Это будет пространное монологическое высказывание. При использовании таких заданий, как неуправляемые интервью, качество образца речи сильно зависит от профессионализма экзаменатора.

Из анализа ситуации использования языка мы знаем, что пилоты и диспетчеры не пользуются монологической речью. Радиообмен – это диалогическая речь, поэтому для нашего теста больше подойдет такое тестовое задание большей или меньшей управляемости как структурированное интервью, которое содержит открытые вопросы или вопросы ограниченного радиуса действия и предполагает развернутые ответы (*extended response*). Структурированное интервью состоит из вопросов, которые заранее тщательно продуманы и отобраны. Экзаменатор должен задать все ключевые вопросы. Активно слушая и реагируя на ответы тестируемого, экзаменатор может задать уточняющие вопросы, запросить детали. В интервью не должно быть вопросов с возможными краткими ответами «да» или «нет». Если все-таки следует односложный ответ, за ним должны следовать уточняющие вопросы “Why?”, “How?”, или просьба “Please, explain”. Интервью не должно состоять из набора случайных вопросов и односложных ответов, не связанных друг с другом. «В речевом продукте выше уровня фразы определенные факты или события связаны между собой разными видами отношений, сравниваются друг с другом, противопоставляются, т. е. находятся в причинно-следственной зависимости...» (*Е.И. Пассов, 1985*). Если в ответах эти связи отсутствуют, если тема не получает развития, трудно оценивать беглость речи, ее организацию и, главное, взаимодействие. В одной фразе информативности мало. Учитывая это, остановим свой выбор на структурированном интервью (*structured interview*), потому что это отличный способ «получить максимум релевантной информации в минимально короткий отрезок времени» (*T. Dudley-Evans and M. J. St John, 2006*).

Другое задание открытого типа, ситуация тестового задания (*simulation task*), или ролевая игра (*role play*), может имитировать профессиональный контекст экзаменуемого. В нашей ситуации таким контекстом является ведение радиообмена. Казалось бы, это самый подходящий способ проверить языковую компетентность экзаменуемого, но мы в п. 4 уже подробно описали, почему нельзя имитировать радиообмен. Как же тогда мы можем

использовать ролевую игру в качестве тестового задания? Когда мы обсуждали применение авиационного языка в разных ситуациях, мы сделали выводы, что только в радиообмене применение языка жестко регламентировано, во всех остальных случаях нет таких жестких требований краткости и однозначности, значит, и таких строгих ограничений. Исходя из этого, мы сможем определить, в каких ситуациях, кроме радиосвязи, могут встретиться лексические домены и коммуникативные функции, типичные для радиообмена. Все инциденты и авиационные происшествия всегда обсуждаются, обстоятельства таких событий, их причины, последствия всегда анализируются на брифингах, разборах. Поэтому мы можем использовать подобные формы ролевых игр.

В ролевой игре нужно предусмотреть ситуационные затруднения, непредвиденный поворот событий, чтобы экзаменуемый смог продемонстрировать свои языковые навыки в полной мере. Примеряя на себя роль, экзаменуемый должен выбирать языковые средства, соответствующие ситуации, т. е. что он сказал бы в подобной реальной ситуации. Выбирая ситуацию для ролевой игры, определяя социальные роли и коммуникативные задачи, надо всегда помнить об аутентичности и ситуации, и коммуникативных задач. Нельзя ставить пилоту задачу успокоить плачущую пассажирку и расспросить ее о причине слез, как было в одном из тестов, разработанных носителями языка.

Некоторые тесты используют жестко структурированные ролевые игры с предварительно записанными репликами одной из ролей. Для проверки фразеологии радиообмена такой вариант подошел бы, учитывая стереотипность радиообмена, но для оценки спонтанной речи при ситуационных затруднениях это никак не подходит. Описывая ситуацию, разработчики должны представлять сценарий развития событий, то есть для экзаменатора должна быть четко прописана легенда.

Несмотря на то, что оценка понимания на слух проводится на протяжении всего теста, в тесте должны быть отдельные задания на аудирование. Во-первых, чтобы добавить для понимания акценты, отличные от акцента экзаменатора, во-вторых, чтобы оценить понимание на слух происходящего в радиоэфире, с помехами и искажениями, т. е. реального радиообмена. Если в тест включаются фрагменты радиообмена для оценки понимания, то это категорически не должны быть вырванные из контекста отдельные предложения. Это должна быть законченная ситуация. Учитывая, что для оценки понимания мы применяем Шкалу ИКАО, то оценивать мы должны понимание не как чисто рецептивный навык. Понимание в ситуации радиообмена всегда связано с реакцией на сказанное: диспетчерское указание – подтверждение указания в виде квитанции – выполнение указания или невыполнение с объяснением причины; запрос информации – передача информации и т. д. Это означает, что и проверка понимания должна быть интерактивной. Для этой цели как нельзя лучше подходят вопросы контроля понимания (*comprehension questions*). Разработчики должны описать подробно, как коррелируются вопросы для контроля понимания с описательными характеристиками каждого уровня понимания в Шкале ИКАО.

Выбрав типы тестовых заданий, можем приступить к определению формата теста. Последовательность заданий должна соответствовать принципу «от простого к сложному». Проведение теста – это пробинг, т. е. *level check*. Каждое задание должно быть сложнее предыдущего. У экзаменуемого должна быть возможность продемонстрировать свой высший уровень владения языком. Если в тесте нет задания, которое позволяет экзаменуемому продемонстрировать умение вести диалог, организовать свою речь при помощи вводных слов, использовать идиомы в своей речи,

стилистически подстраивать свою речь в соответствии с ситуацией, то как можно оценить его уровень владения языком как 5 уровень по Шкале ИКАО?!

Конечно, выбор тестовых заданий, их последовательность, временные рамки – это выбор каждого разработчика тестирования, но любой выбор должен быть аргументирован. Аргументы должны быть убедительными для органов, которые принимают решение о лицензировании или аккредитации теста.

Попробуем обосновать формат гипотетического теста.

Почти все коммуникативные тесты начинаются с так называемого *Warmer*. В русском языке встречаются разные переводы этого термина: речевая разминка, разогрев. И то и другое подходит к характеристике этой части теста. Описывая *interactional language*, мы обращали внимание на неформальный характер подобного общения, на дружелюбность атмосферы подобных бесед. Для начала теста такая форма использования языка как раз подходит – это установление эмоционального контакта. Начинаем с приветствия, представления, просим рассказать о себе. Необходимо помнить, что вопросы не должны касаться личной информации. В этом задании мы получим ответы большей частью в виде домашних заготовок, т. е. заученных фраз.

Продолжим наш пробинг. Логично было бы в следующем задании коснуться лексических доменов и коммуникативных функций общего языка. Эти темы и функции перечислены в Документе ИКАО 9835. Какой тип задания может здесь быть уместен? Пусть это будет структурированное интервью. Вопросы должны быть связаны между собой, служить поводом для полного связного высказывания.

Третье задание должно быть сложнее, можем перейти уже к авиационному контексту. Это тоже может быть структурированное интервью с открытыми вопросами и вопросами ограниченного радиуса действия (описание чего-либо) с учетом предметных знаний. Сколько заданий для аудирования, в каком виде они будут представлены в тесте, конечно, на эти вопросы отвечают разработчики. Выбирая форму и количество заданий, важно понимать, как вы будете потом оценивать понимание. Если у вас несколько текстов для аудирования, несколько фрагментов радиосвязи, как выстраивается приоритетность в оценке? Например, один фрагмент понят полностью, а второй не совсем или совсем не понят?

Нам представляется, что самым сложным и, потому, завершающим, должно быть задание на взаимодействие, например, ролевая игра. Если разработчики заявляют, что тест оценивает 5 уровень по шкале ИКАО, без подобного задания не обойтись.

Определившись с форматом, переходим к временным пределам теста и каждого тестового задания. Исходить здесь необходимо из двух обязательных условий: образец речи должен быть достаточным по времени, но при этом тест должен быть практичным.

3.7. Разработка банка заданий. Спецификации тестовых заданий

Помня о том, что «выполнение экзаменуемым тестовых задач может быть интерпретировано как свидетельство его способности выполнять задачи в ситуации использования языка (*D. Douglas, 2002*), приступаем к созданию банка тестовых заданий.

Каждое задание предваряется инструкцией (*task rubric*), тестируемый должен точно понимать, ЧТО он должен делать при выполнении задания. Инструкция должна быть

четкой и понятной. Переход от одного задания к другому должен быть ясно обозначен экзаменатором.

Warmer может быть одинаковым во всех вариантах. Для составления банка остальных заданий необходимо привлечь команду сценаристов. Скорее всего с написанием сценариев справятся люди, знакомые с ситуацией использования языка, имеющие опыт преподавания авиационного английского языка и радиосвязи. Креативность, воображение, интуиция помогут в такой работе. Найти и отобрать подобных специалистов очень сложно, не каждый преподаватель может работать творчески. Если сценаристы обладают достаточным багажом предметных знаний, можно обойтись на этом этапе без привлечения пилотов и диспетчеров, если же нет, то без консультаций с профессионалами не получится отобразить в тесте авиационный контекст.

Тестовые задания в разных вариантах теста должны быть одинаковой сложности, одинаковыми по времени и структуре, оценивать одинаковые языковые навыки. Для того, чтобы все параллельные варианты тестовых заданий были валидны и одинаковой сложности, для всех заданий должны быть созданы спецификации. Прежде чем перейти к написанию подробных спецификаций, нужно соотнести лексические домены и коммуникативные задачи с описанием уровней в Шкале ИКАО. Разработчики теста должны все время иметь в виду, что чем выше уровень оценивается, тем шире становится список ситуаций в тесте, новые задания прибавляются к тем, которые уже были на предыдущем уровне. В таблицах приведены примеры корреляции лексических доменов и коммуникативных функций со Шкалой ИКАО, это так называемая “elicitation aid” (R. Chatham, из личной переписки), которая служит подсказкой сценаристам и экзаменатору, как в процессе пробинга побудить экзаменуемого продемонстрировать те или иные коммуникативные навыки и лексический запас.

Topics

| Level 1 | Level 2 | Level 3 | Level 4 | Level 5 |
|--|--|---|---|--|
| <ul style="list-style-type: none"> • greetings • introductions | <ul style="list-style-type: none"> • self • colors • time • days of the week • months of the year • dates • family • weather | <ul style="list-style-type: none"> • family • jobs • geography • everyday activities • own background • work • experiences • future plans and expectations • concrete topics • past events • limited work requirements | <ul style="list-style-type: none"> • professional topics, including answering objections, clarifying points, or justifying decisions | <ul style="list-style-type: none"> • topics pertinent to personal and professional needs or experience • complex situations related to professional activities |

Communicative tasks

| Level 1 | Level 2 | Level 3 | Level 4 | Level 5 |
|---------|---|--|--|---|
| | <ul style="list-style-type: none"> • minimal conversation with rehearsed content | <ul style="list-style-type: none"> • simple conversation • request information concerning present, past and future facts • share information concerning present, past and future facts • simple description • request clarification | <ul style="list-style-type: none"> • description (in concrete terms) • directions or instructions • report facts • express necessity and feasibility • relay information • handle situations with a complication | <ul style="list-style-type: none"> • description in detail • hypothesis • dealing with unfamiliar topics and situations • resolve problem situations • interpreting task |

В спецификации каждого тестового задания необходимо указать следующие характеристики задания: время, отведенное на это задание, характеристики инструкции к заданию, ввода данных, ожидаемого ответа и взаимоотношений между вводом данных и ожидаемым ответом, метода оценки. Почему это очень важно? «Потому что мы должны быть уверены, что специальный, или профильный контекст и тест обладают общими существенными характеристиками, и таким образом выполнение тестовых заданий тестируемым может быть интерпретировано как свидетельство его способности выполнять задания в ситуации использования языка вне теста» (D. Douglas, 2002).

L. Bachman (1990) утверждает, что, чем больше черт ситуации использования языка присутствует в тестовых заданиях, тем более «аутентично» эти задания будут выглядеть для тестируемого.

В спецификации необходимо описать инструкцию к заданию (*task rubric*), т. е. процедуру выполнения задания. Из инструкции к заданию тестируемый должен понять, какой тип задания ему предстоит выполнить: описание картинки, интервью, аудиозадание или ролевую игру.

Далее описываем характеристики ввода (*input*), т.е. материала, который служит стимулом для ответа. Содержание ввода ориентирует тестируемого в выборе соответствующих языковых средств для своего ответа. Материалы ввода могут быть визуальными, устными, физическими объектами и т. п.

Следующим разделом спецификации будет описание ожидаемого ответа (*expected response*), его формат и содержание. Содержание ответа должно позволять измерить заявленный конструктор теста. Конечно, ответ тестируемого может отличаться от ожидаемого разработчиками теста, поэтому необходимо уделить достаточное внимание ясности инструкции к заданию и содержанию ввода. Чем четче будет сформулирована инструкция, чем больше содержание ввода будет отражать ситуацию использования языка, тем больше ответ тестируемого будет соответствовать ожиданиям разработчиков теста.

Перейдем к отношениям между вводом и ответом. Эти отношения характеризуются наличием или отсутствием взаимосвязи, взаимосвязь может быть косвенной или прямолинейной (*reactivity and directness*).

Например, в разговоре эти отношения в высокой степени взаимосвязаны (*reciprocal*). Говорящий всегда получает реакцию слушателя об эффективности своего высказывания вербально, при помощи мимики и жестов, повышением голоса и т. п. Бывают случаи, когда никакой взаимной связи нет. Например, когда при телефонном звонке включается автоответчик, при чтении, при прослушивании аудиозаписи – в этих случаях нет никакой реакции, взаимодействие стремится к нулю. При оценке таких навыков как говорение и понимание на слух необходимо иметь в виду, что чем выше степень реактивной активности, или взаимосвязи, между вводом и ответом, тем динамичнее и полнее ответ, тем более образец речи пригоден для оценки.

Прямолинейность (*directness*), отвечает за то, будет ответ прямо или косвенно зависеть от ввода. Если в содержании ввода уже имеются какие-либо данные для ответа, ответ будет прямым. Например, прослушайте текст и ответьте на вопросы. Или вопрос: Can oxygen bottles be transported together with barrels with lubricants? – предполагает прямой ответ. Отношения между вводом и ответом – прямолинейные. Другой пример: Describe your actions in case your cabin crew discover a stowaway on board. В этом случае нельзя сказать, что отношения между вводом и ответом прямолинейные, их можно назвать косвенными. Задание предполагает использование специфической лексики и предметных знаний, которые не содержатся во вводе.

В спецификации тестового задания необходимо указать метод оценки ответа тестируемого.

Ознакомьтесь с примером спецификации тестового задания:

| 3. INTERVIEW, AVIATION ENGLISH | |
|--|--|
| Time allotment | 6 minutes |
| Characteristics of the task rubric | |
| Language | English language |
| Channel | Aural |
| Characteristics of the input | |
| Format | |
| Channel | Aural |
| Form | Language |
| Language | English language |
| Type | Structured topic-based interview |
| Vehicle | Live, face-to-face |
| Characteristics of the expected response | |
| Format | |
| Channel | Aural |
| Form | Language |
| Language | English language |
| Type | Extended production |
| Language skills assessed | Speaking and listening |
| Language elements assessed | Pronunciation, Grammatical knowledge (form and meaning), Vocabulary range sufficient to communicate on topics which characterize day-to-day radio communications, Fluency, Interactional comprehension, Interactions. |
| Scoring method | Analytical and holistic scoring |
| Relationship between the input and response | |
| Reactivity | Reciprocal |
| Directness | Indirect, semi-direct |

Это спецификация интервью на авиационную тематику. На это задание отводится 6 минут. Инструкция к заданию представлена на английском языке в устном формате. Вопросы задаются на английском языке для восприятия на слух. Структурированное интервью состоит из вопросов, тематически связанных между собой. Задание выполняется в режиме реального времени, без посторонних. Характеристика ожидаемого ответа: устный развернутый ответ на английском языке. В ответе оцениваются навыки говорения и понимания на слух, а также следующие языковые элементы: произношение, грамматика (форма и значение), достаточность словарного запаса для ведения радиосвязи, беглость речи, интерактивное понимание и взаимодействие. Метод оценки – аналитический и холистический. Взаимоотношения между вводом данных (вопросами) и ответами – взаимозависимые, косвенные и полупрямые.

Приводим в качестве примера часть анализа ситуации использования языка и анализа потребностей и пример тестового задания, которое разработано с учетом анализа ситуации использования языка и анализа потребностей на основе выше приведенной спецификации.

Аудиозапись реального радиобмена

[audio mp3="https://aeroenglish.ru/wp-content/uploads/2021/02/N130TM-for-analysis.mp3"][/audio]

Target language use analysis and needs analysis

| | |
|-------------------------|--|
| Recording: | communication breakdown in marginal weather |
| Deleted: | pauses and communication with other stations |
| <i>Legend:</i> | |
| P | Pilot(s) |
| C1 | Controller, Vnukovo Precision |
| C2 | Controller, Vnukovo Radar |
| C3 | Controller, Domodedovo Radar |
| C4 | Vnukovo Radar |
| /?/ | unintelligible |
| Text in blue: | Standard Phraseology (ICAO Doc 4444) and subsidiary phraseology (used but not mentioned in ICAO Doc 4444) |
| Text in red: | General English in Aviation Context |
| <u>Underlined text:</u> | Pay attention |
| <i>Comments:</i> | <i>In magenta italics</i> |

| | |
|-------------------------------|---|
| Aircraft registration: | N130TM |
| Year: | 2003 (<i>ICAO Language Proficiency Requirements were introduced in 2008</i>) |
| ATC position(s): | Vnukovo Radar Vnukovo Precision Domodedovo Radar Vnukovo Radar |

Transcript

| No | P/C | Message | Comments | Communicative Functions |
|----|-----|---|--|--|
| 1 | P | Okay, Moscow, November one three zero Tango Mike is heading two four zero, descending down to <u>five thou er</u> to one thousand eight hundred meters. | <i>"Down" is redundant and not specified in ICAO Doc 4444</i> <i>The first sign that the crew mixes SI and metric units of measurements</i> | Inform ATC about heading and descent. |
| 2 | C1 | November one three zero Tango Mike, Vnukovo Radar, turn right, heading three <u>three</u> zero, descend to level one thousand two hundred meters. | | Respond to initial call. Give heading and descent instructions. |
| 3 | P | ...three zero, down to one thousand two hundred meters, roger. | <i>ICAO wording: "descending to". Failure to use call sign in acknowledgement</i> | Receive and confirm heading and descent instructions. |
| 4 | C1 | November one three zero Tango Mike, descend to height six hundred meters, QFE niner seven <u>niner</u> millibars. | | Give descent instructions. Give data for altimeter setting. |
| 5 | P | Roger, nine seven <u>niner</u> millibars, and we're down to six hundred meters, roger. | <i>ICAO wording: "descending to". Failure to use call sign in acknowledgement</i> | Receive and confirm descent instructions and data for altimeter setting. |

| | | | | |
|----|----|--|---|--|
| | | | <i>Failure to use call sign in acknowledgement (repetitive in further communications)</i> | and data for altimeter setting. |
| 6 | C1 | November one three zero Tango Mike, be informed, ceiling seven zero meters. | | Give information about cloud base. |
| 7 | P | Yep, seven zero meters. | | Receive information about cloud base. |
| 8 | C1 | November one three zero Tango Mike, cleared ILS approach for runway zero six, turn final, descend to height four hundred meters, contact Vnukovo Precision, one one eight decimal three. | | Issue approach clearance. Instruct to turn final. Give descent instructions. Hand the flight over to another frequency. Name the ATC unit and frequency. |
| 9 | P | One eight decimal three, good day. | <i>Failure to acknowledge the type of approach, final turn and height</i> | Acknowledge the frequency. |
| 10 | C2 | November Tango Mike. | | Call the aircraft. |
| 11 | P | And one three zero Tango Mike on a missed approach, climbing to twenty six hundred. | <i>Wrong units of measurement</i> | Inform about a missed approach. Inform about climbing. |
| 12 | C2 | Nov... | | |
| 13 | P | ...one three zero Tango Mike, how do you read? | | Query about ability to receive transmissions. |
| 14 | C2 | November one three zero Tango Mike, Vnukovo Radar, turn right, heading one four zero, er <u>descend and maintain</u> six hundred meters height, QFE nine... niner seven eight millimeters. | <i>The pilot busted the height assigned in the go-around procedure</i> | Give heading and descent instructions. Give data for altimeter setting. |
| 15 | P | Heading one four zero, and <u>say again</u> the descent. | <i>N130TM is not sure of height values, probably because of mixing feet with meters</i> | Acknowledge the heading. Request to repeat data for altimeter setting. |
| 16 | C2 | November Tango Mike, descend to six zero zero meters height, QFE niner seven eight millim... bars. | | Repeat descent instructions and data for altimeter setting. |
| 17 | P | Nine seven eight millibars. | <i>N130TM still does not confirm the height value</i> | Acknowledge data for altimeter setting. |
| 18 | P | Er, you gonna take us to our alternate, sir? | <i>N130TM demonstrates unfamiliarity with the UUWW go-around procedure</i> | Query about controller's intention. |

| | | | | |
|----|----|--|--|--|
| 19 | C2 | Would you say again? | <i>"Would you" is redundant</i> | Request to repeat. |
| 20 | P | We request, ah vectors and weather at our alternate. | | Request vectoring and weather information. |
| 21 | P | ...oscow, one three zero Tango Mike, we never did see the field, we er request vectors to our alternate Uniform Uniform Delta Delta. | | Explain reason for a go-around. Request vectoring to alternate airdrome. |
| 22 | C2 | Ah, November one three zero Tango Mike, roger, continue present heading, maintain six zero zero meters. | | State understanding. Inform about vectors and level. |
| 23 | P | ...ix zero zero meters, roger, we're level. | | Acknowledge level flight. |
| 24 | P | And we need the weather for Uniform Uniform Delta Delta. | <i>ICAO: "Request weather at..."</i> | Request weather at alternate. |
| 25 | C2 | Er, stand by one. | | Instruct to wait for a call. |
| 26 | C2 | Er, November one three zero Tango Mike, ah report reason your divert. | <i>Incorrect wording. The controller cannot understand why the aircraft was unable to land in acceptable cloud base and visibility minimums</i> | Ask to clarify reasons for a diversion. |
| 27 | P | Oh we have no, we couldn't get in to Vnukovo, so we need to reroute, we made a missed approach and we need to go where the weather is better, where are you sending people? There must not be anyone got into, er Vnukovo. | <i>"Vnukovo" is mispronounced here and in all further messages. Multiple false starts, inability to state intentions or request information using ICAO phraseology, both here and in further transmissions</i> | Explain reasons for a go-around. Request rerouting to an aerodrome with better visibility. Query about the aerodrome where flights are directed. Doubt a possibility of landing at destination. |
| 28 | P | Are you getting people into Vnukovo this, this evening? | <i>Absence of any response from the controller is a failure of ATC to understand General English both here and in further communications</i> | Query about a possibility of landing at destination. |
| 29 | P | Moscow, one three zero Tango Mike. | | Call the ATC. |
| 30 | C2 | Er, stand by one. | | Instruct to wait for a call. |
| 31 | C2 | November one three zero Tango Mike, are you going to land at Uniform Uniform Delta Delta? | | Query about an intention to land at alternate. |

| | | | | |
|----|----|---|--|---|
| 32 | P | Yes sir, if you can give me the weather, er we need to go somewhere where we can land, we don't have a lot of time, we just came from Japan. | <i>Inability to use ICAO standard phraseology and request information both here and in further transmissions</i> | Answer affirmatively. Request weather information. Inform on time restrictions. Inform on the country where the flight originated. |
| 33 | P | Are you sending the aircraft, who are trying to land at Vnukovo, where are you sending them tonight? | | Query about the aerodrome where flights are directed. |
| 34 | C2 | November one three zero Tango Mike, contact Domodedovo Radar, one two seven decimal seven. | | Hand the flight over to another frequency. Name the ATC unit and frequency. |
| 35 | P | ... two seven seven, thanks. | | Acknowledge the frequency. |
| 36 | P | Moscow Approach, November one three zero Tango Mike, we're at six hundred meters. | <i>Failure to use ICAO standard phraseology.</i> | Call ATC, inform about height. |
| 37 | C3 | November one three zero Tango Mike, ah Domodedovo Radar, good evening, aaah, turn left, turn left, heading ah zero four zero, descend to height ah five zero zeme... zero zero meters by QFE niner eight two millibars. | | Respond to initial call. Give heading and descent instructions. Give data for altimeter setting. |
| 38 | P | Niner eight two millibars, down to five hundred meters, and where are you taking us? | <i>Failure to use ICAO standard phraseology.</i> | Acknowledge data for altimeter setting. Receive descent instructions. Query about controller's intention. |
| 39 | C3 | November one three zero Tango Mike, are you, do you want to land at Domodedovo? | | Query about pilot's intention. |
| 40 | P | Is the weather at Uniform Uniform Delta Delta, er is, what is the weather at Uniform Uniform Delta Delta? | <i>Failure to use ICAO standard phraseology.</i> | Request weather information at alternate. |
| 41 | C3 | November one three zero Tango Mike, roger, runway-in-use one four left, er maintain five thousand, er five, er hundred meters, er heading three six zero, turn left. | | Inform about runway-in-use. Give descent and heading instructions. |
| 42 | P | Okay, heading three six zero and we're going to one four left, would you please give me the ILS frequency. | <i>Compare with standard phraseology:</i> | Receive information about runway in use. Request navaid frequency. |

| | | | | |
|----|----|---|--|--|
| | | | <i>"... approaching one four left, request the ILS frequency".</i> | |
| 43 | C3 | Stand by. | | Instruct to wait for a call. |
| 44 | P | And Moscow, for one three zero Tango Mike, what is the weather at Uniform Delta Delta, the current weather at Uniform Delta Delta? | <i>ICAO: "Request weather..."</i> | Request weather information at alternate. |
| 45 | C3 | November one three zero Tango Mike, ah actual weather for aah Uniform Uniform Delta Delta: visibility six hundred meters, ah RVR one thousand four hundred meters, fog, vethe... vertical visibility eight zero meters, runway wet, braking er coefficient err decimal six. | | Give weather information and runway state. |
| 46 | C3 | November one three zero Tango Mike, heading three two zero. | | Give heading instructions. |
| 47 | P | Heading three two zero, Tango Mike, d'you, do you have traffic heading into the airport at this time sir? | | Acknowledge heading instructions. Query about presence of other aircraft approaching the airport. |
| 48 | P | And do you have the ILS frequency for one four left sir? | <i>ICAO: "Request ILS frequency runway one four left".</i> | Request navaid frequency. |
| 49 | C3 | Er, November one three zero Tango Mike, stand by, I will find out. | | Instruct to wait for a call. Inform about intentions. |
| 50 | P | Thank you. | | |
| 51 | C3 | November one three zero Tango Mike, er turn right, heading zero niner zero. | | Give heading instructions. |
| 52 | P | Zero nine zero, yes sir, right turn, and we need to go to an airport that we can get into er as soon as we can. | | Acknowledge heading instructions. Transmit urgency of the situation. |
| 53 | P | The weather at Uniform Uniform Delta Delta is marginal, and do you have another field close by that has better visibility? | <i>Failure to use ICAO standard phraseology.</i> | Comment on weather. Query about another aerodrome. |
| 54 | C3 | November one three zero Tango Mike, frequency, [ils] | | Give navaid frequency. |

| | | | | |
|----|----|---|---|---|
| | | frequency Uniform Uniform Delta Delta one zero eight decimal five. | | |
| 55 | P | One zero eight decimal five, and the inbound course sir, (pause) ...bound course? | <i>Failure to use ICAO standard phraseology.</i> | Receive navaid frequency. Request inbound course. |
| 56 | C3 | Land course er one three seven, now turn right, heading one three seven, you are on errr final. | | Inform about course. Give heading instructions. Inform about current position. |
| 57 | P | Okay, one three seven. | | Receive course information. |
| 58 | C3 | November one three zero Tango Mike, you are cleared ILS approach, er freque... errr runway one four left, now you are on final. | <i>Numerous signs of uncertainty both here and in further transmissions</i> | Issue approach clearance. Inform about current position. |
| 59 | P | Okay, you got the... roger, cleared on... | | |
| 60 | P | We gonna need, we gonna need some vectors to go around, er we're asking you if the weather is, are people landing at the airport sir? | <i>Failure to use ICAO standard phraseology.</i> | Request vectoring. Request a go-around. Request weather information. Request information whether the aerodrome operating normally. |
| 61 | P | Moscow Approach, November one three zero Tango Mike, we need vectors back to final, we just got the frequency, we were not set up for the approach. | | Request vectoring. Inform about a missed approach (in the Past). |
| 62 | C3 | November one three zero Tango Mike, errr say again, please, er you are now to, you are left from land cross. | | Request repetition. Inform about present position. |
| 63 | P | We cannot make the approach at this time sir, we're not set up, we need vectors back for final, thank you. | <i>ICAO: "Unable approach, request vectors back for final".</i> | Inform about inability for a maneuver and provide reason. Request vectoring. |

Etc.

Target Language Use Analysis

The recording (made in 2003) exemplifies a situational and, thus, linguistic complication in marginal weather conditions. The pilot, an EFL speaker, uses General English instead of both standardized ICAO phraseology and when ICAO phraseology does not suffice. Rather often, his language possesses features of spoken language: "gonna", "wanna"; direct word order in interrogative sentences ("Er you gonna take us to our alternate sir"); false starts ("We gonna need

we gonna need some vectors to go around er we're asking you if the weather is are people landing at the airport sir"), etc.

Inability of the pilot to use metric measuring units and correctly translate them into feet (probably because of fatigue after a long-haul flight) resulted in a misunderstanding when relatively good visibility was erroneously taken for a value below pilot's (or company) minimums. The situation was aggravated by controllers' inability to identify the misunderstanding that was caused by very limited ability to understand and use General English. An accident similar to Avianca Flight 052 that crashed in poor weather conditions on Long Island after running out of fuel was narrowly avoided.

The recording demonstrates general proficiency of both pilots and ATC officers in ICAO standard phraseology. However, both almost fail when using General English in aviation context (pilots – use of colloquialisms, inability to rephrase; controllers – inability to understand and use General English in aviation context). Inability of the pilot to follow standardized phraseology has become the starting point of communication breakdown.

Language to be tested

1. The object of testing should be General English in aviation context, not ICAO standardized phraseology. If ICAO standardized phraseology is to be tested, the best place for doing it is a flight/ATC simulator. Simulation of radio communications in routine situations, urgencies and emergencies had been a part of RTF training course and exit testing at Moscow ATC Center since 1983, long before ICAO language proficiency requirements were developed. However, language used at the simulator follows the requirements of RTF register that cannot be evaluated by the ICAO LPR Scale.
2. Basing on this recording, test tasks should incorporate the following topics (domains): marginal weather, captain's minimums, aerodrome minimums, units of measurement, lack of fuel, diversion.
3. Basing on this recording, the following communicative skills should be tested:

Pilots

- Request/receive information concerning weather conditions.
- Receive/acknowledge current weather information and forecast.
- Monitor/acknowledge meteorological broadcasts (ATIS, VOLMET).
- Request/transmit meteorological information en-route or in a certain point of the route.
- Transmit pilot's intentions.
- State the problem.
- Request appropriate assistance.
- Request rerouting or diversionary landing.
- Advise controller of remaining fuel.

Controllers

- Brief pilot on weather data in specified format.
- Issue significant weather advisory or update to others.
- Formulate weather broadcast.

- Evaluate and inform pilot of alternate routes based on weather, aeronautical information, pilot preference, and pilot/aircraft limitations.
 - Receive/acknowledge information on the problem.
 - Receive information/acknowledge information/query about intention.
 - Receive request for appropriate assistance; acknowledge it.
 - Receive request/give instructions and appropriate information for rerouting or diversionary landing.
 - Request/ receive information of remaining fuel.
 - Give instruction to monitor meteorological broadcasts (ATIS, VOLMET).
 - Request/transmit meteorological information en-route or in a certain point of the route.
4. Basing on this recording, the following grammatical structures appropriate for communicative tasks should be tested:
- Modal verbs (can, may, must, have got to, should, ought to, would, could, might, needn't, don't have to, mustn't);
 - Numbers (cardinal and ordinal);
 - Passive voice (Simple Present, Simple past);
 - Question words for describing things and for requesting information (What? Who? Which? When? Why? Where? How?);
 - Tenses (Present Simple, Present Continuous, Simple Future (will) and Future – to be going to);

Basing on this recording, the following vocabulary needed for communicative tasks should be tested: weather, marginal weather, visibility, cloud base, ceiling, weather minimum, meters, feet, divert, fuel, short of fuel, low on fuel, etc.

На основании результата анализов разработано следующее задание:

Interview. Aviation English (6 min)

Now we are going to discuss Aviation English issues. Let's talk about weather phenomena and rerouting.

- Could you speak about weather phenomena that affect flight operations?
- What do you know about captain's and aerodrome minimums?
- What will the pilot do if the visibility and cloud base at his destination are below minima?
- Please recall other situations when the crew divert to an alternate.

This is the end of Part 3.

Экзаменатор обязательно должен задать вопросы из тестового задания. Но учитывая, что экзаменатор должен активно слушать ответ тестируемого, он может и должен задать дополнительные вопросы, что-то уточнить. Если в ответе на первый вопрос тестируемый забыл некоторые метеоявления, можно спросить про них. Вопрос про минимумы можно персонализировать, т. е. спросить капитана о его минимуме или минимуме аэродрома базирования, диспетчера – о минимумах аэродромов в его воздушной зоне. Последний пункт может быть расширен вопросом про подобные случаи из предыдущего опыта тестируемого или его коллег. Таким образом, в ходе выполнения задания мы получим распространенный ответ с неподготовленной речью, позволяющий сделать предположение о том, справится ли тестируемый с языковыми и ситуационными затруднениями при управлении воздушным судном или управлением воздушным движением в сложных метеорологических условиях вне теста. В задании отражены и

лексические домены, и коммуникативные задачи, а полученный образец позволит оценить все элементы Шкалы ИКАО.

Еще один пример спецификации тестового задания.

| AUDIO TASK 2 | |
|--|--|
| Time allotment | 5 minutes |
| Characteristics of the task rubric | |
| Language | English language |
| Channel | Aural |
| Characteristics of the input | |
| Format | |
| Channel | Aural |
| Form | Language |
| Language | English language |
| Type | Real radio communication extract played two times and comprehension questions |
| Vehicle | Reproduced and live, voice-only |
| Characteristics of the expected response | |
| Format | |
| Channel | Aural |
| Form | Language |
| Language | English language |
| Type | Limited production and extended production |
| Language skills assessed | Listening comprehension and speaking |
| Language elements assessed | Pronunciation, Grammatical knowledge (form and meaning), Vocabulary range sufficient to communicate on topics which characterize non-routine and emergency radio communications, Fluency, Interactive comprehension. |
| Scoring method | Analytical and holistic scoring |
| Relationship between the input and response | |
| Reactivity | Reciprocal |
| Directness | Direct |

Это – спецификация аудиозадания. На него отводится 5 минут. Экзаменатор предваряет задание устной инструкцией на английском языке. Фрагмент реальной радиосвязи на английском языке воспроизводится дважды, потому что при реальном ведении радиосвязи всегда существует опция запроса повторения передачи (*Say again*). Контроль понимания проверяется при помощи вопросов контроля понимания, заданных в режиме реального времени. Задание выполняется при отсутствии визуального контакта. В результате выполнения этого задания должны быть получены устные ответы ограниченного радиуса действия и развернутые ответы на английском языке. В этом задании оцениваются понимание на слух и говорение, а также такие элементы Шкалы ИКАО как произношение, грамматика (форма и значение), словарный запас, необходимый для коммуникации в нестандартных и аварийных ситуациях при ведении радиосвязи, беглость речи и интерактивное понимание. Метод оценки – аналитический и холистический. Взаимоотношение между вводом данных (аудиозапись и вопросы) и ответом – взаимозависимые, прямые.

Пример аудиозадания:

Audio task 2 (5 min)

Now you are going to listen to a recording. You can take notes. You will listen to the recording twice. After you have listened to the tape, you will be asked several questions. Do you understand? The call sign is Speedbird 873.

Аудиозапись реального радиообмена для аудиозадания

[audio mp3="https://aeroenglish.ru/wp-content/uploads/2021/02/Speedbird-873.mp3"][/audio]

Вопросы контроля понимания можете придумать самостоятельно.

Такие спецификации разрабатываются к каждому заданию. После этого сценаристы могут приступить к созданию банков заданий. На начальном этапе их не должно быть много, так как в процессе испытания бета-версии могут быть внесены изменения в формат теста, изменены задания и тайминг.

3.8. Испытание бета-версии теста

Первые опытные версии заданий могут быть опробованы на преподавателях, которые обучают пилотов и диспетчеров. После создания бета-версии ее необходимо опробовать на целевой аудитории теста, т. е. пилотах и диспетчерах. В популяцию для испытания бета-версии должны быть включены экзаменуемые с разным уровнем, который может быть установлен по результатам обучения или каких-либо предыдущих экзаменов. Количество экзаменуемых должно быть репрезентативным (чем больше группа испытуемых, тем результат достовернее). Все участники испытания сдают несколько разных вариантов с разными экзаменаторами и рейтерами. После прохождения каждого варианта заполняется опросник с вопросами о понятности инструкции к заданиям, о достаточности времени для выполнения заданий, о степени трудности параллельных форм заданий, о степени аутентичности заданий. Такие же анкеты заполняются экзаменаторами и рейтерами. После изучения статистических данных принимается решение о внесении изменений в тест и задания или оставления формата без изменений, что, кстати, бывает крайне редко.

4. Подготовка экзаменаторов и рейтеров

В зависимости от численности популяции теста надо определить количество персонала, необходимого для функционирования теста: технический персонал, администрация, экзаменаторы и рейтеры.

Отбор экзаменаторов и рейтеров представляет собой очень непростой процесс. Поскольку тестирование по Шкале ИКАО – это квалификационное тестирование, ответственность людей, вовлеченных в тестирование, очень высокая. Это должны быть люди, владеющие английским языком не ниже 5 уровня, обладающие достаточным объемом предметных знаний, хорошо знакомые с ситуацией использования языка и, конечно, с высокими моральными качествами. По данным профессора R. Chatham (из личной переписки), во время его работы в Лингвистическом институте министерства обороны США только 30

процентов кандидатов в рейтеры OPI (Oral Proficiency Interview) после прохождения подготовки успешно сдавали выпускной экзамен и получали свидетельство рейтера.

Численный состав команды экзаменаторов и рейтеров зависит и от процесса тестирования. Разработчики должны принять решение о том, сколько экзаменаторов/рейтеров принимают участие в получении образца речи (один или два), сколько оценивают образец речи.

Программа подготовки рейтеров и экзаменаторов включает в себя теоретическую и практическую части.

Теоретическая часть состоит из знакомства с Этическим кодексом Международной ассоциации языкового тестирования, со Шкалой ИКАО, форматом теста, техническими приемами получения образца речи, пригодного для оценки, критериями оценки образца речи и каждого элемента языка в отдельности.

Практическая часть состоит из ознакомления с образцами речи каждого уровня и оценки речевых образцов, из проведения тестов для получения образца речи, обсуждения неправильного поведения экзаменатора во время проведения теста, оформления оценочных листов и сопроводительной документации, обучения работы с оборудованием, применяемым для проведения теста. Главные задачи курса — это (1) научить экзаменаторов получать образец речи, пригодный для оценки и (2) добиться от рейтеров адекватной, заслуживающей доверия оценки образца речи. Совершенно очевидно, что такой объемный курс не может уложиться в 20 академических часов. Изучение теоретической части и отработка практических навыков получения образца речи и адекватной его оценки требуют, по нашему опыту, не менее 70 часов.

Для экзаменаторов и рейтеров должны быть написаны пошаговые инструкции проведения теста и работы с образцами речи, в которые должны быть включены брифинг перед началом теста, подтверждение личности экзаменуемого, заполнение сопроводительной документации, мероприятия, обеспечивающие сохранность материалов тестирования, механизмы взаимодействия с администратором тестирования или другим ответственным лицом.

4.1. Инструкции для экзаменаторов/рейтеров по получению пригодного для оценки образца речи

Мы уже писали о том, что качество образца речи очень во многом зависит от способностей экзаменатора/рейтера его получить. Обучение экзаменаторов/рейтеров затратно по времени, а, следовательно, и финансово, однако на этом категорически нельзя экономить. Экзаменатор должен понимать, от чего зависит качество образца речи, и как получить пригодный для оценки образец речи.

Доброжелательная, сдержанная манера экзаменатора индуцирует аналогичное поведение тестируемого. Обращаться к тестируемому надо так, как он представился. Во всех частях теста, кроме тех, что проводятся без визуального контакта, должен быть визуальный контакт с тестируемым, особенно в начале теста. Экзаменатор должен избегать как оценочных суждений, так и конфронтационных вопросов.

Экзаменаторы – это большей частью преподаватели, и очень важно научить их отказываться от поведения преподавателя. Любой учитель, естественно, хочет помочь разрешить любую ситуацию, а экзаменатор должен сознательно воздерживаться от этого.

Самая распространенная ловушка, в которую попадают многие экзаменаторы, – желание нарушить молчание. «Молчание во время ожидания ответа – лучший друг экзаменатора» (*R. Chatham, из личной переписки*). Молчание в паузах hesitation не должно прерываться повторением или разъяснением вопроса. Нельзя подсказывать или подводить к правильному ответу. Необходимо дать шанс тестируемому самому справиться с ситуацией.

При вербальном взаимодействии наблюдается такой феномен как «принцип синхронности»: при замедлении речи одного из собеседников речь другого тоже замедляется. В тесте, как только тестируемый сталкивается с ситуационным или лингвистическим затруднением, его речь замедляется, экзаменатор же не должен подстраивать свою речь ни в коем случае.

Так как тест должен оценивать только спонтанную, неподготовленную речь, от экзаменатора требуется умение быстро распознавать по незначительным вербальным и невербальным проявлениям, реакциям и деталям подготовленную речь. Как только тестируемый начинает избегать визуального контакта, отводит взгляд в сторону, это означает поиск вербальных паттернов, то есть попытки вспомнить заученные и шаблонные фразы. Рейтеры/экзаменаторы должны быть знакомы с характеристиками устной и письменной речи. Если в ответе тестируемого присутствуют сложные предложения с грамматическими формами, характерными для письменной речи, то это всегда подготовленная речь. Такая форма типична, например, для презентаций, но не встречается в реальной жизни в устной речи. Активное слушание помогает переключить тестируемого с подготовленной на спонтанную речь: можно повторить конец фразы, произнесенной экзаменуемым, и уточнить информацию или перефразировать.

Во время обучения экзаменаторов необходимо добиться безусловного выполнения этих требований. Тех, кто неспособен придерживаться этих требований, нельзя допускать к проведению теста. Самые распространённые нарушения в работе экзаменаторов – это подсказки, постоянные возгласы одобрения после каждого высказывания тестируемого, повторение вопросов без просьбы тестируемого об этом, удовлетворение заученными наизусть ответами и т. п.

К сожалению, встречаются просто психологически не готовые к такой деятельности люди. Можем привести примеры из нашего опыта. Девушка-экзаменатор позволяла себе флирт с тестируемыми противоположного пола. Другая дама просто не давала открыть рот тестируемому: образец речи содержал только 30 процентов речи тестируемого, остальные 70 процентов составляла речь экзаменатора. Еще раз напомним, что меньше трети обучаемых рейтеров/экзаменаторов могут быть допущены к работе. К сожалению, часто можно понять, что человек не справится с работой, только в процессе обучения.

4.2. Инструкции для рейтеров по оценке образца речи

Разработчики тестирования должны прописать в Паспорте тестирования процедуру оценки. Процедура эта должна быть абсолютно прозрачной, чтобы исключить коррупционную составляющую. Должно быть прописано, кто назначает экзаменаторов/рейтеров, как выбирается вариант теста, сколько рейтеров оценивает образец речи, сколько времени это занимает, что происходит в случае расхождения оценок и т. д.

В целях объективности получения оценки лучше, если образец речи оценивают не менее двух рейтеров. Чтобы тестирование сделать более практичным, можно проводить оценку

двумя рейтеерами: один рейтеер проводит тест, а потом оценивает образец речи, другой только оценивает. Желательно, чтобы во втором случае это была анонимная оценка (*blind assessing*). Необходимо, чтобы тест не проводили и не оценивали преподаватели, которые принимали участие в обучении тестируемого.

Разработчики тестирования сами решают, какая документация сопровождает оценивание, это могут быть оценочные листы (*scoring sheets*) с перечнем ошибок, обоснование оценки (*rating sheets*) с описанием каждого элемента Шкалы ИКАО и общей оценки уровня владения языком по Шкале ИКАО. Документация должна быть полной и не вызывать сомнений в объективности оценки. Много лет назад диспетчер одного из диспетчерских центров привез из Великобритании сертификат, в котором черным по белому было написано, что у него четвертый уровень владения языком по Шкале ИКАО, потому что не третий.

Критерии оценки образца речи по всем элементам Шкалы ИКАО

Критерии оценки возникают из анализа ситуации использования языка и, соответственно, описания конструкта. Конструкт является ключом к пониманию метода оценки и критериев правильности в тесте определения уровня владения общим и профильно-ориентированным языком. Описание конструкта служит теоретическим обоснованием того, какие языковые аспекты необходимы для выполнения коммуникативной задачи.

Анализ каждого элемента Шкалы ИКАО и соотнесение образца речи с соответствующим описанием каждого уровня (*holistic descriptor*) позволяет рейтееру обосновать свою оценку каждого элемента примерами из речи тестируемого. Это непростая задача, так как каждый из оцениваемых элементов должен быть детально изучен. Необходимо также определить практические соображения, которые рейтееры должны иметь в виду в процессе оценки.

Произношение

Произношение – один из самых сложных элементов для рейтеерской оценки. С одной стороны, произношение включает в себя интонацию, громкость речи, логическое ударение, ритм, манеру говорения и акцент. С другой стороны, оценка произношения субъективно зависит от ознакомленности рейтеера с акцентом говорящего. Можно привести пример, когда на одном из семинаров-практикумов ИКАО образец речи французского пилота все присутствовавшие оценили, как 3 уровень по произношению, а рейтеер, для которого французский был вторым родным языком, оценил его как 5 уровень.

Если тестируемый и экзаменатор – носители одного родного языка, то экзаменатор не должен ориентироваться только на то, как он понимает речь тестируемого. Обязательно при оценке произношения необходимо учитывать следующие факторы.

Способность произносить отличительные звуки в словах с разным значением (*meaning-distinguishing sounds*), такие как долгие/краткие гласные, звонкие/глухие согласные, межзубные звуки, которые используются для распознавания различных слов (*feet/fit, search/surge, foot/food* и т. п.), по этой причине ошибки в фонологической форме не могут считаться незначительными, так как они напрямую влияют на ясность и могут привести к ложной интерпретации (*J. Purpura, 2005*). Эти ошибки считаются грубыми (*global errors*). В России рейтеер, оценивая произношение соотечественников, может непреднамеренно пропустить фонетические ошибки тестируемого, такие как отсутствие разницы между долгими и краткими гласными, оглушение звонких согласных в конце слова, интонацию, характерную для русского языка, неправильное произнесение межзубных звуков. Причина

кроется в предположении рейтера, что если речь понятна ему, то она будет понятна и другим пользователям английского языка. Чтобы избежать подобных ошибок в оценке произношения, рейтер должен постоянно иметь в виду фактор «ознакомленности» с акцентом.

Любой тест должен иметь задания, которые позволяют оценить уровень произношения с большой долей точности. Задания должны позволять получить достаточный объем речевого материала для оценки всех компонентов произношения.

Понимание фразы, контекста и ситуации в целом зависит от способности произносить и понимать звуки, которые образуют фонологическую форму слова, а, следовательно, его значения.

Главным критерием в оценке произношения является не сходство с произношением носителя языка, а понятность речи для членов авиационного сообщества. Существуют причины социального и психологического порядка, по которым некоторые неносители языка не хотят, чтобы их путали с носителями языка. Например, если неноситель языка допустит какую-либо культурологическую ошибку или по незнанию нарушит нормы вежливости, ему это простят как неносителю языка. Если акцент не препятствует пониманию, нельзя снижать оценку даже за ярко выраженный акцент, потому что акцент является частью национальной идентичности каждого человека (*S. Luoma, 2004*).

Грамматические структуры

Грамматические структуры и предложные схемы, которые употребляет тестируемый, определяются коммуникативными задачами.

Описания 3 и 4 уровней в Шкале ИКАО соотносятся с простыми грамматическими структурами и предложными схемами, в то время как употребление сложных структур соотносится с описанием 5 и 6 уровней. Также в описании каждого уровня приведено допустимое количество незначительных и грубых ошибок на каждом уровне. Так как целью теста является определение коммуникативной компетентности тестируемого по шести элементам Шкалы ИКАО, включая грамматические структуры, рейтеру необходимо сконцентрироваться не на знании грамматических правил, а на способности использовать грамматические структуры в речи. Рейтер оценивает этот элемент в коммуникативном акте, т. е. учитывает взаимное влияние грамматической формы, буквального, интенционального и прагматического значений. Форма определяется как морфология и как синтаксические схемы, т. е., как слова образуются и как они соединяются между собой. Этот аспект соотносится преимущественно с лингвистической точностью (*J. Purpura, 2005*). Про значения мы говорили в п. 2.1.

При оценке грамматической формы рейтеру следует обратить внимание на то, как грамматическая ошибка влияет на значение, т. е. искажает ли ошибка речевую интенцию. Например, вопрос тестируемого: "What eruption means?" – неправильный по форме. Мы можем предположить, что тестируемый может знать, что подобные вопросы требуют вспомогательный глагол "to do", но он/она не в состоянии употребить его в неподготовленной речи. В любом случае речевая интенция "to specify the meaning of the word 'eruption'" очевидна, несмотря на ошибку в грамматической форме. Ошибки такого рода считаются «локальными», или незначительными.

Другой пример, когда ошибка в грамматической форме не позволяет определить речевую интенцию говорящего: "I have received failure on board is no problem to fly is possible to ground". Ошибки, которые искажают значение высказывания или речевую интенцию

говорящего, называются «глобальными», или грубыми. На 4 уровне такие ошибки случаются редко (не более 10%).

Приведем еще примеры грубых ошибок. В определенных случаях употребление притяжательного местоимения женского рода “her”, вместо “his”, например, в описании требований террористов или буйных пассажиров может привести к тому, что террорист или пассажир мужского рода может быть воспринят службой безопасности как женщина.

Иногда случается, что грамматическая форма абсолютно правильная, но она не передает речевую интенцию говорящего. Например, рассказывая о прошлом, тестируемый говорит: “Do you need another destination?” Правильная грамматическая форма не выполняет коммуникативную задачу в описании прошедших событий, когда диспетчер спросил пилота, хотели ли они следовать на запасной аэродром.

В другом случае тестируемый объяснил инцидент с отключением двигателя следующим образом: “Engine was stopped”. Выбор грамматической формы (страдательный залог глагола) говорит о том, что двигатель отключил экипаж. На самом деле, обсуждалась ситуация, когда в полете неожиданно заглох двигатель. Или другой пример: собирая информацию об уже произошедшем событии, тестируемый опять выбирает неправильную грамматическую форму: “And tell me when you have a TA on your TCAS, do you change your vertical speed?”

Эти примеры должны привлечь внимание рецензентов в тех случаях, когда выбранная форма не соответствует коммуникативной задаче.

Несмотря на то, что форма и значение тесно связаны друг с другом, рецензент должен различать эти компоненты, так как это влияет на оценку элемента Структуры.

Словарный запас

Оценивая словарный запас, рецензент не оценивает количество слов, которые знает тестируемый. В тесте словарный запас оценивается с точки зрения его достаточности для обсуждения общих, конкретных и связанных с работой тем, насколько уместен лексический выбор для данного контекста и ситуации использования языка.

Лексика, используемая в авиационном английском языке, может быть условно поделена на три категории:

1. Слова из общего языка, которые в авиационном английском языке имеют такое же значение, как и в общем языке: weather, wind, aerodrome, route и т. п. Эта лексика имеет отношение к доменам общего языка (common and concrete topics).
2. Слова, которые в общем языке имеют одно значение, а в авиационном английском языке – другое: “runway” как “long stage that models walk on during a fashion show” (подиум) и “runway” as “a hard level roadway where aircraft can take off and land” (взлетно-посадочная полоса; “taxi” как “car that may be hired, along with its driver, to carry passengers to any specified destination” (автомобиль такси) и “taxi” как “to move along the ground, especially before takeoff and after landing” (рулить); “climb” как “to go up mountains or rocks” (подниматься в горы) и “climb” как “to gain height» (набирать высоту), и т.п.). Это – псевдо-техническая лексика (*semi-technical or sub-technical vocabulary*).
3. Техническая лексика, или терминология (technical vocabulary): instrument approach, cruising flight level, vertical situation indicator, primary radar, side-step maneuver, и т. п.

В тесте должны быть задания, позволяющие тестируемому продемонстрировать владение лексикой всех категорий. Рейтер должен уметь оценить случаи, когда выбор лексической единицы не соотносится с коммуникативной задачей и может привести к недопониманию, например, употребление слова “dispatcher” вместо “controller”.

Очень важно оценить способность тестируемого перефразировать при незнании каких-то слов. Зная, что употребление идиом относится рейтерами к более высокому уровню, тестируемые иногда заучивают какое-то количество идиом и вводных слов наизусть, но они не всегда могут их употребить правильно в речи, часто вставляя их не к месту. Рейтер должен быть готов к таким ситуациям и оценивать не только знание, но и умение применить эти знания.

Переключение кода, когда тестируемый переходит на родной язык, рассматривается как грубая ошибка.

Беглость речи

Sari Luoma (2004) говорит, что «если рейтерам не дать словесного описания беглости речи и попросить их выбрать наиболее «беглый» из нескольких образцов речи, то они будут склоняться к несогласию друг с другом как в определении, так и в методах ее оценки. Определения беглости речи часто включают в себя ссылки на плавность, скорость речи, отсутствие чрезмерных пауз, отсутствие раздражающих маркеров хезитации, длительность высказываний и их связность. Эти характеристики сложные, потому что это не просто описание речи говорящего, но и описание восприятия ее слушателем».

Например, «наличие избыточных пауз» — это описание речи говорящего, но базируется оно на восприятии его речи слушателем.

Основная характеристика беглости речи относится к временным аспектам речи, таким как соотношение речи и пауз, и частоте маркеров потери беглости речи, таких как хезитации, повторения и самокоррекции. Существует четыре типа хезитации: (1) незаполненные паузы (молчание); (2) паузы, заполненные звуками “uh”, “um”, “er”, “ah”, “mm” или словами “well”, “anyway”, “let me think”, “it’s an interesting question”; (3) повторения, когда тестируемый повторяет фразу или часть фразы несколько раз; (4) фальстарты, когда тестируемый начинает говорить, внезапно замолкает, а потом заменяет предыдущее слово или фразу другим словом или фразой.

Чем выше уровень владения языком, тем выше скорость речи тестируемого, тем короче паузы хезитации. Такие паузы всегда семантически обоснованы. Слушатель может воспринимать подобные паузы как размышления над содержанием высказывания, а не как поиск подходящего слова или грамматической структуры. Чем выше уровень владения языком, тем точнее и полнее высказывания тестируемого.

Употребление вводных слов и связок тоже характеризуют беглость речи. Чем богаче и разнообразнее речь тестируемого в выражении речевой интенции, в реакции на высказывания экзаменатора, в степени обязательности или неуверенности, в запросе подтверждения и т. п., то есть в случаях, где вводные слова и связки наиболее уместны, тем выше уровень беглости речи.

Беглость речи напрямую зависит от произношения, грамматики и словарного запаса. Чем хуже произношение, тем труднее артикулировать малознакомые слова. Чем больше пробелов в грамматике, чем беднее словарный запас, тем больше времени требуется для поиска подходящих языковых средств.

При этом рейте́р не должен оценивать беглость только по скорости речи. Когда речь тестируемого относительно медленная, но паузы в ней контекстуально обоснованы, высказывания логически завершены, маркеры хезитации не раздражают, то можно предположить, что тестируемый так говорит и на родном языке. В подобных случаях нельзя занижать оценку беглости речи.

Понимание

Понимание на слух – навык, необходимый для успешной коммуникации. Оценивая понимание, необходимо учитывать особенности ситуации использования языка в радиообмене: понимание всегда интерактивно, то есть понимание выражается вербально или действием. Анализ элементов, составляющих понимание, показывает, что понимание на слух невозможно без достаточной компетенции в фонологии, морфологии, синтаксисе, словаре, организации речи и т. д.

Если беглость речи напрямую зависит от произношения, грамматики и словарного запаса, то уровень понимания не всегда совпадает с оценками этих элементов. На понимание влияют так называемые компенсаторные факторы. В реальной радиосвязи такой важный компенсаторный фактор как визуальный контакт отсутствует. С другой стороны, профессиональный опыт тестируемого начинает в понимании на слух играть важную роль. Предметные знания и предыдущий опыт позволяют тестируемому воспользоваться преимуществом языковой догадки и «включить» механизм прогнозирования в случаях, когда ему не хватает языковых ресурсов.

При оценке понимания рейте́р должен учитывать факторы, которые препятствуют пониманию. Темп речи экзаменатора и в записях для аудирования не должен превышать 170 слов в минуту (средний темп диалогической речи) и не должен быть менее 100 слов в минуту (рекомендация ИКАО для радиообмена). Используемый для аудирования акцент должен соответствовать описанию произношения на 4 уровне по Шкале ИКАО, т. е. должен быть понятен членам международного авиационного сообщества.

Понимание должно оцениваться на протяжении всего теста. Необходимо проследить, все ли вопросы экзаменатора были поняты с первого предъявления, а если нет, то указать причину затруднения понимания. Если понимание в аудиозаданиях проверяется при помощи вопросов контроля понимания, рейте́р должен привести доказательства соответствия оценки понимания описательной характеристике уровня в Шкале ИКАО.

Взаимодействие

Оценивая шестой элемент Шкалы ИКАО, рейте́р должен определить, насколько успешно тестируемый применяет навыки говорения и понимания на слух в коммуникативном акте и для решения коммуникативных задач, с которыми он сталкивается. Ответы, вопросы и комментарии тестируемого определяют эффективность взаимодействия и могут быть оценены по таким критериям как незамедлительность, информативность и, самое важное, адекватность. Успешное взаимодействие всегда имеет цель, а результат его непредсказуем и достигается в ходе коммуникации, как в реальной жизни.

Успешность коммуникативного акта зависит от того, владеет ли тестируемый языковыми средствами для того, чтобы проверить, уточнить, подтвердить. Рейте́р должен понимать, что взаимодействие зависит от других пяти элементов Шкалы ИКАО, и оценка взаимодействия не может быть выше оценок других элементов. Тестируемый, с трудом артикулирующий английские слова, с недостаточным набором грамматических структур и

предложных схем, с недостаточным словарным запасом, не способный выдать связное высказывание, не может взаимодействовать эффективно.

Оценка уровня владения языком по Шкале ИКАО

Рейтеры (не менее двух) прослушивают анонимную запись теста. Рейтер должен отметить в оценочном листе все ошибки с указанием, грубая это ошибка или незначительная, обосновать оценку каждого аудиозадания. После этого рейтер сравнивает оценочный лист с описанием каждого элемента Шкалы ИКАО, обосновывает оценку каждого элемента и, определяет общий уровень владения языком в соответствии с холистическими дескрипторами. Общий уровень владения языком определяется по низшему уровню одного из шести элементов.

Образец Rater's Rating Sheet

Rater's Rating Sheet

| | | | |
|------------------|----------------|-------------------|----------------|
| Date of Testing: | Deleted | Candidate's Code: | Deleted |
| Date of Rating: | Deleted | Rater's Code: | Deleted |

| Time | Observation | Element |
|---------------|--|---|
| Task 1 | | |
| 0137-0356 | I graduated technical university 8 years ago...to be a cadet [wrong stress]. Now I'm studying at the third course. If I finish flying school this month, I think I will try to start my work as a pilot this summer...It takes me about one day. | 2+, 1+, 1- 2+, 1+, 2+ 2+, 1+, 2+, 4+ |
| 0720-0802 | When I thought about entered university...there [z] were very bad situation in aviation uh and my father said that [z] it would be very hard to <i>pause</i> to join any air company after you uh finish uh flying school... | 2-, 1-, 2- 2+, 1-, 2+ 4-, 2+, 2- |
| 0821-0851 | So, I'm going to return my flying school. I will waiting for results and um I <i>pause</i> I have passed all my exams uh and uh I have to fly about 10 hours to graduate flying school. | 2+, 2-, 2- 4-, 2+ 2+, 2- |
| Task 2 | | |
| 0904-0936 | I know visual and instrument flights. Um so according to the purpose there are commercial [k] flights, some special flights connecting with [s] special purposes, maybe emergency flights and other. | 4+, 1+, 3+ 3+, 1-, 3-, 1-, 3- |
| 1002 | So um, can you rephrase your question, please? | 6+ |
| 1019-1110 | So, so there are different problems and one of them for example is the bad weather condition. So, for example, unruly passengers...misunderstanding between pilots and between Air Traffic Controllers...some technique problems, some engine failure...instrument failure, fire onboard. | 1+, 2+, 4+ 5-, 5+, 3+ 1+, 3+ 2-, 3+ |
| 1122-1212 | Um I know some of them, PAN PAN PAN...I don't remember but I think it's for urgency situation...SOS [] maybe I know that in case of emergency situation pilots have to set 7400 code on transponder and this is one of the signs of the emergency situations.... | 3+, 5-, 2+ 5+, 1- GL 2+, 1+, 3+, 4+ |
| 1213 | Oh yeah MAY DAY signal we have to use in the case of emergency situation. | 3-, 3+ |
| 1226-1333 | So maybe um <i>pause</i> difference is maybe in distress situation it's possible to proceed flight...pilots have to make an emergency landing. Um okay so in the case of uhh engine failure of uh one engine uh aircraft it's impossible to proceed flight and uh the first procedure is to initiate an emergency landing. | 4-, 5+, 6+ 2-, 4- 5+, 6+ 4+, 1+, 2+ |
| 1341-1413 | So as I um as I know maybe in the case of the same engine failure but for multi-engine aircraft um pilots can proceed their flight maybe to their airport destination and maybe to the nearest airport... | 5+, 6+ 2- 2+, 1+, 4+ |
| 1446-1550 | Due to uh flammable [] materials...and during uh fire situation it's very hard to breathe [s] sometimes and uh pilots have to use their oxygen uh uh facilities, oxygen system...the first procedures is to extinguish fire and if it's not possible to proceed they have to make make emergency landing. | 1- 1-, 5+ 2- 1+, 4+ |
| Task 3 | | |
| 1559-1653 | A runway incursion is a situation when <i>pause</i> uh pilots uh initiate uh rolling to the runway without any permission from air traffic controllers. So maybe it's connected with [z] air traffic controllers too but...as I know yes. | 6+, 1+, 4- 5+, 6+, 4+ 5- |
| 1730-1842 | So pilots uh have to initiate missed approach in the case of any runway uhh in the case when someone uh is on the runway...in what situation? By uh air traffic controller uh so when air traffic controller said to do it, so when uh the attitude maybe speed and altitude and speed of the aircraft are not safety [] for landing. | 1+, 6+ 5-, 6- 6+ 1- GL 1+, 2-, 6+ |
| 1848-1923 | Um these height is called decision altitude...no, it depends on different things, uh for example visibility and uh captain's minimums for landing and uh that's all I think. | 1-, 1+, 5+ 2+, 5+ |
| 1940-2027 | Um because the take-off and landing is one of the important parts of the flight and as I know the ATC could not able to talk with [z] pilots uh before they maintain...Sorry, can you rephrase your question, please? | 5-, 4+ 1- 6+ |

| | | |
|---------------|--|--|
| 2042-2104 | So I think in the case of some emergency and situation and maybe uh when it's not safety to proceed next landing or take-off maybe. | 5+, 6+ 2- |
| Task 4 | | |
| 2350-2432 | Um some of the passengers suffering <i>pause</i> and due to as I understood due to some something uh with [s] his sugar level. So I have understood that it was a woman not some...just one. | 5-, 4- 1-, 5+ 5+, 6+ |
| 2439-2506 | Um so um yes uh the [z] um crew request but when uh they landed only when they landed the doctor treated this woman. | 1-, 4- 5+, 6+ |
| 2515 | So I don't know about it. | 5-, 6- |
| 2523-2542 | She suffer from some kind of uh uh with something connected with sugar maybe sugar level of her blood in her blood. | 2- 5+ |
| 2558-2622 | It's very hard for me to say uh what they do but they try to <i>pause</i> they try to help her and they make they made her sugar level normal. | 4- 5+, 6+ |
| 2627 | She continue she continued her journey. | 2-, 2+ |
| Task 5 | | |
| 2906-2927 | So uh the captain of Delta 31 um request emergency service and police uh due to uh one of the passengers uh reclaim that he has a bomb. | 4+, 5+, 6+ |
| 2937-3010 | Uh so I think...because the stand of this airplane will be another I mean that...the airplane must have to be stand...in special stand. | 4- 2-, 3-, 4- |
| 3018-3028 | Yes, it's a positive because they uh I understood that it's not a bomb. | 5+, 6+ |
| 3057-3140 | The air traffic controller said that it's uh will be able to uh uh to so let me rephrase. So I'm I've heard that the airplane have landed and the situation was soft. | 4- 2-, 5- 3- |
| Task 6 | | |
| 3303-3336 | Can you describe the tale of this event? Um so what type of aircraft <i>pause</i> it is is it? | 2+, 3+, 6+ 2-, 2+ |
| 3350-3414 | Who was captain and co-pilot? So what was weather condition? The situation happened during landing...? | 2-, 6+ |
| 3430 | It was a problem with engines or maybe with controls? | 5+, 6+ |
| 3447 | So, how old is the captain, do you know? | 2+, 6+ |
| 3512-3522 | Did they pass over tower to confirm that something wrong with [z] wheels? | 1-, 5+, 6+ |
| 3600 | Did they make a landing? | 2+, 6+ |
| 3611 | What about passengers? | 6+ |
| 3628 | Was there a fire after this emergency? | 2+, 6+ |
| 3808-3925 | So, um it was um uh emergency landing...first destination was uh...some other town Lithuania but they missed approached and uh flew back due to one the main gear is not rejected and...one of the gears uh <i>pause</i> wasn't uh eject yeah, no. One of the gears is not get down. | 4-, 5+ 4-, 5+, 6- 3+, 6+ |
| 3926-4105 | They the pilots didn't initiate landing...due to short uh of the runway and after that they flew back and made a emergency landing uh without one of the main gear. The passengers were ok and um no one were injured but during landing rolling uh the aircraft was uh rolling out from the runway as I have understood and so now the investigation group is working and um there are no any results about this situation now...So, yes as I have understood it was damaged...I think the wing is exactly damaged. | 4-, 6- 5+, 6+ 5+, 6+ 2-5+, 6+ 2-, 5+, 6+ 5+, 6+ |

| Pronunciation | Structure | Vocabulary | Fluency | Comprehension | Interactions |
|---------------|-----------|------------|---------|---------------|--------------|
| 4 | 4 | 4 | 4 | 4 | 4 |

ICAO Score: 4

Notes and Observations: The test-taker's existing command of English and his performance in the test demonstrate a level 4 on the ICAO scale. His pronunciation is clear and, though it is affected by his native

language, it does not interfere with ease of understanding. He makes grammar errors, especially later in the test when he encounters unexpected situations. However, these mistakes are not numerous.

Pronunciation: The test-taker's pronunciation, stress, rhythm, and intonation are influenced by his first language (i.e. 0720-0802, 0904-0936, 1341-1413, 1446-1550, etc.) but only sometimes interfered with ease of understanding (1122-1212). Most of the time, his pronunciation was clear, correct, and completely understandable.

Structure: The test-taker's basic grammatical structures and sentence patterns were used creatively and were usually well-controlled (i.e. 0137-0356, 1019-1110, 1122-1212, 1848-1923, etc.). Some errors occurred, especially as he encountered unpredictable and unexpected situations (i.e.). He frequently used the form "flying school" instead of "flight school" (i.e. 0821-0851) and "safety" instead of "safe" (i.e. 1559-1653, 2042-2104). However, these mistakes did not often interfere with meaning. He also corrected himself on occasion (i.e. 2627, 3303-3336).

Vocabulary: The test-taker's vocabulary range and accuracy are usually sufficient to communicate effectively on common, concrete, and work-related topics (i.e. 0904-0936, 1019-1110, 1122-1212, 1226-1333, 1848-1923, etc.). He did need to be reminded of the MAY DAY call, but he knew what it was when the examiner mentioned it (1213). He showed that he was able to paraphrase when he lacked vocabulary (3808-3925).

Fluency: He produced stretches of language at an appropriate tempo (i.e. 0137-0356, 0904-0936, 1019-1110, 1122-1212, 1446-1550, 1848-1923, 2906-2927, etc.). His fluency did suffer and he used more fillers when he had to produce spontaneous interaction (i.e. 2350-2432, 2558-2622, 2937-3010, 3057-3140). However, he was mostly able to communicate effectively.

Comprehension: His comprehension was mostly accurate on common, concrete, and work-related topics (i.e. 1226-1333, 1341-1413, etc.). He was able to understand the main ideas and some details from the audio tasks (i.e. 2350-2432, 2439-2506, 2627, 2906-2927, 3018-3028). However, he did miss some important details and would need to use clarification strategies in the future. When he did not understand the examiner, he was able to ask her to clarify her speech (i.e. 1002, 1940-2027, etc.). During the simulation task, he did relay some correct and useful information. However, he was not able to communicate the details very well or immediately when he retold the story (3808-3925, 3926-4105).

Interactions: For most of the test, the test-taker's responses were appropriate, informative, and timely (i.e. 0137-0356, 1226-1333, 1341-1413, 2906-2927, 3018-3028, etc.). He was able to maintain and initiate exchanges rather well during the simulation task by correctly asking many relevant questions (i.e. 3350-3414, 3430, 3447, 3512-3522, 3600, 3611, etc.).

Такой формат оценки образца речи содержит доказательства в виде примеров, как положительных, так и отрицательных. На основании фактической информации рейтеер делает вывод, каким холистическим дескрипторам, т. е. какому уровню владения языком по Шкале ИКАО соответствует образец речи тестируемого. Такой метод не вызывает сомнений в результатах тестирования ни у тестируемых, ни у работодателя, ни у контролирующих органов.

Если у рейтеера возникнут какие-либо сомнения или замечания по проведению теста, поведению экзаменатора или качеству записи, он может их изложить письменно в листе обоснования оценки.

В документации теста должен быть подробно описан порядок действий в случае несовпадения оценок двумя рейтеерами: что делается в этих случаях, кто определяет конечную оценку?

5. Мероприятия по обеспечению надежности теста

Мы уже писали, что валидация теста – это не однократная акция. Валидация подтверждается на протяжении всего времени использования теста. Надежность теста тоже должна поддерживаться все время применения теста.

В этой главе мы остановимся на такой характеристике надежности теста, как качество работы рейтееров. Рейтер считается ненадежным, если он/она не может придерживаться стандартов оценки, критерии оценки применяет непоследовательно, его оценки зачастую не совпадают с оценками других рейтееров. Контроль качества оценки образцов речи должен осуществляться на постоянной основе. Даже если руководитель проекта или старший рейтеер добились во время первоначального курса обучения адекватной и единообразной оценки от всех рейтееров, вовсе не факт, что это качество сохранится на долгое время. На качество оценки влияют многие факторы, например, какое количество оценок проводится рейтеером в определенное время. Известны случаи, когда рейтеер проводит оценку только несколько раз в год. Конечно, гарантировать качество при такой нагрузке просто невозможно.

Каким образом осуществляется контроль качества оценки образцов речи, т. е. надежности теста? Обсуждение контроля качества работы рейтееров не может обойтись без двух терминов: *intra-rater reliability (IntraRR)* и *inter-rater reliability (InterRR)*.

В первом случае речь идет о степени постоянства в оценках одного и того же рейтеера, т. е. как рейтеер оценивает один и тот же образец речи в разные промежутки времени. В этом случае минимальная вариативность допускается при оценке языковых элементов, но не допускается разница в оценке общего уровня владения языком.

Во втором случае имеется в виду степень консенсуса между рейтеерами, т. е. как два или более рейтееров оценивают один образец речи. Здесь также допустима минимальная вариативность в оценке элементов, и не допускается разница в оценке общего уровня.

Еще при первоначальном обучении руководитель проекта или команда старших рейтееров отбирает образцы речи для оценки. Образцы речи сначала оценивает руководитель проекта или старший рейтеер, знакомит будущих рейтееров с принципами и критериями оценки. Потом образцы речи с разным уровнем обсуждаются в группе, и только после этого рейтееры оценивают образцы речи самостоятельно. Каждый рейтеер должен обосновать свою оценку, если она не совпадает с оценками других рейтееров и с оценкой старшего рейтеера. В процессе обучения необходимо добиться, чтобы оценки уровня владения языком не были ни занижены, ни завышены. На первых этапах старший рейтеер должен контролировать работу рейтееров, а в дальнейшем необходимо определить форму и периодичность контроля и КПК всех экзаменаторов/рейтееров.

Когда образец речи оценивается независимо двумя рейтеерами, то степень согласия или несогласия в оценке всегда очевидна. Если два рейтеера по-разному оценили один и тот же образец речи, то это должно быть сигналом для руководства проекта о необходимости проведения мероприятий по стандартизации процесса оценки. Должна быть предусмотрена процедура принятия окончательного решения в отношении оценки: кто привлекается в качестве «третьего судьи», чей голос является решающим? Какие механизмы контроля качества работы рейтееров существуют? Попробуем описать некоторые из них. Желательно, чтобы оценка образца речи проводилась рейтеерами независимо друг от друга, т. е. должен быть исключен контакт между ними. Это – забота администрации тестирования, при современных средствах связи и технической оснащенности это не трудно делать в разных локациях с определенной степенью анонимности.

Администрация тестирования или руководитель проекта должны определить временные промежутки сессий стандартизации, на которых проверяется степень согласованности оценок всех рейтееров. Проведение таких сессий офлайн возможно, но трудно осуществимо, если коллектив рейтееров многочисленный, и рейтееры находятся в разных городах (иногда странах). Поэтому проведение подобных мероприятий онлайн удобнее. Опишем возможный механизм стандартизации работы рейтееров. Раз в полгода или при возникновении случаев несогласованности оценок руководитель проекта выбирает образец речи и через администратора под разными кодами отправляет всем рейтеерам, кроме того, кто проводил тест. Все оценки администратор предоставляет руководителю проекта, который проводит анализ и принимает решение в отношении рейтееров, которые занижали или завышали оценки. Ненадежные рейтееры дополнительно изучают образцы речи на диске ИКАО и отобранные руководством образцы речи в качестве примеров, критерии оценки каждого элемента и общего уровня. Если после внеочередного курса повышения квалификации рейтеер продолжает демонстрировать большую вариативность в оценках, то он отстраняется от работы.

Подобным образом можно проверить и степень постоянства оценки одного рейтеера (*intra-rater reliability*). Рейтееру через значительные промежутки времени отправляется образец речи, ранее им оцененный, но под другим кодом (это возможно при достаточной рейтеерской нагрузке). Результаты оценки сравниваются. При неудовлетворительном результате можно предусмотреть КПК, описанный выше.

В качестве примера приведем результаты одной сессии стандартизации оценок:

| Rater | Assessment | Comment |
|-------|------------|---|
| A | 433443 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. A four in comprehension is incorrect as the test-taker failed to get important details in Audio task 1 and 2. |
| B | 333333 | Precise rating |
| D | 333333 | Precise rating |
| F | 333333 | Precise rating |
| G | 333333 | Precise rating |
| H | 433433 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. |
| I | 434343 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in lexis is not proved by the rater's scoring sheet. A four in comprehension is incorrect as the test-taker failed to get important details in Audio task 1 and 2. |
| J | 333433 | A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. |
| K | 433443 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in comprehension is incorrect as the test-taker failed to get important details in Audio task 1 and 2. |
| L | 433433 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. |
| N | 333433 | A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. |
| Q | 322332 | Two in lexis and interaction is a real underestimation of the candidate. The rater requires additional training. |
| P | 333433 | A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. |
| R | 433433 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. |
| S | 333333 | Precise rating |
| T | 433333 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. |
| U | 433433 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. |
| X | 434544 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in lexis is not proved by the rater's scoring sheet. A five in fluency is a sign of completely inadequate judgment. A four in fluency means the rater failed to identify parts in the test where speech tempo was well below 100 words per minute, with false starts, pauses and sounds of hesitation. A four in comprehension is incorrect as the test-taker failed to get important details in Audio task 1 and 2. A four in interactions is inadequate as the test-taker made two global mistakes with multiple local ones, and questions asked by the test-taker in Part 6 were prepared and often did not fit the situation. |

| | | |
|----|--------|---|
| AS | 333333 | Precise rating |
| AR | 434334 | A four in pronunciation shows that the accent of the test-taker was familiar to the rater. A four in interactions is inadequate as the test-taker made two global mistakes with multiple local ones, and questions asked by the test-taker in Part 6 were prepared and often did not fit the situation. |

Каждый рейтеер получает комментарий руководителя проекта или старшего рейтеера и рекомендации. А рейтеерам Q и X необходимо пройти внеочередной КПК.

В тестологии существует несколько различных подходов к оценке стабильности и надежности работы рейтееров, например, Cohen's kappa, Rasch model и другие. Коэффициент Cohen's kappa — это статистические данные, которые используются для измерения стабильности и надежности (или ненадежности) между разными рейтеерами, оценивающими качественные категории. Cohen's kappa – довольно точный метод, однако он сложен в использовании, поскольку в статистику входят и те случаи, когда одинаковые оценки рейтееров вызваны не столько надежностью их работы, сколько простым совпадением.

Cohen's kappa использует следующую градацию надежности работы рейтееров:

| Value of Kappa | Level of Agreement | % of Data that are Reliable |
|----------------|--------------------|-----------------------------|
| .40–.59 | Weak | 15–35% |
| .60–.79 | Moderate | 35–63% |
| .80–.90 | Strong | 64–81% |
| Above .90 | Almost Perfect | 82–100% |

Следует отметить, что процентные соотношения при использовании других методов определения надежности работы рейтееров отличаются незначительно. Вероятно, разработчикам тестирования и руководителям проекта необходимо с учетом этих исследований, но, тем не менее, самостоятельно установить пороговую величину, ниже которой работа рейтееров не может считаться надежной.

Математическая теория, лежащая в основе моделей Раша, является, в общем смысле, частным случаем обобщенной линейной модели. Однако существуют важные различия в интерпретации параметров модели и ее философском значении, которые отделяют сторонников модели Раша от традиции моделирования ответов на вопросы. Работать с такими сложными моделями могут профессиональные специалисты, однако их привлечение к работе в рамках тестирующей системы незамедлительно сказывается на стоимости прохождения теста, делая его неконкурентоспособным.

Поэтому из различных способов оценки InterRR наиболее практичной оказывается линейная модель сравнения оценок рейтееров в отношении одних и тех же кандидатов. Такая модель не требует значительных затрат и позволяет достаточно оперативно реагировать на возможные негативные тенденции в работе рейтееров. Пороговой величиной для определения надежности работы рейтееров считается значение 86% соответствий. Что это означает на практике?

Коэффициент InterRR (*inter-rater reliability*) можно определить путем соотношения совпадения оценок к общему количеству проведенных тестов. Например:

| Year | Number of Tests | Percent Agreement | InterRR |
|------|-----------------|-------------------|---------|
| 20XX | 574 | 78% | .90 |
| 20XX | 870 | 81% | .87 |
| 20XX | 640 | 79% | .88 |
| 20XX | 300 | 88% | .94 |

Как можно трактовать такие результаты? На протяжении 4 лет эксплуатации тест показывает высокую степень надежности. При вариативности коэффициента стабильности работы рейтегов (*InterRR*) в разные годы надежность теста не опускалась ниже допустимого порога.

6. Заключение

В этой работе мы попытались объяснить, почему для тестирования по Шкале ИКАО нельзя воспользоваться широко известными тестами для определения уровня владения общим английским языком. Пользуясь определением профессионального дискурсивного сообщества, мы описали ситуацию использования языка и привели отличительные черты радиообмена. На основе анализа потребностей было определено содержание теста и конструктор теста. Подробно описан и процесс валидации.

В данной работе подробно описаны критерии отбора персонала для разработки и применения тестирования, потому что от этого зависят и качество теста, и репутация тестирующей организации. Подготовка экзаменаторов и рейтегов, контроль качества получения и адекватной оценки образцов речи тестируемых находится в зоне ответственности руководителей проекта. Критерии оценки каждого элемента Шкалы ИКАО и общего уровня должны быть хорошо знакомы каждому сертифицированному рейтегу.

Вы могли ознакомиться с типами тестовых заданий, примерами спецификаций тестовых заданий. Наличие спецификаций тестовых заданий является основой для создания банка валидных параллельных вариантов теста.

Приступая к разработке тестирования, необходимо обладать достаточными теоретическими знаниями в области тестирования, поэтому мы кратко изложили особенности тестирования в области специальных языков, но очень надеемся, что список литературы пригодится и тем, кто только собирается приступить к разработке тестирования по Шкале ИКАО, и тем, кто уже применяет тестирующие системы на практике.

Анализ применяемых для определения уровня владения общим и авиационным языком пилотами и диспетчерами тестов показывает, что валидность и надежность некоторых вызывает сомнения. Надеемся, что эта работа поможет оценить применяемые тесты с точки зрения их соответствия требованиям, предъявляемым к тестированию в области профильно-ориентированных языков.

Все цитаты приведены в переводе авторов.

Библиография

1. ICAO Annex 1 – Personnel Licensing.
2. ICAO Annex 10 – Aeronautical Telecommunications. Volume II – Communication Procedures including those with PANS status.
3. ICAO Doc 4444 PANS ATM – Air Traffic Management.
4. ICAO Doc 9432 – Manual of Radiotelephony.
5. ICAO Doc 9713 – International Civil Aviation Vocabulary
6. ICAO Doc 9835 – Manual on the Implementation of ICAO Language Proficiency Requirements.
7. ICAO Circ 318 – Language Testing Criteria for Global Harmonization.
8. ICAO Working Document – Proficiency Requirements in Common English Study Group (PRICESG/1). Summary of Discussions and Conclusions.
9. ICAO Policy on Language Proficiency Testing, Montreal 03/2006.
10. Air Traffic Control Test Preparation. Learning Express, 2009.
11. Eurocontrol. Air-Ground Communication Safety Study, 2004.
12. ICAO. Proficiency Requirements in Common English Study Group, First Meeting – Summary of Discussions and Conclusions, 2000.
13. IATA, IFALPA, IFATCA, Phraseology, 2011
14. Alderson, J. Charles et al. Language Test Construction and Evaluation. Cambridge University Press, 2006.
15. Bachman, L. F. Language Testing in Practice, 1990
16. Bachman, L. F. Statistical Analyses for Language Assessment. Cambridge University Press, 2004.
17. Bhatia, V. Applied genre analysis and ESP. In T. Miller (ed.) Functional approach to written text: Classroom applications. Washington D.C.: United States Information Agency. 1997.
18. Bhatia, V.K. Analyzing Genre: Language Use in Professional Settings. London: Longman. 1993.
19. Buck, Gary. Assessing Listening. Cambridge University Press, 2001.
20. Chatam, R., G. Henning, H. Mogilka, R. Ramos, S. Tomas. Language Tasks in Air Traffic Control English Language Project (ATCELP): Identifying Basic English Language Proficiency for International Air Traffic Controllers, 1999.
21. Chatam, R. личная переписка.
22. Crystal, D. English as a Global Language. Cambridge: Cambridge University Press, 1997.
23. Cushing, S. “Fatal words”. Communication Clashes and Aircraft Crashes. The University of Chicago Press, 1994.
24. Cushing, S. Pilot – Air Traffic Control Communications: It’s Not (Only) What You Say, It’s How You Say It. Flight Safety Digest. Flight Safety Foundation, 1995.
25. Defense Language Institute’s English Language Center, Human Resources Research Organization – English Language Proficiency of International Air Traffic Controllers: A Review and Guidelines for Appropriate Measures, 2000.
26. Defense Language Institute’s English Language Center, Human Resources Research Organization – Air Traffic Control English Language Project (ATCLEP) I: Identifying Basic English Language Proficiency for International Air Traffic Controllers, 1999.
27. Defense Language Institute’s English Language Center, Human Resources Research Organization – Air Traffic Control English Language Project (ATCLEP) II: An Analysis and Validation of Language Proficiency Measurement Models, 1999.
28. Douglas, D. Assessing Languages for Specific purposes. Cambridge University Press, 2002.
29. Dudley-Evans, A., St John, M.J. Developments in English for Specific Purposes. Cambridge University Press, 2006.
30. ENAC, dept LH – Levels of English (comprehension and spoken communication) for Trainee and Qualified Air Traffic Controllers, 2001.
31. Hayes, Michael D. Instrument Oral Exam Guide. Aviation Supplies and Academics, Inc, 1992.

32. Hayes, Michael D. Private Oral Exam Guide. Aviation Supplies and Academics, Inc, 1992.
33. Hutchinson, T. and Waters, A. Performance and competence in ESP. Applied Linguistics, 1981.
34. Hutchinson, T. and Waters, A. English for Specific Purposes. Cambridge University Press, 2006.
35. Hymes, D. On Communicative Competence. Sociolinguistics. Harmondsworth, Middlesex: Penguin, 1972.
36. Illman, P.E., Pouzar, J. The Pilot's Radio Communications Handbook. TAB Books, 1988.
37. Luoma, Sari. Assessing Speaking. Cambridge University Press, 2004.
38. Mell, J. What is Not Standard in Real Radiotelephony? Presented at the 4th International Civil Aviation English Association Forum in Paris, France 1991 and reprinted in Doc 9835.
39. Oksman, Pasi. A Model for the Analysis of Aviation Radiotelephony. University of Jyväskylä, 1989.
40. Purpura, J. E. Assessing Grammar. Cambridge University Press, 2005.
41. Raimund Fridrich, Communication is the key to a safer world, 2004.
42. Read, J. Assessing Vocabulary. Cambridge University Press, 2005.
43. Swales, J. M. Genre Analysis in academic and research settings. Cambridge: Cambridge University Press, 1990.
44. Swales, J. M. Other Floors, Other Voices: A textography of a small university building. Mahwah, NJ: Lawrence Erlbaum, 1998.
45. Swales, J. M. The Concept of Discourse Community. Composition Forum 37, Fall 2017.
46. Teasdale, A. Content validity in tests for well-defined LSP domains: an approach in defining what is to be tested. Cambridge University Press, 1996.
47. Vatnsdal, A. O. A register analysis: The language of air traffic control. Occasional Papers in Systematic Linguistics, Vol.1, 1987.
48. Арнольд И. В. Стилистика. Современный английский язык. Флинта, 2006.
49. Пассов Е. И. Коммуникативный метод обучения иноязычному говорению. Москва Просвещение, 1985.
50. Поляков О. Г. Английский язык для специальных целей: теория и практика. НВИ-Тезаурус 2003.
51. Приказ Росавиации "Об утверждении Методических рекомендаций о тестировании членов летных экипажей воздушных судов гражданской авиации Российской Федерации на определение уровня владения авиационным английским языком по Шкале ИКАО" №353-П от 26 апреля 2018 г.