

Часть речи английского термина	Длина в словоформах	Часть речи русского соответствия	Длина в словоформах	Количество терминов такого типа	Пример английского термина	Пример русского термина
N	3	N	3	578	theme rheme relation	тема-рематическое отношение
N	1	N	2	535	abessive	абессивный падеж
V	1	V	1	499	adjectivize	адъективировать
N	3	N	2	359	suppositional subjunctive mood	гипотетическое наклонение
N	2	N	4	347	ritual language	язык ритуального речевого поведения
N	2	N	1	318	vowel mutation	переогласовка
N	3	N	4	277	subject relative clause	субъектное относительное придаточное предложение
N	3	N	5	119	phrase structure rule	правило развертывания по непосредственно составляющим
N	2	N	5	117	native signer	пользующийся языком жестов с рождения
N	4	N	4	100	standard average European Language	усредненный европейский литературный язык

Часть речи английского термина	Длина в словоформах	Часть речи русского соответствия	Длина в словоформах	Количество терминов такого типа	Пример английского термина	Пример русского термина
N	1	N	3	98	complementizer	морфема комплетивного отношения
A	2	A	2	69	context free	контекстно-свободный
A	2	A	1	57	East Slavic	восточнославянский
N	4	N	3	57	adjective-based adverbial phrase	отадъективная наречная составляющая

3.4. ПРОБЛЕМЫ СООТНЕСЕНИЯ ТЕРМИНОВ В МИКРООБЛАСТИ ТЕСТОЛОГИИ

Тестирование языковой компетенции представляет собой относительно новую подобласть прикладной лингвистики и лингводидактики с относительно автономно развивающейся терминологией. Особое значение корректность перевода и/или передачи терминов этого микрополя получила с внедрением в практику оценивания языковой компетенции и уровня достижений единого государственного экзамена. Некорректность некоторых широко используемых переводов, отсутствие специализированных словарей и справочников приводят к рассогласованию в понимании конкретных методических и лингвистических критериев, аспектов и принципов тестологии.

Сегодня в нашей стране принципам и практикам тестирования в целом начинает уделяться серьезное внимание. Необходимость

разработки подходов к способам контроля и оценивания была осознана с внедрением единого государственного экзамена (ЕГЭ). Результаты ЕГЭ являются де-факто пропуском для учащихся в высшие учебные заведения. Эта значимость ЕГЭ предъявляет его разработчикам повышенные требования с точки зрения основных характеристик теста. Особое внимание к ЕГЭ вполне оправдано и тем фактом, что ЕГЭ стал частью общественной жизни, а следовательно, население должно испытывать доверие к процедуре и результатам, которые представляет этот тест.

В зарубежной традиции единый государственный экзамен рассматривался бы как *high stake test*, т. е. тест, результаты которого имеют высокую степень важности для испытуемых. К созданию апробации, администрированию и безопасности таких тестов предъявляются особые требования. Тест, результаты которого имеют низкую степень значимости для испытуемых, в зарубежной традиции называется термином *low stake test*. Возможно, что учет степени значимости в описании единого государственного экзамена привел бы к тому, что уже на начальном этапе подход к ЕГЭ был бы иным. Рассмотрим (табл. 10) варианты перевода терминов, определяющих степень значимости теста.

Таблица 10

Варианты перевода терминов *high stake test* и *low stake test*

Русский термин	Термин на английском языке
high stake test	Особо значимый тест
low stake test	Незначимый тест

Внедрение ЕГЭ не было успешным, вызвав критику со стороны родителей, испытуемых, учителей, преподавателей вузов. Тем не менее ЕГЭ все же выполняет роль единого стандарта образования и является реализацией общего права на получение высшего образования, поэтому его форма проведения в виде теста является на данный момент единственно возможной. Основная проблема состоит не столько в формате ЕГЭ, сколько в том, какие данные этот экзамен предоставил, насколько качественным является собственно наполнение теста. Не будем забывать, что ЕГЭ создавался в условиях отсут-

ствия традиций разработки больших тестовых комплексов, поэтому неудивительно, что опыт ЕГЭ оказался не вполне удачным.

Рассмотрим далее проблемные области, которые были выделены по результатам ЕГЭ:

- результаты тестов не являются надежными;
- процедура тестирования не вызывает доверия у населения и экспертного сообщества;
- низкий уровень развития тестологической науки;
- недостаточное использование зарубежного опыта при разработке тестов.

Кроме этих проблемных областей отдельно стоит вопрос о терминологическом соответствии систем русской и зарубежной традиции. Рассмотрим некоторые аспекты сложности этих соотношений.

Прежде всего, попытаемся проследить соответствие между такими терминами, как *testing*, *assessment* и *evaluation* и русским термином *тестирование*. Отметим некоторые основные положения тестирования.

Тестирование должно предоставить информацию о следующем:

- об уровне владения иностранным языком (сведения о степени сформированности навыков и умений, в соответствии с выбранным критерием адекватности);
- о прогрессе, который сделал тестируемый для достижения поставленной цели;
- о знании языка (насколько оправдался прогноз испытуемых о собственном знании иностранного языка) [95, с. 10].

Исходя из типов информации, которые предоставляют процедуры тестирования, можно выделить функции, которые выполняет лингвистическое тестирование в обучении и жизни — контроль и сертификация уровня владения языком. Под сертификацией понимается процесс прохождения тестирования, построенного на авторитетных языковых стандартах и документации его результатов в случае преодоления необходимого минимума [116, с. 13–15]. Сегодня именно ЕГЭ является примером сертификации, которая оказывает серьезное влияние на будущее тестирование. Каждый этап теста, используемого для тестирования, требует точности при разработке теста, соответствия целям теста, процесса выверки и стремления к сбалансированности самого теста. Таким

образом, к понятию «тестирование» еще относится и процесс создания тестов.

Помимо этого тестирование также является методом контроля, который определяет самую суть тестирования. Это заставляет нас внимательно проанализировать понятие «контроль». Прежде всего, контроль рассматривается как обязательный компонент учебной деятельности. Благодаря контролю совершенствуются приемы управления процессом овладения знаниями, навыками и умениями на иностранном языке. Контроль позволяет получить обеим сторонам процесса обучения информацию о том, какова степень сформированности изучаемых компонентов, а также выявить основные трудности освоения языка [63, с. 81].

Контроль может также рассматриваться и как процесс определения уровня знаний, навыков, умений обучаемого, устанавливаемого в результате выполнения им устных и письменных заданий, тестов и формулирование на этой основе определенной оценки за определенный раздел программы, курса, периода обучения [2, с. 123].

При более общем подходе контроль понимается как процесс, в ходе которого должны быть получены результаты, позволяющие выявить на определенном этапе обучения уровень сформированности тех или иных навыков и осуществить коррекцию имеющихся ошибок у обучающихся. В первом определении принципиальным является включение тестовой практики в понятие контроля, во втором случае важным является упоминание о возможности коррекции ошибок обучаемых, что подразумевает временную протяженность осуществления контроля и его включение в систему обучения.

Временная протяженность контроля позволяет ему реализовывать следующие функции:

- функция обратной связи;
- функция управления;
- функция оценки;
- мотивационная функция.

Выделение функций контроля является очень важным моментом. Любая из функций контроля реализуется в различных формах контроля. И именно разное соотношение этих функций в процессе контроля и определяет отличия форм контроля друг от друга. Соотношение функций контроля может служить базой для определения

различных терминов, связанных с процессом тестирования (например, тест от контрольной работы при их формальном совпадении).

Для определения термина *тестирование* под контролем мы будем понимать вид учебной деятельности, направленный на установление индивидуального качественного уровня владения иностранным языком, на основе количественной оценки речевых действий обучаемого, выявления объектов для коррекционных мероприятий и разработки дальнейшей стратегии иноязычного образования [88, с. 23]. Выбор данной трактовки связан с тем, что в этом определении присутствуют все ключевые компоненты, которые составляют суть процесса контроля:

- включение в процессе учебной деятельности;
- оценка речевых действий обучаемых;
- возможность коррекции траектории обучения на основе полученной количественной и качественной информации;
- указание на все основные функции контроля, мотивации, оценки, управления.

Это определение поможет выявить и отрицательные формы контроля. Например, для объективности контроля важной является проблема субъективности оценки, что, в основном, зависит от выбора объекта контроля. Если объектом контроля является коммуникативная компетенция, то определенная степень субъективности будет заложена в контроль изначально [127, с. 17].

Определив термин *контроль*, мы можем перейти к анализу термина *тест*. Различные определения этого термина, как правило, имеют существенные недостатки. Так, например, в английском языке *test* может обозначать как тестовый комплекс, так и любую другую форму контроля, просто контролирующее мероприятие. По аналогии с английским языком термин *тест* используют для обозначения различных заданий, выполняемых в классах, включая перевод с русского на английский язык, ответы на вопросы и т. д. [89, с. 36]. Включение таких заданий в понятие *тест* приводит к его перегруженности, требующей постоянного уточнения в контексте.

Нельзя забывать о том, что тест, как и любая другая форма контроля, характеризуется целью, структурой, способом анализа результатов. При выборе одной из этих характеристик в качестве доминанты под тестом понимается:

- одна из наиболее доступных форм использования математических методов для оценки процесса усвоения знаний и формирование умений [40, с. 39];
- объективный контроль, охватывающий самые различные языковые аспекты, диагностирующий количественные и качественные отклонения от ожидаемого на конкретном этапе уровня владения языковыми шкалами [87, с. 39];
- метод контроля знания, позволяющий одновременно проверить всех испытуемых, которые выполняют данный тест в течение одинакового временного отрезка с одинаковым по объему и сложности материалом.

Данные определения представляются нам очень узкими, подчеркивающими только какой-то определенный компонент понятия «тест». Подобное сужение понятия выпускает из рассмотрения измерительную специфику теста и его структурные особенности. Данные трактовки представляют разрозненные реализации функций контроля. Для определения термина *тест* необходимо также рассмотреть его отличия от таких терминов, как *тестовое задание* и *задание в тестовой форме, контролирующее задание*. Итак, под тестом мы будем понимать форму контроля, использующуюся для выявления у испытуемых степени лингвистической и/или коммуникативной компетенции и представляющую комплекс заданий, который состоит из частей для тестирования отдельных видов речевой деятельности, прошел апробацию и предлагает получение результатов, которые могут быть оценены по установленным критериям [46, с. 40]. Тест может реализовывать все функции контроля. Это определение предпочтительно тем, что в нем задана цель тестирования, его структура, отмечена измерительная специфика, показана обязательность наличия четких критериев оценивания. Выделение этих необходимых критериев оценивания приводит нас к необходимости вычисления валидности и надежности теста, а также апробации тестовых комплексов. Это определение успешно вписывается в понятие контроля, которое мы рассмотрели ранее. Оно помогает разграничить понятия *тест* и *контрольная работа, тестовое задание* и *задание в тестовой форме* или просто *задание*. Контрольная работа будет отличаться от теста тем, что, во-первых, она контролирует сформированность коммуникативной компетенции на гораздо более узком участке, чем

тест, контрольная работа обычно включает в себя меньше материала. В отличие от контрольной работы тест является не просто набором заданий, но использует статистически просчитанное предположение о достоверной информации. Наличие апробации также говорит о том, что тесты имеют тенденцию проверять более широкий спектр компетенций, в то время как контрольная работа направлена на проверку специфической компетенции. Именно поэтому в зарубежной традиции термину *test* противопоставляется термин *quiz* для обозначения «контрольной работы».

На основе данного рабочего определения можно также развести понятия *тестового задания* и *задания в тестовой форме*. Тестовые задания будут являться частью измерения, у них будет свой вычисленный уровень сложности, они будут попадать в тест с серьезной доказательной базой в виде статистических данных. Задания в тестовой форме и тестовые задания также различаются с точки зрения цели. Тестовые задания, являясь частью теста, не только контролируют, но и мотивируют обучаемых. Использование же заданий в тестовой форме необходимо для реализации управляющей формы контроля и функции обратной связи. Для этого удобна тестовая форма. Но использование тестовой формы не сделает из контрольной работы надежный измерительный инструмент. Это важное различие позволяет нам разобраться с употреблением таких терминов, как *test task, item* (тестовое задание), *task* (задание или задание в тестовой форме).

Еще одну группу терминов, которая требует серьезных комментариев, составляют варианты и соотношения терминов *assessment, evaluation* и *testing*. В широком смысле эти слова являются синонимичными, тем не менее следует отметить наличие важных различий. Так, наиболее общим из этих трех терминов является термин *evaluation*, который обозначает целые системы контроля для курсов, учебных программ. *Assessment* и *testing* используются более узко, для именованья использования конкретных средств в процессе контроля.

При этом понятию *assessment* более соответствует термин оценивание, обозначающий варианты контроля, которые не входят в структуру тестирования. Термин *testing* следует переводить как тестирование, в том значении, которое было рассмотрено выше.

Результаты сопоставления русских и английских терминов представлены в табл. 11.

Таблица 11

Сопоставление основных терминов, связанных с понятием тестирования

Русский термин	Термин на английском языке
test	тест
evaluation	контроль
assessment	оценивание
testing	тестирование
quiz	контрольная работа
test task, item	тестовое задание
task	задание или задание в тестовой форме

От четкости определения основных характеристик теста и связанных с ними понятий зависит установление соотношений между другой группой терминов: *validity*, *валидность*, *надежность* и их производные. Рассмотрим эти основные характеристики. Традиционно применительно к тестам используют две основные характеристики — валидность и надежность. Первая из них представляет характеристику теста, которая исследует тест на его эффективность в различных его аспектах. В самом общем виде, валидность определяют исходя из соответствия результата теста с одной или несколькими заданными целями. Например, тест достижений должен представлять все данные для того, чтобы сделать вывод о достижениях испытуемых [91].

Процедурам валидации должно подвергаться содержательное наполнение теста. В ходе этой процедуры мы должны понимать, насколько полно тест охватывает материал, для проверки которого он был создан. Как правило, в лингвистических тестах содержательная валидность определяется использованием соответствующего языкового материала.

Кроме того, валидность исследуется в области конструкта (*test construct*). Под конструктом понимается способность или признак, который измеряется в ходе теста. Тестовые задания должны стремиться раскрыть требуемые свойства конструкта и следовать его основным психофизическим характеристикам.

При разработке теста важно понимать, что тест может оказывать на испытуемого стрессовое воздействие. В тесте не должно быть ничего такого, чтобы вызывало у испытуемого стресс. Задания должны быть правильно расположены и оформлены, инструкции должны быть четкими и ясными, сам формат тестирования должен быть нейтральным. Здесь оценка валидации связана с внешними факторами теста.

Таким образом, следует различать разные типы валидности: в области содержания тест должен соответствовать содержательной валидности, в области конструкта — конструктивной валидности, в области соответствия результата и цели — сопоставительной или прогностической валидности, в области оформления — внешней валидности [92]. Описанным видам валидности соответствуют следующие англоязычные термины и их русские эквиваленты (табл. 12).

Таблица 12

Соответствие английских и русских терминов, связанных с валидностью

Термин на английском языке	Русский термин
criterion-related validity	валидность внешнего критерия
discriminant validity	дискриминантная валидность
divergent validity	дивергентная валидность
convergent validity	конвергентная валидность
content validity	содержательная валидность
construct validity	конструктивная валидность
prognostic/ predictive validity	прогностическая валидность
concurrent validity	сопоставительная валидность
face validity	внешняя валидность

В табл. 12 представлены те английские термины, которые в нашей традиции употребляются редко. Прежде всего, это термин *criterion-related validity*, которому соответствует перевод *валидность внешнего критерия*. Фактически, этот термин может быть использован как синоним термину *сопоставительная валидность*,

которая может обозначать сравнение с определенным внешним критерием.

Другие три термина (*divergent validity*, *discriminant validity*, *convergent validity*) связаны с понятиями дивергентной и конвергентной валидности, они называют аспекты конструктивной валидности, показывая, что данный тест находится в близкой связи с результатами теста, измеряющего схожий конструкт (конвергентная валидность) и насколько далеко он находится от результатов теста, который измеряет несходный конструкт (дивергентная валидность, дискриминантная валидность).

Каждый из рассмотренных терминов представляет валидность как сумму различных характеристик, как составные части общей валидности теста. Такой подход требует определения валидности как единой тестовой характеристики, что позволяет обеспечивать социальную значимость тестирования.

Обобщенный подход к валидности, состоящий в том, что валидность представляет собой единство содержательного, процессуального структурного обобщенного внешнего и последовательного аспектов, предлагает, в частности, С. Мессик [119, с. 75–83].

Еще одной важной характеристикой оценки тестов является *надежность*. Для начала определим то, что мы понимаем под надежностью. Надежность — это характеристика, которая позволяет определить точность и устойчивость измерений теста. Надежность характеризует тест, с одной стороны, с позиции его баланса, т. е. тест должен равномерно распределять задания, чтобы не было очевидных провалов с точки зрения правильности/неправильности. С другой стороны, надежность показывает, насколько тест зависит от условий его проведения и зависит ли. Надежность для теста необходима, но надежность не считается достаточным условием для того, чтобы считать тест валидным. Более того, вполне возможна ситуация, когда валидный тест не обладает высокой степенью надежности.

Обыкновенно степень надежности выражается в коэффициенте надежности [106, с. 370]. Чем ближе этот показатель к единице, тем более надежным считается тест. Однако следует отметить, что коэффициент надежности для тестов, измеряющих разные навыки и умения, будет существенно отличаться. В частности, тест, измеряющий грамматические навыки, будет считаться достаточно надежным, если

он будет иметь показатель коэффициента 0.95. А для теста, проверяющего умения говорения, высокой степенью надежности будет считаться результат 0.85.

В зарубежной терминологии термину *надежность* соответствует термин *reliability*. С термином *reliability* связан ряд терминов, которые требуют уточнения: *split half reliability* и *inter-rater reliability*. Термин *inter-rater reliability* соответствует термину *надежность рецензоров*. Под надежностью рецензоров понимается степень доверия к тому, как эксперты оценивают поведение тестируемых. Особенно актуальна разработка этого термина в рамках поиска критериев и процедур коммуникативного оценивания.

Термин *split half reliability* соответствует методике расщепления теста. Методика предполагает, что тест должен быть разделен на две половины. Если коэффициент корреляции между двумя этими половинами близок к единице, то делается вывод о том, что тест является надежным.

Интересным видится ситуация с вариантами перевода термина Item Response Theory (IRT). Следование концепции этой теории сделало возможным разработку тестовых комплексов, предоставляющих надежную информацию о конструктах теста. Оценивание, осуществляющееся по законам IRT, требует от разработчиков измерительных мероприятий доказательств эффективности применяемых средств и их связи с целями, одним из которых является надежность.

Перевод терминологической аббревиатуры IRT предлагает много разных вариантов. Эти варианты связаны с тем, что популярность и распространение теории сделали его скорее брендом, номеном, чем терминологическим сочетанием. Варианты перевода на русский язык представлены в табл. 13.

Таблица 13

Варианты перевода аббревиатуры IRT на русский язык

Термин	Перевод
Item Response Theory	теория латентных черт
	теория характеристических кривых заданий
	теория моделирования и параметризации педагогических тестов

Термин	Перевод
Item Response Theory	современная теория тестов
	теория ответа на задание
	математическая теория педагогических измерений

Рассмотрим эти эквиваленты. Перевод этого термина как *теория ответа на задание* выглядит как просто пословный перевод тех слов, которые входят в название. Полученный русский вариант не раскрывает специфику этой теории. Именно поэтому эквивалент использовать не рекомендуется.

Термин *современная теория тестов* также считается не очень удачным, поскольку эта теория является не единственной, она имеет свои достоинства и недостатки. Кроме того, другие теории создания тестов пока еще не потеряли возможности развития. Да и само понятие «современности» является достаточно условным.

Использование в качестве варианта перевода *теории латентных черт* имеет серьезные недостатки. Дело в том, что этот вариант уводит нас от основной концепции теории, концентрируясь на том, что в широком смысле теория IRT — это частная реализация методов латентно-структурного анализа. Но, как мы определили выше, концепция теории IRT включает не только математико-статистический аппарат. *Теория моделирования и параметризации тестов* — также неудачный вариант для перевода названия теории и фиксации его терминологического статуса. Так как следует учитывать, что теория IRT рассматривается гораздо шире, чем просто теория создания тестов, сегодня она включает в себя и оценивание. Самым лучшим представляется вариант *математическая теория педагогических измерений*. Этот вариант предполагает оценивание как измерение, в нем отмечена специфика метода измерения, выделена необходимость педагогической составляющей.

Мы рассмотрели лишь некоторые примеры сложностей, возникающих при сопоставлении терминов в области лингвистического тестирования. Без сомнения, требуется дальнейшая работа по уточнению наполнения терминов в сфере тестирования. Гармонизация терминологии в этой области приведет к более простой адаптации и использованию зарубежного опыта тестирования в российских ус-

ловиях. Важным следствием такой гармонизации станет пересмотр принципов, которые сегодня используются отечественными специалистами для составления тестов.

3.5. ХАРАКТЕРИСТИКА СЛОВАРЯ ПОДЪЯЗЫКА «ПРИКЛАДНАЯ ЛИНГВИСТИКА», ИЗВЛЕЧЕННОГО ИЗ МАССИВА СПЕЦИАЛЬНЫХ ТЕКСТОВ

Язык филологии, как все специальные языки, включает как общенаучную (*проблема, система, анализ, синтез*), общелингвистическую (*язык, текст, парадигма, синтагма, синтаксис*), так и отраслевую (*фонема, аллофон, граммема, синтаксема, жанр, стиль, поэтика, разметка, корпус, лингвистический автомат*) терминологическую лексику. То же верно и для подъязыков филологии, при этом очевидно, что специальные узкоотраслевые тексты, тематика которых к тому же ограничена условиями дискурса (материалы конференций, тематические сборники или предметно-ориентированные журналы), так называемые псевдопараллельные тексты, представляют собой релевантную эмпирическую базу для извлечения специализированной терминологии и создания отраслевого словаря. В качестве такой базы был создан массив (корпус) текстов научных докладов VI Международной конференции «Прикладная лингвистика в науке и образовании», проходившей 5–7 апреля 2012 г. в Санкт-Петербурге.

На основе массива, объем которого около 60 000 с/у, был определен состав микрополя «прикладная лингвистика», включающий 8 вершинных узлов, соответствующих 8 направлениям прикладной лингвистики, проблемы которых составили тематику конференции (рис. 5).

Структура микрополя открывает возможность описания частных подъязыков, например подъязыка лингводидактики или подъязыка лингвостатистики. Однако на данном этапе поставим задачу более широко: определим некоторые общие характеристики извлеченного списка терминов и другой специальной лексики (в частности номенов) с точки зрения задач наполнения отраслевого словаря, а именно:

- каноничности (общепринятости) термина;
- представленности в других лингвистических словарях;

