

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное
учреждение высшего образования

Национальный исследовательский университет
«Высшая школа экономики»

Факультет гуманитарных наук
Образовательная программа
«Фундаментальная и компьютерная лингвистика»

КУРСОВАЯ РАБОТА

На тему «Разработка теста на знание лингвистической
терминологии в Item Response Theory»

*Тема на английском: Development of a Linguistic Terminology Test in
the Item Response Theory Approach*

Студент 3 курса
группы БКЛ161
Змановский Никита
Сергеевич

Научный
руководитель
Козлов Алексей Андреевич
старший преподаватель
Школы лингвистики

Москва, 2019 г.

Оглавление

Оглавление	2
1. Введение	2
1.1. Как оценить знание предметной области?	2
1.2. Знание предметной области равно знанию её терминов?	3
1.3. Прикладное обоснование исследования и постановка задач	5
2. Обзор существующих подходов	6
2.1. Тестирование	6
2.2. Тестирование для оценки словарного запаса	6
3. Разработка теста на знание лингвистической терминологии	7
3.1. Составление терминологического словарика	7
3.2. Выбор терминов для включения в тест	8
3.3. Формулировка задания и составление опросника	11
4. Анализ результатов пилотного исследования теста на знание лингвистической терминологии ..	12
3.4. Результаты пилотного исследования	12
3.5. Анализ результатов пилотного исследования	12
3.6. Ограничения и перспективы развития теста	12
5. Заключение	13
Литература	13
Приложение	14

1. Введение

1.1. Как оценить знание предметной области?

Перед любым преподавателем рано или поздно встаёт вопрос: как оценить эффективность своих занятий? Эффективным в некоторой деятельности можно назвать человека, который добивается в ней хороших результатов.

Образовательными результатами часто называют знания, умения и навыки, полученные учениками и ученицами в процессе обучения. Значит, эффективность преподавания можно оценить как функцию от образовательных результатов.

Что при этом важнее: знания или навыки? Для определённых областей учения довольно легко сформулировать ожидаемые образовательные результаты в терминах умений и навыков. Например, в обучении игре на музыкальных инструментах или плаванию. Однако для многих гуманитарных наук такие формулировки даются труднее. Формирование каких, к примеру, умений и навыков ожидается от окончивших курс философии? При этом определение

знаний, которые учащиеся должны получить в ходе такого курса, вызывает гораздо меньше затруднений (хотя может не быть консенсуса у разных специалистов насчёт конкретного набора фактов и сведений, необходимых для успешного освоения курса).

В рамках данной работы мы будем исходить из презумпции первичности знаний как основного образовательного результата, по крайней мере при освоении такой гуманитарной науки, как лингвистика. Иными словами, мы будем предполагать, что, оценив уровень знаний в данной области, мы сможем разделить учащихся по уровню компетентности в лингвистике, не оценивая при этом их умения и навыки, например, навыки лингвистического анализа языковых единиц разных уровней, фонетической транскрипции, построения синтаксических деревьев.

Несомненно, такое сильное предположение необходимо в конечном итоге поставить под сомнение и дополнительно проверить. Мы ожидаем, что по результатам валидации нашей методики оценки компетентности в лингвистике мы сможем сказать, противоречат ли факты нашему исходному предположению.

Чтобы оценить качество полученных знаний, очень часто прибегают к тестированию — стандартному способу контроля знаний в педагогике. Однако у предметных тестов есть неизбежные проблемы: трудность в составлении равных по сложности вариантов тестовых заданий, невозможность охватить весь пройденный материал в содержании заданий, шанс списывания.

Мы предлагаем новый способ контроля знаний — тест на знание терминологии, который сформирован по образцу тестов словарного запаса в рамках современной теории тестирования (Item Response Theory). Для создания такого теста используются достижения современной лингвистики и психометрии.

1.2. Знание предметной области равно знанию её терминов?

Многие области современной лингвистики прочно связаны с когнитивной психологией. Лингвистические исследования позволяют пролить свет на то, как происходит процесс познания, как языковые знания — ментальный лексикон и грамматические умения — встраиваются в общую структуру знаний человека.

Как лингвистика может помочь в разработке тестов, оценивающих компетентность в некоторой предметной области?

Поскольку мы исходим из предположения, что ключевым образовательным результатом в процессе обучения некоторым гуманитарным наукам (в частности лингвистике) являются знания, можно сделать ещё допущение и предположить, что знание некоторой предметной области равно знанию терминов данной области. Термины — это слова, единицы лексического уровня данного языка, хотя и особые по свойствам. Значит, для оценки знания термина мы можем воспользоваться теми же методами, которыми пользуются психолингвисты и когнитивные лингвисты для оценки знания слова (word knowledge).

Разумно предположить, что чем шире круг терминов, знакомых некоторому человеку, тем больше у него знаний. Мы будем называть круг терминов, знакомых человеку, терминологическим лексиконом данного человека. Терминологический лексикон можно уподобить лексикону как индивидуальному словарному запасу человека, так как, в сущности, терминологический лексикон есть часть общего лексикона человека. В таком случае можно воспользоваться существующими методиками оценки словарного запаса для оценки терминологического лексикона.

Предположение о том, что знание некоторой области — это то же, что знание терминов данной области, не менее сильное, чем наша исходная презумпция, в некоторой степени даже сильнее. Каковы наши аргументы в пользу такого предположения?

Во-первых, можно отметить, что память о различных фактах, сведениях, осознанно усвоенной информации часто вербализована. В современных когнитивных моделях долговременной памяти выделяют два основных модуля — эксплицитная память (декларативная память) и имплицитная память. Эксплицитная память содержит сведения, которые доступны сознанию, их можно намеренно (осознанно) вспомнить (Reisberg 2013). Нас больше всего интересует такой блок эксплицитной памяти, как семантическая память, хранящая сведения о мире, объектах, концептах и понятиях. Существует множество различных моделей семантической памяти (Jones et al. 2015), но многие из них исходят из предположения о том, что знания о мире в человеческой памяти часто связаны с конкретными словами и концептами. Например, сетевая модель семантической памяти Teachable Language Comprehender представляет семантическую память в

виде сети, в узлах которой находятся слова, отображающие некоторые концепты (Collins, Quillian 1969).

Во-вторых, существующие инструменты компьютерной лингвистики позволяют создавать базы знаний в некоторой области — лингвистические онтологии — на основе автоматического анализа текстов на естественном языке, в том числе через анализ терминологии, встречающейся в этих текстах (De Nicola et al. 2009). Подобные методы инженерии онтологий также основаны на предположении, что структуру знаний можно уподобить некоторой структуре из языковых единиц, репрезентирующих знания. Значит, в прикладной лингвистике уже используется редукция системы знаний к системе терминов, репрезентирующих эти знания.

Каково на самом деле устройство структуры наших знаний некоторой предметной области, действительно ли его можно представить в виде структуры, изоморфной лексикону, — это отдельные актуальные вопросы для исследований в психолингвистике и когнитивной науке, которые выходят за рамки нашей работы.

1.3. Прикладное обоснование исследования и постановка задач

Целью данной работы является **создание прототипа теста на знание лингвистической терминологии** по модели тестов на определение словарного запаса. Исходя из основного предположения данной работы, такой тест позволял бы оценить уровень лингвистической компетенции человека, не прибегая к специальным вопросам «на знание».

Новизна работы состоит в использовании в качестве матричной модели тесты словарного запаса. Предположение о том, что структуру знаний некоторой области можно уподобить структуре её тезауруса, не ново, однако применение этой презумпции с прикладной целью составления теста относительно ново.

В качестве предметной области была выбрана лингвистика в связи с личными интересами авторов данной работы.

Целевая аудитория теста — школьники и школьницы старших классов, которые принимают участие в школьных предметных олимпиадах по русскому языку и лингвистике. Естественно предположить, что у таких школьников и школьниц размер терминологического лексикона в области лингвистики будет

больше, чем у сверстников, не занимающихся углублённо лингвистикой и русским языком. В перспективе тест можно провести одновременно с олимпиадными состязаниями, оценив надёжность связи результатов теста и результатов олимпиадных состязаний.

Небольшое количество времени, которое тратится на тест, а также возможность его повторного прохождения позволяют использовать тест одновременно как средство входного контроля и как средство оценки образовательных результатов.

2. Обзор существующих подходов

2.1. Тестирование

Существует множество разных способов контроля знаний и оценки того, достигнуты ли образовательные результаты: устные экзамены, сдача нормативов (для обучения в спорте и др.), демонстрация уровня мастерства, созданных в процессе обучения работ (для обучения в сфере искусства), тесты и др. Один из самых распространённых способов — стандартизированное тестирование.

Стандартизированное тестирование может предполагать разные формат тестовых вопросов. Например, это могут быть вопросы с выбором из множества ответов (multiple-choice task), а также вопросы с вариантами ответа «да/нет» (yes-no questions).

2.2. Тестирование для оценки словарного запаса

Оценить словарный запас, то есть размер лексикона человека, довольно трудная задача. Самый точный способ такой оценки — использовать какой-нибудь словарь знакомого человеку языка и попросить его для каждого слова в словаре отметить, знает он это слово или нет. Проблема такого способа оценки словарного запаса состоит в том, что его очень трудно использовать на практике: временные затраты на подобную процедуру огромны и едва ли сопоставимы с выигрышем в точности по сравнению с другими методами, аппроксимирующими размер словарного запаса по какой-либо иной величине, которую проще эмпирически установить.

Современные тесты словарного запаса исходят из предположения, что знание человеком лексемы определённой частотности (можно определить по

корпусу языка) с высокой вероятностью свидетельствует о том, что человек знает и другие слова, которые встречаются в языке примерно с той же частотой, а также наверняка знает слова, которые встречаются чаще (Read 2000). На практике это означает, что при оценке словарного запаса можно оценивать знание не всех слов языка, а знание выборки из N слов, которая репрезентативна относительно структуры лексической системы данного языка (например, в ней есть мало высокочастотных слов и много низкочастотных). Тогда можно установить некоторую функцию, устанавливающую соответствие между знанием слов из выборки, которые являются словами одного частотного диапазона, и размером лексикона.

Как понять, знает ли человек некоторое слово? Разные люди могут вкладывать разный смысл в концепт «знать слово». В лингвистике под знанием слова часто понимают знание множества разнородной информации о слове: его произношении, семантике, сочетаемости, морфологическом статусе и др. Нас прежде всего будет интересовать знание семантики слова, поскольку наша работа исходит из предположения о том, что концепты, понятия, декларативные знания можно редуцировать до слов, их репрезентирующих, мы отвлекаемся от внешней формы слов и их особенностях, важных для языковой системы, но не отражающих особенности системы знаний, которые они репрезентируют.

В своей работе мы опирались на тест словарного запаса, представленный проектом MyVocab.info (Головин 2015).

3. Разработка теста на знание лингвистической терминологии

3.1. Составление терминологического словника

В качестве основы для составления словника мы взяли авторитетный словарь лингвистической терминологии — «Лингвистический энциклопедический словарь» под ред. В. Н. Ярцевой (ЛЭС). Выбор этого словаря обоснован его энциклопедичностью, широтой словника, включающего в том числе и очень специальные, малоизвестные лингвистические термины.

Хотя основная аудитория теста — школьники и школьницы, мы решили не основывать наш словник на школьных лингвистических словарях, поскольку термины, толкуемые в этих словарях, часто не выходят далеко за рамки школьной

программы по русскому и иностранным языкам. Из этого следует, что дифференцирующая способность заданий, основанных на подобном словнике, будет ниже.

Мы не использовали в работе известный в педагогической среде «Энциклопедический словарь юного филолога (языкознание)» (Панов 1984), поскольку его словник значительно меньше необходимого для нашей целей размера, многие термины, которые есть в ЛЭС в виде отдельных статей («Эллипсис») или в тексте других статей (клитика), ни разу не встречаются в тексте «Энциклопедического словаря юного филолога (языкознание)».

3.2. Выбор терминов для включения в тест

Одна из ключевых задач, стоящих перед нами, состояла в отборе терминов для теста. Разумеется, не все термины, которые толкуются в ЛЭС, подходят для составления теста, ориентированного на старшеклассников и старшеклассниц, пусть и занимающихся углублённо лингвистикой и русским языком.

Был проанализирован весь список терминов, которые толкуются в отдельных (чаще всего одноимённых им) статьях ЛЭС. В ЛЭС насчитывается 1399 словарных входов — от статьи «Абазинский язык» до статьи «Яфетическая теория».

Из всех словарных входов при отборе были исключены следующие заголовки статей:

- заголовки, не являющиеся терминами («Язык и мышление», «Язык и общество»),
- заголовки статей, состоящие из чересчур больших по объёму терминов («Понимание речи»),
- общенаучные термины, не имеющие особого лингвистического смысла («Эстетический идеализм», «Статистический метод»),
- некоторые общенаучные термины, имеющие слишком прочные внелингвистические ассоциации, даже если у них существует специальное лингвистическое значение («Аналогия», «Атрибут»),
- названия языков и таксонов в генеалогической классификации языков («Абазинский язык», «Финно-угорские языки»), а также названия частных лингвистик («Арабистика», «Индология»),

- названия лингвистических школ, научных объединений и других культурных реалий, упоминаемых в контексте истории лингвистических учений («Академия наук СССР», «Античная традиция», «Виноградовская школа», «Казанская лингвистическая школа»),
- названия различных языковых законов и тенденций («Лескина закон», «Гримма закон»),
- названия большинства систем письма (исключение было сделано для статей «Кириллица», «Глаголица»).

Из групп однокоренных терминов, объединяющихся в тематические группы, был выбран только один термин. Например, из группы терминов «Агглютинативные языки», «Агглютинация теория», «Агглютинация» был отобран только термин «Агглютинация».

На этапе первичного отбора было отсеяно довольно много терминов. Из общего числа 1399 словарных входов ЛЭС в качестве терминов-заданий был отобрано 360 терминов.

К сожалению, такое количество терминов всё равно оказалось слишком велико для нашей цели. Мы решили пойти по пути сокращения списка за счёт пропорционального его уменьшения в 3 раза — до 120 терминов.

Все термины были условно расклассифицированы по тематическим группам:

1. Фонетика, фонология, просодия. 57 терминов.
2. Морфемика, морфонология, некоторые аспекты теории словообразования. 32 термина.
3. Морфология, грамматическая семантика, названия грамматических категорий и грамматических значений, грамматическая типология языков. 73 термина.
4. Синтаксис, синтаксические связи в предложении, члены предложения. 42 термина.
5. Лексический уровень языка, лексикология, лексическая семантика, семантические отношения. 52 термина.

6. Стилистика, риторика, теория коммуникации, литературная и нелитературные формы бытования языка. 32 термина.
7. Письменность, теория и история письма, системы письма, особые письменные символы. 37 терминов.
8. Общая теория языка, общее языкознание, общие для разных областей лингвистики термины. 35 терминов.

Каждый элемент мог попасть только в одну группу. В спорных случаях решение о присвоении термину той или иной тематической группы принимал разметчик.

На этапе вторичного отбора из получившегося набора размеров в 360 терминов мы отобрали в 3 раза меньше, то есть 120 терминов, которые попали в итоговую версию теста. Отбор шёл пропорционально размеру каждой тематической группы в терминах. Например, в тематической группе «Фонетика» из первичных 57 терминов в итоговую версию теста были отобраны 19 терминов. Полный список из 120 отобранных терминов доступен в Приложении.

Вторичный отбор терминов основывался на личной языковой и педагогической интуиции авторов работы с тем условием, чтобы в каждой тематической группе было как минимум 2 «простых» и 2 «сложных» термина.

Простыми мы считали термины, включённые в школьную программу по русскому языку и иностранным языкам, они же с высокой вероятностью встретятся в школьных лингвистических словарях (например, «Вводные слова»).

Сложными мы считали термины, которые встречаются в текстах для школьников довольно редко по причине своей специальности (например, «Сандхи»).

Остальные принципы отбора не были сформулированы строго, что является одним из существенных ограничений данной работы.

Чтобы проконтролировать внимательность отвечающих и предотвратить «механическое» прохождение теста, 8 терминов из теста (по одному из каждой тематической группы) были заменены на псевдослова — последовательности символов, которые внешне напоминают лингвистические термины, на основе которых они составлены (например, вместо «Имперфект» в тест добавлено слово «Импрефект»). Если человек, проходящий тест, отметит такой термин как тот,

который он знает, то результаты данного человека будут считаться недостоверными, поскольку мы не можем сказать наверняка, не отмечал ли он другие слова по тому же принципу, то есть неосознанно, не проверяя себя дополнительно.

3.3. Формулировка задания и составление опросника

Пилотная версия опросника была запущена на платформе «Google Формы» и включала вопросы социодемографического характера (возраст, гендер), а также вопросы об участии в олимпиадах, помимо собственно теста на знание лингвистической терминологии.

Задание к тесту было сформулировано следующим образом:

«Задание: для каждого термина отметьте, знаете ли вы его.

Что значит «знать термин»? Будем считать, что это подразумевает следующее:

- 1. вы хотя бы раз видели/слышали термин,*
- 2. вы можете дать примерное толкование термина своими словами без обращения к терминологическому словарю и (желательно) вы можете привести хороший пример, иллюстрирующий данный термин.*

Например, прочитав термин «плеоназм», вы можете без словаря вспомнить, что так называют речевое излишество — ситуацию, когда в речи появляются лишние слова (вроде «июнь-месяц» или «масло масляное»).

NB! В тесте задаётся вопрос только про лингвистическое значение термина. Очевидно, что под «корнем», «обстоятельством» и «эрозией» здесь понимаются только лингвистические понятия, а не быденные понятия или термины других наук.

Пожалуйста, НЕ отмечайте как знакомые те термины, которые вы где-либо видели/слышали, но значения которых не знаете! В рамках нашего теста знакомство с внешней формой термина не считается его знанием.

В тесте есть несуществующие термины, будьте внимательны. Если вы отметите несуществующий термин, ваши результаты будут недействительны».

Такая формулировка акцентировала внимание на том, что под знанием слова (термина) в контексте теста понимается именно знание семантики данного

лингвистического термина, концепта, который за ним стоит, но необязательно знание других его особенностей (произношения, сочетаемости, др.).

4. Анализ результатов пилотного исследования теста на знание лингвистической терминологии

3.4. Результаты пилотного исследования

В пилотном исследовании приняло участие 63 человека. Из них мы отобрали тех, кто сейчас учится в школе, — 34 человека. Затем мы исключили из анализа результаты тех школьников и школьниц, кто отметил, что знает какой-либо из несуществующих псевдотерминов. Было сделано исключение для псевдотермина «глагольница», поскольку его отметили как знакомый сразу 10 респондентов (очевидно сильное влияние термина «глаголица», на основе которого и создан данный псевдотермин). После такого отсева остался 31 человек. Анализ проводился для данной выборки ($N = 31$).

3.5. Анализ результатов пилотного исследования

Мы провели статистический анализ полученных результатов с помощью latent trait моделей (LTM) современной теории тестов. Были протестированы две модели: модель с параметром угадывания (guessing parameter) и без такого параметра. Согласно тесту соотношения правдоподобия, модель с параметром угадывания лучше описывает полученные данные.

Результаты пилотного исследования однозначно свидетельствуют о том, что термины в нашем тесте отличаются по сложности для респондентов. В рамках модели с параметром угадывания мы оценили параметры сложности и угадывания для каждого термина из нашего теста. Самыми сложными терминами оказались термины «изафет», «референция» и «гипонимия». Подробные результаты анализов доступны в Приложении.

3.6. Ограничения и перспективы развития теста

Основное ограничение данной работы связано с тем, что многие операции по отбору терминов опирались на интуицию авторов. Перспективным направлением дальнейших исследований и разработок может быть создание алгоритмов по автоматическому составлению словников на базе онтологий и

частотных словарей тематических корпусов (например, корпус научных и научно-популярных текстов по лингвистике).

Другим важным ограничением теста является его унифицированность. Для многих подобных тестов современная теория тестов предлагает надстройку в виде адаптивного компьютерного тестирования, в ходе которого материал для тестирования подбирается по ходу процесса тестирования с учётом успешности прохождения предъявленных респонденту тестовых заданий. В контексте нашего теста это возможно, если последовательно увеличивать сложность терминов, предлагая сначала «простые» термины, а затем всё более и более «сложные» — до тех пор, пока вероятность знания новых предъявляемых в тесте терминов не приблизится к нулю (для данного конкретного респондента).

5. Заключение

В результате работы по данному проекту был подготовлен и протестирован прототип теста на знание лингвистической терминологии, разработанный по модели тестов словарного запаса, который может в дальнейшем использоваться в педагогической практике как средство оценки лингвистических знаний.

Литература

- Головин 2015 — Г. В. Головин Измерение пассивного словарного запаса русского языка // Социо-и психолингвистические исследования. – 2015. – №. 3. – С. 148-159.
- ЛЭС — Лингвистический энциклопедический словарь / Гл. ред. В. Н. Ярцева. — М.: Сов. энциклопедия, 1990. — 683 с.
- Панов 1984 — Энциклопедический словарь юного филолога (языкознание) / Сост. М. В. Панов. — М.: Педагогика, 1984. — 352 с., ил.
- De Nicola et al. 2009 — A. De Nicola, M. Missikoff, R. Navigli. A software engineering approach to ontology building // Information systems. – 2009. – V. 34. – №. 2. – P. 258-275.
- Jones et al. 2015 — M. N. Jones, J. Willits, S. Dennis. Models of semantic memory // Busemeyer J. R. et al. The Oxford handbook of computational and mathematical psychology. – Oxford University Press, 2015. – P. 232-254.
- Collins, Quillian 1969 — A. M. Collins, M. R. Quillian. Retrieval time from semantic memory // Journal of verbal learning and verbal behavior. – 1969. – V. 8. – №. 2. – P. 240-247.
- Read 2000 — J. Read. Assessing Vocabulary. Cambridge University Press. 2000. 294 p.

Reisberg 2013 — D. Reisberg (ed.). The Oxford handbook of cognitive psychology. – Oxford University Press, 2013. – P. 206-216.

Приложение

Исходные данные пилотного исследования, скрипт для анализа, а также полный список терминов, которые использовались в тесте, доступны в специальном репозитории на платформе GitHub:

<https://github.com/vekkle/lingvotest> .