

УДК 930.272:2-264

В.М. Лурье

**СОЛУНСКАЯ ЛЕГЕНДА:
ОРИГИНАЛЬНОЕ СЛАВЯНСКОЕ ПРОИЗВЕДЕНИЕ
ИЛИ ПЕРЕВОД С СИРИЙСКОГО?
КОЛИЧЕСТВЕННЫЙ ПОДХОД К ОЦЕНКЕ ПРАВДОПОДОБНОСТИ
АЛЬТЕРНАТИВНЫХ ГИПОТЕЗ¹**

1. Введение

Для научных дискуссий в гуманитарных науках довольно типичны две одинаково неконструктивные манеры ведения: либо отказывать гипотезам в сколько-нибудь серьезном праве на обладание истиной, либо считать все гипотезы, пока они не доказаны, одинаково плохими (или пусть даже одинаково хорошими: тут важно, что «одинаково»). Обе манеры сводятся к тому, что научные споры, опирающиеся на гипотезы, якобы, не имеют серьезных перспектив.

Мы не будем говорить о том, что такое «доказательство» в гуманитарных науках, то есть чем гипотеза отличается от теории. Сделаем вид, будто мы это понимаем. Нас будет интересовать оценка таких гипотез, которые никому и не представляются доказанными.

Ниже я постараюсь предложить некоторую методику, основанную на известных положениях индуктивной логики, но адаптированную к материалу гуманитарных исследований. В гуманитарных исследованиях, хотя, как правило, не проводят экспериментов, но точно так же, как, допустим, в астрономии делают наблюдения. И нельзя сказать, что с результатами этих наблюдений не работают количественными методами. Как раз работают, хотя отчасти в духе немецкой шутки о сути научного метода: *einmal ist keinmal, zweimal ist immer*. Если правда, что эту поговорку придумали крестьяне для интерпретации своих естественнонаучных наблюдений над явлениями природы, то это лишний раз напоминает нам, что и естественные науки в недалеком прошлом были такими же, как и гуманитарные в настоящем. В естественных науках (за исключением астрономии) только недавно, в XVII в., как-то научились преодолевать приблизительный характер любого измерения реальной величины и приходиться от неточных измерений к математически точным выводам. В гуманитарных науках это оказалось не так просто, потому что и приблизительность измеряемых параметров, как правило, выше: она, в целом, соответствует принципиальной нечеткости человеческого мышления и естественного языка. Но сегодня мы знаем, что расчеты можно вести и с нечеткими величинами.

Практический пример, который будет разбираться ниже, разумеется, взят из моих собственных научных тем. В качестве примера он оказался очень удобен. Неудивительно, что потребность в оценке гипотез, относящихся к этому примеру, привела к разработке предлагаемой здесь методики.

¹ В настоящей статье использованы материалы исследования, выполняемого при поддержке Российского фонда фундаментальных исследований; проект № 13-06-00832, «Описание логики естественного языка как системы модальных операторов».

Методика разработана только в виде эскиза, когда многие решения легко принять *ad hoc*, тогда как следовало бы их рассматривать на разных примерах и обсуждать подробности теоретически. Вместо таких теоретических исследований я буду просто указывать те логические теории, которые, на мой взгляд, тут следовало бы применить. Я буду очень рад, если кто-нибудь воспользуется этими идеями для уже профессиональной разработки статистических методов оценки степени достоверности логических выводов в гуманитарных науках.

2. Индуктивная логика и научные гипотезы

Индуктивная логика позволяет оценивать правдоподобие некоторого утверждения в зависимости от правдоподобия предыдущего утверждения. Это, в свою очередь, позволяет ее применить для оценки прочности логических цепочек в науке. В частности, для оценки различных научных гипотез. В этом разделе мы дадим лишь самый общий (без формул) очерк соответствующего подхода (более подробно и с формулами см. ниже, особенно в разделе 10²).

Разумеется, индуктивная логика не может подменить собой никакое научное исследование в конкретной области. Это означает, что оценка правдоподобия той или иной научной гипотезы как абсолютной величины в индуктивной логике невозможна. Для этого надо обладать таким знанием, которого ни у кого не может быть никогда, а именно, знанием вообще всех возможных гипотез относительно данного предмета.

Но зато индуктивная логика хорошо справляется со сравнением двух любых альтернативных гипотез относительно одного и того же предмета. Она не может сказать, какая из них правильная (может быть, неправильны обе), но она может привести здравые аргументы в пользу одной из них относительно другой, либо показать, что правдоподобие альтернативных гипотез более-менее одинаковое.

Чтобы к сравнению научных гипотез можно было применить логическую методику, нужно оценивать результаты либо экспериментов, либо наблюдений (что более привычно для гуманитарных наук) по таким параметрам, которые должны меняться в зависимости от принятой гипотезы. Чем сильнее выбор гипотезы влияет на прогнозируемое значение этих параметров и чем больше набрано самих наблюдаемых результатов, тем более достоверным окажется сравнение.

Дальнейшая идея проста и естественна: правдоподобнее будет та из гипотез, в которой наблюдаемые значения зависимых от нее параметров чаще совпадают с прогнозируемыми. На основании этих идей можно построить некоторую вычислительную модель. Основой для нее служит теорема Байеса, которая связывает между собой две разных вероятности (так называемые «условные вероятности», т.е. вероятности одного события при условии наступления другого события): события *A* при условии состоявшегося события *B* и события *B* при условии состоявшегося события *A*. Если события *A* и *B* совершенно никак друг с другом не связаны, то вероятности каждого из них никак не будут зависеть от того, состоялось ли другое событие или нет. Но если связаны, то будут зависеть друг от друга

² В качестве общего введения можно особо рекомендовать *Hawthorne J. Inductive Logic // The Stanford Encyclopedia of Philosophy (Winter 2014 Edition) / E.N. Zalta (ed.) [Электронный ресурс] // <http://plato.stanford.edu/archives/win2014/entries/logic-inductive/>*. В качестве введения в соответствующий логико-философский подход см. особо последнюю книгу Ричарда Джеффри (1926–2002), завершённую за пять дней до смерти: *Jeffrey R. Subjective Probability: The Real Thing. Cambridge: Cambridge UP, 2004.*

весьма ощутимо, и эта зависимость как раз и описывается теоремой Байеса. Поскольку любая гипотеза содержит ряд взаимосвязанных утверждений, то правдоподобие гипотезы в целом будет определяться рядом зависимых друг от друга вероятностей.

Логика не налагает никакого ограничения на тот способ, которым мы будем измерять «количество» вероятности. В традиционной математической теории вероятностей для этого используются действительные числа в интервале от 0 до 1, но они почти никогда не подходят для описания процессов, происходящих за пределами нашего воображения. В жизни мы крайне редко имеем дело с абсолютно точно определенными количествами. В древности считалось, что абсолютно точные измерения возможны только в астрономии, а теперь (то есть уже давно) мы знаем, что и в астрономии они невозможны.

Все количества, которые мы измеряем в реальном мире, — это некоторые интервалы с размазанными краями. Но это не означает, что эти неточные количества нельзя измерять, и на основании таких измерений выводить точные законы природы. Когда человечество (в лице Галилея) сделало такой вывод, произошла научно-техническая революция XVII в.³

С начала XX в. существует специальная теория статистической обработки данных, которая позволяет жестко отсеять лишь практически важную часть этого интервала (например, так называемый метод наименьших квадратов). Так поступают в естественных науках.

В гуманитарных областях существует та же самая проблема нечеткости («размазанности») количеств, но при этом нечеткость настолько велика, что обычные статистические процедуры не смогут дать удовлетворительного приближения. Эти обычные процедуры основаны на том, чтобы заменить нечеткое значение четко определенным интервалом, за пределами которого находится тот разброс наблюдаемых величин, которым можно пренебречь. В гуманитарных областях заменить нечеткую величину четким интервалом и просто отбросить оставшуюся нечеткость не получится. Нужно работать с нечеткими количествами такими, каковы они суть, — то есть не пытаться уйти от их нечеткости.

Если наши количественные оценки не могут быть выражены иначе, нежели в таких понятиях, как «много», «мало», «далеко», «близко», «часто», «редко», то для такого рода количественных показателей нужно использовать особую математику, основанную на так называемой нечеткой логике и теории нечетких множеств. Ее мы и постараемся применить в предлагаемой ниже методике.

3. Две гипотезы относительно «Солунской легенды»

Переходя к описанию примера, который мы будем рассматривать, можно сделать предупреждение о том, что для описания методики безразлично, реальный он или выдуманный, но автору (то есть мне) это небезразлично. Впрочем, здесь я буду использовать только конечные результаты моего содержательного анализа различных гипотез о происхождении «Солунской легенды» (далее СЛ)⁴. Те, кто считает, что я неправильно интерпретирую связь различных гипотез с наблюдаемыми данными, не согласятся с моими

³ См. особ. *Koyré A. Études galiléennes* (Histoire de la pensée. № 15). Paris: Hermann, 1966.

⁴ *Lourié B. The Slavonic Solunskaja Legenda* ("The Thessalonican Legend") and Its Syriac Original // *Bible and Syriac Studies in Context* / C.B. Horn (ed.) (Eastern Mediterranean Texts and Contexts). Warwick, RI: Abelian Academic. (forthcoming). Здесь же см. и подробную библиографию.

выводами относительно СЛ, но это никак не должно сказаться на их оценке предлагаемой методики (просто для них она будет показана как бы на воображаемом примере).

Проблема языка оригинала СЛ была поставлена еще в XIX в. в связи с крайней невразумительностью ее текста в ключевой сцене. Текст этот следующий:

S свытъкъ съчици скокине съ҃҃ле свезане

k сборокъ с'чице со куки и жголѣ свр'зана

T зборкъ съчици скокине соупоуль свезаноу

x зборкъ съчицискокине соугоуль звезаноу

p зборкъ съ чици с кокине с оугоуль свезану

N зборакъ съ числи бльгарскими

Сиглы средневековых (XV—XVI вв.) рукописей даны заглавными буквами, сиглы списков XIX в. с утраченных средневековых рукописей — строчными буквами. Рукописи распадаются на два семейства, из них ближе к общему архетипу всех рукописей семейство **SNk**, дальше — **Txp**. Рукопись **T** древнейшая (конец XV в.), рукопись **S** (XVI в.) — наиболее исправная с точки зрения близости к общему архетипу.

Редактор **N** переписал темное место по-своему, но не очень впопад. Список **x** приведен согласно всем его изданиям, первые два из которых (1856 и 1859 гг.) осуществлены самим открывателем этого списка И. Хаджиконстантиновым-Джинотом. При этом, однако, сохранился рукописный список Хаджиконстантинова, с которого текст был напечатан в 1856 г., и в этом списке вместо «соугоуль» ясно читается «соупоуль». Видимо, Хаджиконстантинов колебался в прочтении своей (нам уже недоступной) средневековой рукописи, где буквы «г» и «п» писались похоже. Минимально отличающийся от **x** список **p** (изданный в 1877 г. Порфирием Успенским, но вскоре забытый, пока о нем не пришлось напомнить мне) подтверждает чтение «соугоуль» (на котором, напомним, остановился и Хаджиконстантинов).

Последнее слово в этом фрагменте (во всех рукописях, кроме **N**) означает нечто связанное. Первое слово — также очевидно славянское, но с определением его значения возникают трудности. Между первым и последним словами заключены три сегмента, то есть, очевидно, три значимых слова с предлогами или без, которые и составляют главную загадку. По контексту можно сказать, что речь идет о каких-то предметах, которые на Кирилла, главного героя СЛ, сбросил ворон (ворон в **SNk**, но голубь в **Txp**). Эти предметы упали к нему «на крыло», то есть куда-то в область плечевого пояса, он переложил их себе за пазуху, а там они вошли в его тело и исчезли, но после этого Кирилл забыл греческий язык и стал знать славянский.

Вся эта сцена имеет место в следующем контексте (сейчас мы позволим себе некоторый объяснительный пересказ части СЛ, учитывающий подробности и контекст, не отмеченные прежними исследователями, но и не зависящие от гипотезы относительно языка оригинала).

Удрученный невозможностью, как ему кажется, выучить славянский язык, Кирилл в воскресный день идет в Фессалонике в церковь. Уже на выходе из церкви он видит

каркающего ворона, который и сбрасывает на него эти предметы. Кирилл подсчитывает их количество (то ли 32, то ли 35; разногласия в списках), после чего впоследствии создает азбуку для славян с точно таким же количеством букв (средневековые славянские предания о количестве букв в алфавите знают оба числа, и 35, и 32; количество букв алфавита — это всегда некое символическое число, мало связанное с реальной орфографией).

Легенды о вороне (но не голубе!), сидящем на церкви и наделенном сверхъестественными способностями, засвидетельствованы (правда, только арабомусульманскими авторами) для разных регионов христианского мира, включая окрестности Константинополя (хотя для Фессалоники таких легенд мы не знаем). Связь изучения алфавита для священных писаний с воскресной литургией хорошо известна из разных полуофициальных богослужебных чинопоследований, встраивающих в контекст воскресной Евхаристии (или еврейского праздника Пятидесятницы, дарования Завета Моисею — Шавуот) символическое поедание букв. В СЛ, однако, этот образ усиливается, т.к. речь идет не о поедании букв, а о прямом их доступе к сердцу.

Тот факт, что тут имеется в виду попадание предметов, взаимно-однозначно связанных с буквами, именно в сердце, видно как из стандартных для христианских и иудейских традиций молитв о просвещении именно сердца для уразумения священной азбуки, так и из более широкого контекста нашего эпизода внутри самой СЛ. В более широком контексте СЛ эксплицитно представляет своего Кирилла мессианской фигурой наподобие нового Моисея, который дает новому (и эсхатологически последнему верному) народу новый Завет. В самом начале СЛ говорит, что Кирилл должен будет дать славянам «Закон». В конце она говорит, что Кирилл учил славян лишь немного, но они сами понимали очень многое. Это явная реминисценция пророчества Иеремии (31, 34) о новом Завете: «И не научит кийждо ближняго своего и кийждо брата своего, глаголя: познай Господа. Яко вси познают Мя от мала даже и до великаго их». Но новый Завет, согласно Иер. 31, 33 и 2 Кор. 3, 3, пишется на сердце, а не на скрижалях каменных.

По сделанным в 2003 г. подсчетам В.М. Загребина, существовало не менее 11 способов реконструкции темной фразы СЛ, а он сам предложил еще один, 12-й. Среди прежних 11 способов 8 (а также девятый загребинский) исходили из возможности прочтения темного места на славянском языке (возможно, с примесью греческого), 2 — на еврейском языке, 1 (предложенный мной в докладе 1994 и статье 1996 г.) — на сирийском языке. В последней работе о СЛ я предлагаю несколько иной способ прочтения этого места на сирийском языке; это уже, как минимум, тринадцатая попытка реконструкции. Характерно, что нет попыток реконструкции темного места на основе греческого языка. По ряду причин, на которых здесь можно не останавливаться, гипотеза о наличии у СЛ греческого оригинала не имеет сторонников. Обе гипотезы о наличии оригинала еврейского также давно уже не имеют сторонников, и мы их рассматривать не будем.

Идея о том, что СЛ изначально была написана на славянском, а темное место в ней возникло вследствие некоей порчи славянского текста, в настоящее время представлена в виде двух групп гипотез: «старых» восьми гипотез и «новой» одной В.М. Загребина. Различие между «старыми» и «новой» методологическое: «старые» просто подбирали похожие слова и не делали особых предположений относительно механизмов их искажения до того вида, который они получили в рукописях СЛ; В.М. Загребин подошел противоположным образом, подобрав в качестве объяснений ко всем странностям славянского текста те или иные явления в тех или иных южнославянских языках.

Гипотеза В.М. Загребина оказалась слишком жестко собранной и поэтому разбивающейся при столкновении с реальностью. Для достижения идеала «полного

объяснения» В.М. Загребину пришлось ввести ряд крайне маловероятных предположений (так, для двух из трех ключевых слов он реконструирует такие формы, которые, по его признанию, не зафиксированы ни в одном из славянских языков, хотя и не противоречат морфологии некоторых из них) и при этом выбрать чтения тех рукописей, которые, по мнению исследователей текстологии СЛ, отстоят дальше от общего архетипа дошедших списков и поэтому являются *a priori* менее достоверными. Многие из этих предположений независимы от остальных, чем дополнительно снижается правдоподобность всей конструкции. Поэтому за научную строгость гипотезы (в смысле отсутствия в ней необъясненных мест) пришлось заплатить ее правдоподобием: чем больше мелких деталей в твоей гипотезе, тем хуже она работает, то есть меньше похожа на правду. Если же из гипотезы Загребина убрать все мелкие детали, все эти маловероятные допущения относительно существования именно в СЛ нигде более не зафиксированных славянских слов или особых значений известных славянских слов, то в ней останутся, в основном, лишь такие элементы, которые встречались уже в «старых» гипотезах о славянском происхождении текста.

Конечно, именно славянские гипотезы являются *a priori* самыми правдоподобными, но только не в варианте Загребина. Чтобы вообще отказаться от славянского подхода, нужны очень веские аргументы. Поэтому сравнение предложенного мною сирийского подхода со славянским будет иметь в виду не гипотезу В.М. Загребина (слишком легкого спарринг-партнера для любой иной гипотезы), а некоторую обобщенную и идеализированную «славянскую гипотезу». Это будет такой умственный конструкт, который заведомо правдоподобнее всякой реальной славянской гипотезы, не только существующей, но и даже возможной теоретически⁵. В рамках этой нашей идеальной «славянской гипотезы» мы в каждом случае выбираем самое лучшее из возможных «славянских» объяснений каждого отдельного факта, не утруждая себя необходимостью связать результаты таких отдельных выборов между собой (в любой реальной гипотезе это сделать необходимо, и это всегда вызывает значительные затруднения). Таким образом, мы будем оценивать потенциал славянского подхода по такой методике, по которой он будет заведомо сильно завышен. Что касается сирийского подхода, то мы будем ориентироваться на его новую версию, предложенную мною в статье 2015 года.

4. Области сопоставления гипотез

Гипотезы сопоставляются на основании значений тех параметров, которые от этих гипотез должны зависеть. Все прочие факты, которые объясняются при учете обеих гипотез одинаково, во внимание не принимаются.

Вопрос о языке оригинала какого-либо документа никогда не может быть окончательно решен на основании только текстологических и лингвистических данных, так как необходимо еще учитывать культурный контекст. В случае документа агиографического, которым является СЛ, этот культурный контекст — прежде всего, контекст других агиографических традиций, а, во вторую очередь, контекст общеисторический. Но для того или иного вывода о языке оригинала хорошая гипотеза, основанная только на данных

⁵ Его максимальная апостериорная правдоподобность, то есть максимальное соответствие результатам реальных наблюдений, будет, разумеется, достигаться ценой его абсолютной нереалистичности: в этом конструкте будут разрешены любые сочетания по смыслу несочетаемых слов, т.е. в целом он будет бессмысленным. Но мы сделаем вид, будто не замечаем его бессмысленности.

текстологии и лингвистики, является необходимой, хотя и не достаточной предпосылкой. Этим и определяется важность обсуждаемого сейчас вопроса — сравнения двух подходов, но лишь в пределах текстологии и лингвистики, игнорируя все остальные контексты.

Оставаясь в рамках текстологии и лингвистики, мы можем только предложить некую реконструкцию для темного места в СЛ и затем посмотреть, насколько она будет гармонировать с текстом легенды как целым. За пределы текста СЛ мы выходить не будем, хотя для полноценной интерпретации СЛ это было бы необходимо.

В рамках лингвистики и текстологии мы можем выделить четыре следующих области для сравнения:

1. Фонологическо-лексикологическая: реконструкция слов, изначально бывших в славянском тексте темного места (эти слова, хотя и оказались в славянском тексте, не обязательно должны быть славянскими; так, в некоторых из «старых» гипотез одно из трех слов могло быть греческим, а в нашем сирийском подходе — все три сирийские);
2. Лексикологическая: реконструкция изначально подразумевавшихся значений этих слов (подобрать слово бывает легче, чем выбрать его значение!);
3. Контекстуальная: уместность реконструкции темного места в рамках СЛ как целого;
4. Текстологическая: определение причины «темноты» темного места.

Сразу скажем, что мы можем предварительно упростить себе задачу, сократив сопоставление двух наших подходов во второй области (лексикологической). Здесь оба подхода имеют, приблизительно, одинаковую эвристическую силу, сталкиваясь с однотипными трудностями. — Но иначе и быть не могло, так как место это недаром является темным. Речь идет о том, что, независимо от конкретной гипотезы, мы будем иметь дело с какими-то необычными значениями слов, хотя сами слова могут быть (а могут не быть) обычными.

Главным камнем преткновения во всех гипотезах является первый из трех «темных» сегментов — същици. В рамках славянского подхода в отношении этого сегмента имеется два типа гипотез: либо это минимально искаженное в поздней (но еще средневековой) южнославянской орфографии слово «сучки», либо это очень сильно искаженное слово со значением «дощечки» («дщицы»). В первом случае остается не очень понятным, как и почему эти сучки соотносятся с буквами, но зато легче понять, почему они входят в тело Кирилла, будучи взятыми за пазуху: пусть они и не заостренные специально, но все-таки, некоторым образом, колющие предметы. Во втором случае, помимо того, что следующий сюжетный ход выглядит совершенно немотивированным (плоские дощечки не очень легко ассоциировать с чем-то проникающим в тело), возникает проблема домысливания связи дощечек с буквами. Разумеется, дощечки сразу представляются аналогами скрижалей Моисея, но это довольно существенное исправление текста средствами воображения: текст, если понимать его буквально, соотносит буквы алфавита с самим количеством предметов, которые Кирилл пересчитывает сразу по получению. Нельзя сказать без натяжки, что Кирилл пересчитывает не сами предметы, а какие-то буквы на них. Все эти трудности обеих гипотез примиряются в нашей идеализированной славянской гипотезе: предположим, она имеет в виду некие предметы, которые образуют счетное множество со взаимно-однозначным соответствием элементам множества букв алфавита, но при этом название этих предметов передается в первом «темном» сегменте почти без искажения, а контекстуально их образ встраивается в символический ряд «нового Завета» не хуже (но и не лучше) гипотетических

дощечек-скрижалей (если Завет новый, то наличие для него каких-то отдельных от «скрижалей сердца плотяных» скрижалей несколько сомнительно).

Сирийская гипотеза (в ее нынешнем варианте) состоит в следующем. Сброшенные предметы суть связки (одна или несколько) птичьих перьев для письма. Традиционно самыми лучшими перьями такого рода считались перья ворона или лебедя (а перья голубя — непригодными вообще вследствие их мягкости). Второй и третий «темные» сегменты читаются как записанные славянскими буквами почти без искажения два сирийских слова, означающих «связку» («сгуло», мн. ч. «сгуле»), перевязанную ремешками (обычно кожаными) — «скокине» (< *съкокинъ / *сакокинъ). Первый сегмент интерпретируется как минимально искаженная передача слова «цуцийот» (с заменой сирийского окончания множественного числа *-йот* славянским). Все три слова образуют одно словосочетание, которое славянский переводчик, согласно гипотезе, не смог понять и просто транслитерировал (возможно, понимая его грамматику и постаравшись ее передать через множественное число для первого слова).

Тут проблема в том, что слово «цуцийот» нигде не фиксируется в значении «птичье перо для письма». Но эта проблема не так велика, как может показаться на фоне латыни или славянских языков, так как в сирийском не существовало специального термина для такого пера (в отличие от лат. *penna*). В отношении к птичьим перьям для письма сирийская традиция занимала промежуточное положение между византийской (где такие перья не применялись вообще, и специального термина для них не существовало вовсе; для письма использовались только тростниковые перья) и латинской (где такие перья применялись издавна, а особенно популярными стали с начала VI в.). В сирийской традиции они применялись с начала VI в. включительно до X или даже XI вв., но потом вновь вышли из употребления. Что касается славянской традиции, то относительно ее самых ранних этапов ситуация неясна, а русская с древнейших времен использует только птичьи перья, совершенно отказавшись от тростниковых; этот факт можно считать доказательством раннего использования птичьих перьев также и у южных славян, от которых на Руси были заимствованы традиции письменности.

В сирийской традиции для обозначения пера не тростникового, а птичьего нужно было использовать то или иное сочетание из двух слов. По контрасту с этим, слово «трость» (тростник) означало сразу и растение, и тростниковое перо. Слово «цуцийот», обычно обозначающее различные продолговатые и при этом мягкие предметы, фиксируется в одном из сирийских текстов VI в. в качестве обозначения мягких тростниковых соцветий (метелок), внешне похожих на птичьи перья для письма. Конечно, соцветие-метелка — это не та часть тростникового стебля, которая годится для изготовления калама, но зато она и метафорически (по внешнему сходству), и метонимически (по семантической смежности — принадлежности к тростнику как материалу для изготовления пера для письма) подходит для обозначения писчего пера из пера птицы. Даже если для сирийского языка такое обозначение не было обычным (о чем мы, строго говоря, судить не можем, не имея достаточно репрезентативного массива текстов, т.к. птичьи перья для письма слишком давно вышли из употребления у сирийцев), оно могло легко возникнуть *ad hoc* в высоко-поэтическом образном строе легенды. Кроме того, оно хорошо согласуется и с последующим сюжетным ходом — вхождением (заостренных предметов, какими являются перья для письма) в тело Кирилла, а также с появлением птицы и особенно ворона (чьим перьям могли бы составить конкуренцию только перья лебедя) и с «упаковкой» сбрасываемых предметов именно в виде связки (оптовые количества птичьих перьев для письма также упаковывались в связки).

Но следует признать, что и наша сирийская гипотеза использует такое значение ключевого слова, которое нельзя вычитать из словаря. Если ее сравнивать только в этом отношении с гипотезой славянской в ее идеализированном виде, то силы окажутся приблизительно равны. Поэтому из дальнейшего анализа мы эту область (лексикологическую) исключаем.

Заодно это избавит нас от необходимости оценивать еще один довольно сложный параметр — степень пробабилистической зависимости между значениями слов, выбранных для каждого из трех сегментов: очевидно, что она велика, а ниже мы будем использовать формулы, простота которых подразумевает пробабилистическую независимость друг от друга результатов наших наблюдений. Для остальных областей наблюдения предположение о независимости наблюдаемых параметров будет вполне разумным.

Для дальнейшего формализованного сопоставления гипотез у нас остаются три области. Далее при описании методологии я буду рассчитывать на читателя, хотя бы смутно помнящего программу старших классов средней школы по математике, однако, даже те читатели, которые ее не помнят, смогут при желании следить за изложением.

5. Понятие правдоподобия предположения

Правдоподобие P (*likelihood*) определяется как особого рода вероятность, определяющая степень строгости логического следования, если она зависит от случайных событий. Если предположение h_1 существенно правдоподобнее, чем h_2 , то это означает, что его правдоподобие существенно выше, а именно:

$$P[e | h_1 \cdot b \cdot c] \gg P[e | h_2 \cdot b \cdot c],$$

где e — это результат наблюдения при условиях c , который должен, согласно гипотезе h_1 или h_2 , каким-то определенным образом зависеть от этой гипотезы. Это определение так называемой правдоподобности прямого вывода (*direct inference likelihood*), которая представляет собой полностью объективную величину.

Все это при условии одинакового «бэкграунда» b (то есть знаний о данной области, на которые выбор гипотезы не влияет), а также при условии, что интерпретации наблюдений ведутся по одинаковым правилам⁶.

Пользу и неочевидность такого подхода легко пояснить на конкретном примере.

Мы будем говорить о гипотезах, имеющих одинаковый «бэкграунд» b , то есть общие представления о текстологии и лингвистике, а также одинаково подходящих к интерпретации результатов⁷.

Рассмотрим такой пример, к которому можно сразу применить оценку численными методами, а именно, различие в одной согласной («г» или «п») в сегменте «соуголе». Для сирийского подхода жизненно важно настаивать на аутентичности «г». Для славянского — желательнее настаивать на аутентичности «п» (тогда «сопоуле» можно трактовать как искажение «с полу» или чего-то подобного, то есть «надвое» или «пополам»), но возможно смириться и с «г» (тогда речь пойдет о чем-то «с углов», что труднее вписать в общую

⁶ То есть одни и те же факты не будут трактоваться совсем по-разному сторонниками разных гипотез. Это условие можно нарушить, но тогда мы будем сопоставлять субъективные представления о правдоподобии, соответствующие разным школам, и нам придется оперировать не общими для всех значениями P , а значениями P_α , P_β и т.д.

⁷ Поэтому будем оперировать только P , а не P_α , P_β и т.д.

картину, и что предполагает еще более сильные искажения текста). Нашей задачей будет установить, какой вклад в общее целое вносит только одна эта составляющая — выбор в качестве аутентичного чтения либо «г», либо «п». Интуитивно может показаться, что, коль скоро «п» в одном списке, а «г» в четырех, то правдоподобности соответствующих предположений об аутентичности каждого из чтений должны отличаться в четыре раза. Но на самом деле это не просто не так, а далеко не так. Насколько же именно далеко?

Для оценки правдоподобия гипотезы об аутентичном чтении мы обращаемся к пяти спискам: это пять наблюдений при пяти условиях $c_1 \dots c_5$ (тут «условие» = «отдельный список»), при которых мы получаем пять результатов наблюдений (outcomes) $e_1 \dots e_5$. Из этих результатов четыре «г» и один «п». Если бы результаты имели четкое распределение по семействам рукописей, мы не могли бы их рассматривать как независимые. Следовало бы учитывать возможность того, что ошибка принадлежит общему протографу одного из семейств, которому могло больше повезти в выживании. Таким образом, нужно было бы учитывать влияние принятой нами гипотезы на сами условия наблюдения⁸. Но в нашем случае этого не требуется, т.к. четыре рукописи, содержащие «г», распределены по обеим редакциям СЛ. Следовательно, не должно быть большой ошибки в том, чтобы считать результаты наблюдений статистически независимыми (и, следовательно, не учитывать влияние принимаемой гипотезы на сами условия наблюдения).

Независимо от того, которое из этих чтений аутентично, в рукописной традиции существует определенная вероятность искажений. Скажем, для традиции небрежного переписывания не особенно важных текстов ошибку можно принять вероятной для каждого десятого списка. Для нашего *lectio difficilior* вероятность ошибки при переписывании довольно мала (может быть, 0,05%, например), но предположим ее неправдоподобно большой: пусть ошибка считается допустимой для каждой пятой рукописи (= 20%)⁹. Тогда гипотеза об аутентичности либо «г», либо «п», должна теоретически предполагать, что нужное чтение встречается в 80% рукописей. В наших реальных пяти рукописях мы видим иное, и это влияет на оценку правдоподобия гипотез. Понятно, что гипотеза с «п» видится существенно менее вероятной, если смотреть только на разночтения рукописей, но нам важно оценить, насколько именно. Чем больше мы допускаем случайности для этих разночтений, тем меньше должны различаться правдоподобности гипотез, но мы увидим, что даже при огромном (и выходящем за пределы естественного) допуске 20%-ной случайной ошибки различия правдоподобности весьма велики.

Правдоподобие отдельной гипотезы тогда вычисляется по формуле (1) (которую желающие могут пропустить, т.к. нам будет нужна не она сама, а следующая формула,

⁸ Формально это можно было бы сделать любым образом, при котором ожидаемость (вероятность появления) данного разночтения будет определяться как условная вероятность, зависящая от вероятности появления этого разночтения в соответствующем субархетипе. Последнюю (независимую) вероятность нужно было бы как то оценить, потом в рамках данного семейства оценить условную вероятность данного разночтения, а итоговая вероятность (ожидаемость) данного разночтения была бы равна произведению его условной вероятности на вероятность этого разночтения в архетипе. В таком случае, $P(a) = P(a | A) \times P(A)$, где $P(A)$ — вероятность появления данного разночтения в субархетипе (пропорциональна числу субархетипов), $P(a | A)$ — вероятность данного разночтения в рамках субархетипа, а $P(a)$ — итоговая вероятность данного разночтения.

⁹ Назначение этих вероятностей происходит по процедуре, описанной ниже (раздел 7) для терм-множества значений величины m_i . На данном этапе читатель может удовлетвориться интуитивным пониманием этой процедуры. Говоря более строго, значение p или q (при условии, что гипотез может быть только две) должно сильно отличаться от 0,5 (в любую сторону), т.к. оно показывает меру неслучайности наблюдаемых значений выбранных параметров, тогда как значение 0,5 соответствует чистой случайности их обоих.

полученная на ее основе, — формула для сопоставления двух гипотез (2); см. ниже):

$$(1) \quad P[e | h_r \cdot b \cdot c] = \frac{n!}{m! \times (n-m)!} \times r^m (1-r)^{n-m}$$

где (для нашего примера) $r = 0,8$ (наше предположение о проценте правильных, согласно выбранной гипотезе, чтений от общего числа), n — количество рукописей, содержащих интересующее нас чтение в том или ином его виде, m — количество рукописей, имеющих правильное чтение согласно данной гипотезе (знак ! — знак факториала, то есть произведения всех целых натуральных чисел от 1 до указанного включительно; при этом $0! = 1 = 1!$)¹⁰.

В этой формуле не учитывается априорная вероятность (*prior probability*) данного предположения, которая, в принципе, может серьезно влиять на правдоподобие предположения в целом. В нашем примере мы вправе предположить, что априорные (без учета той или иной гипотезы) вероятности появления «п» или «г» одинаковы, а потому могут не учитываться при сопоставлении гипотез. Поэтому в формуле выше приводится только расчет апостериорной составляющей (*posterior probability*) правдоподобности гипотезы.

Расчет этой величины для нас, впрочем, бесполезен, и мы не будем его проводить, т.к. всё, что касается полезности гипотез, познается логически только в их сравнении. А для сравнения апостериорных правдоподобностей двух гипотез мы получаем тогда такую формулу, определяющую уже не сами правдоподобности как абсолютные величины, а их соотношение:

$$(2) \quad \frac{P[e^n | h_q \cdot b \cdot c^n]}{P[e^n | h_r \cdot b \cdot c^n]} = \frac{q^m (1-q)^{n-m}}{r^m (1-r)^{n-m}}$$

где, предположим, r — это ожидаемое количество «п» в рамках гипотезы h_r (т.е. «п» аутентичное чтение, его ожидаемая вероятность 80%, т.е. r полагаем равным 0,8), q — ожидаемое количество *все того же чтения* «п», но теперь уже в рамках гипотезы h_q об аутентичности чтения «г»; положим для простоты, что $q = 0,2$, что соответствовало бы возможности ошибки только между «п» и «г», т.е. похожими кириллическими буквами); n — общее число сопоставляемых рукописей (оно равно 5), m — число таких рукописей, в которых реально наблюдается «п» (оно равно 1).

Остановимся подробнее на величинах r и q . Возможность использовать биномиальное распределение вероятностей ограничивает нас оценкой только таких событий, в которых возможны лишь два исхода (аналогично результату бросания монеты). Фактически, мы оцениваем правдоподобие двух гипотез: будто «п» выпадает с ожидаемостью 80% (h_r) и будто «п» выпадает с ожидаемостью 20% (h_q). По условиям нашей задачи, последняя гипотеза эквивалентна предположению, что «п» выпадает лишь в пределах погрешности рукописной традиции, тогда как аутентичным чтением является некое «не-п», и оно уже имеет 80% ожидаемость. Так как мы условно приняли, что других ошибок, кроме путаницы между «п» и «г», быть не может, то «не-п» у нас означает «г», то есть $r + q$ мы принимаем равным 1. Предположение о невозможности других ошибок, строго говоря, является

¹⁰ Это так называемая биномиальная формула для расчета вероятностей. Она применима только к таким ситуациям, когда вариантов исхода для наблюдения или эксперимента только два. Первый член справа, так называемый биномиальный коэффициент, выражает число сочетаний по m из общего числа элементов n .

неверным, но его целесообразно принять, поскольку вероятности подобных ошибок намного (на порядки) меньше, чем ошибка вследствие путаницы похожих букв «п» и «г».

Получаем для соотношения правдоподобия двух гипотез: $0,2^1 (1-0,2)^{5-1} / 0,8^1 (1-0,8)^{5-1} = 64$. Предположение об аутентичности чтения «г» правдоподобнее не в четыре, как могло бы показаться, а в 64 раза (4^3). Ровно такой же результат мы получим, если слегка изменим условия, а именно учтем чтение рукописи Константинова-Джинота (с «п»), но при этом будем учитывать его же публикацию с «г» как отдельный список. Тогда у нас $n = 6$, $m = 2$, а соотношение апостериорных правдоподобностей не меняется.

Еще раз изменим условия, чтобы оценить влияние предположения о тщательности переписчиков. Предположим, у нас $n = 5$, $m = 1$, но теперь $r = 0,9$, $q = 0,1$; это соответствует всего лишь не особенно небрежному переписыванию. Тогда соотношение правдоподобия вырастет до $9^3 = 729$ раз. А если сделать поправку на тщательность переписывания *lectio difficilior* — хотя бы до $r = 0,95$, $q = 0,05$, — то оно достигнет величины $19^3 = 6859$ раз.

Из разобранного примера мы можем сделать вывод лишь о том, что любые гипотезы, опирающиеся на чтение «соупоуле» вместо «соугоуле», сильно вредят своей правдоподобности. Но, возможно, они это делают так, чтобы где-то в другом отношении что-то выигрывать так сильно, что эта потеря компенсируется. Наша идеализированная славянская гипотеза теряет на «соупоуле» довольно мало, т.к., в отличие от любой реальной гипотезы, она имеет возможность следовать за наиболее правдоподобными с текстологической точки зрения решениями (но все-таки, как мы увидим, что-то теряет все равно).

Остановимся еще раз на величинах r и q . Предположим теперь, что мы обязаны учитывать не два варианта ошибки, а неопределенно много, то есть сумма r и q (ожидаемостей появления «п» и «г» соответственно гипотезам об аутентичности каждого из этих чтений) будет существенно меньше 1, т.к. теперь мы не сможем принять, что «г» это просто «не-п». Тогда нам будет очень трудно найти величины r и q , то есть правдоподобность появления чтения «п» или «г» в зависимости от гипотезы. Степень аккуратности переписывания (которую мы использовали выше, оценив ее в 0,8, а потом в 0,9 и 0,95), сама по себе, не зависит от наших гипотез относительно аутентичного чтения, и поэтому никак не скажется на сравнении их правдоподобия, если «г» перестанет быть «не-п». Если мы не найдем какого-то другого способа предсказывать вероятность появления того или иного чтения в зависимости от гипотезы (а сделать это нам будет трудно), то тогда весь метод неприменим.

Итак, наш метод довольно легко применим только тогда, когда мы можем свести все разнообразие возможных гипотез только к одной паре, то есть принять в качестве разумного приближения $r + q = 1$. В других случаях его применить труднее, т.к. параметры r и q надо будет задать каким-либо специальным образом.

На практике принятие условия $r + q = 1$ означает не то, что все остальные гипотезы абсолютно исключены, а то, что — на основании каких-либо априорных, но разумных соображений — мы считаем, что их правдоподобностями можно пренебречь по сравнению с двумя сопоставляемыми величинами (и, разумеется, если апостериорно мы и на самом деле не встретили ничего, кроме двух оцениваемых вариантов, «п» и «г»). Зато, если мы можем упростить изучаемую ситуацию до эквивалентности между r и q ($r = 1 - q$), то фактически мы получаем некоторую эмпирическую константу (за нее можно принять либо r , либо q), которая является характеристикой когерентности текста (в наших примерах это 0,8, 0,9 и 0,95) либо, напротив, некогерентности (в наших примерах 0,2, 0,1 и 0,05). При рассмотрении случайных ошибок при переписывании мы имеем частный случай когерентности —

аккуратность переписывания. Но критерий когерентности, как мы увидим, можно будет применять на совершенно всех уровнях построения текста.

Когерентность текста оценивается независимо от сопоставляемых гипотез, но чем выше когерентность, тем сильнее расходятся апостериорные правдоподобности для одного и того же числа наблюдений n и одного и того же числа положительных исходов наблюдений (т.е. исходов в пользу одной из пары гипотез) m .

Весь разобранный только что пример был нужен для того, чтобы показать в общих чертах и без обращения к специальным методам нечетких логик, как устроены методы сопоставления альтернативных гипотез в зависимости от наблюдаемых результатов.

6. Сопоставление славянской и сирийской гипотез: общие соображения относительно метода

Обсуждая гипотезы в связи с конкретными наблюдениями над СЛ, мы, в принципе, делим все факты на три категории: безразличные для данной гипотезы, благоприятные для нее и неблагоприятные. Правдоподобности всех гипотез, кроме славянской и сирийской, мы принимаем практически равными нулю, исходя из априорных соображений.

Мы по-прежнему можем считать, что все наблюдаемые факты, если они как-то зависят от одной из сопоставляемых гипотез, делятся на благоприятные для, допустим, славянской гипотезы и неблагоприятные для нее. Относительно сирийской гипотезы те же факты будут иметь обратное значение (мы опять принимаем, что $r + q = 1$).

За нами остается выбор константы когерентности текста (либо r , либо q , поскольку эти величины у нас эквивалентны), которая теперь уже будет учитывать как его сохранность в рукописной традиции, так и его внутреннюю связность, включительно до, например, взаимосвязанности сюжетных ходов. Эта константа оценивается независимо от гипотезы относительно языка оригинала¹¹.

Традиционно считалось, что у СЛ она довольно низкая (текст, якобы, составлен наивно и неряшливо и не менее неряшливо переписывался), то есть, допустим, 0,8. После исследований СЛ в контексте славянской апокалиптики (особенно Анисавы Милтеновой) стало ясно, что текст отнюдь не наивен и не прост. Позже и я постарался показать, что когерентность текста СЛ, напротив, крайне высокая: насколько можно судить без обращения к той или иной гипотезе относительно языка оригинала, текст весьма целен концептуально, последователен и хорошо сохранился. Его низкая оценка ранними исследователями происходила из того, что они не понимали в должной мере его библейско-паралитургического символизма именно нового Завета по Иеремии, а не просто Завета

¹¹ Когерентность текста удобно рассматривать как некую обобщенную величину тогда, когда его бытование в рукописной традиции примерно соответствует его изначальной целевой аудитории, т.е. когда можно предположить, что переписчики относились к своему делу с той же тщательностью или небрежностью, что и авторы и редакторы. Для СЛ это выглядит правдоподобным утверждением, но читатель может сам пересчитать мои оценки, меняя константу когерентности для разных групп наблюдений. Классический пример текста, где подобное допущение заведомо недопустимо, — канонические евангелия: в дошедших до нас рукописях они переписывались тщательно, но сами дошедшие редакции составлялись иногда довольно небрежно (очень много взаимоисключающих утверждений, например, в географических подробностях...). Несмотря на большое сходство общей для всего произведения и его рукописной традиции константы когерентности с известным параметром «средней температуры по больнице», для маленьких текстов, переписывавшихся с более-менее одинаковой аккуратностью, этот критерий может быть удобен. В других случаях нужно назначать разные константы когерентности для разных групп эмпирических данных.

Моисеева, а поэтому смотрели на многие подробности сюжета как на «лишние детали», т.е. элементы информационной энтропии. Если же прочитать символический ряд сюжета, то все детали оказываются осмысленными и на месте, и ни одна из них не является необязательным элементом художественного украшения¹².

Проблемы возникают из-за того, что наблюдаемые факты бывают благоприятными или неблагоприятными в разной степени, и когда нужно учесть одновременно вклад множества фактов, то мы не можем просто так взять и воспользоваться приведенной выше формулой биномиального распределения вероятностей.

«Условиями» c_i для каждого наблюдения у нас будет наличие такого параметра текста, значение которого сильно зависит от выбранной гипотезы (те параметры текста, значения которых от выбора гипотезы не зависят, во внимание не принимаются). Количество таких набранных «точек для наблюдения» по-прежнему образует некое целое число n . Чем больше таких точек мы наберем, тем надежнее будет наше сравнение.

Но зато число m , т.е. число положительных, с точки зрения одной из сопоставляемых гипотез, исходов наблюдения, у нас перестанет быть натуральным числом и вообще перестанет быть числом в обычном математическом смысле. В отличие от ситуации с бросанием монеты или даже с выбором разночтений рукописей между «п» и «г», мы не сможем использовать четкие «да» и «нет» при ответе на вопрос, в пользу какой гипотезы свидетельствует данный исход данного наблюдения. Даже наше предельно сильное «да» не будет иметь абсолютного, 100%-го характера, т.к. оно не будет исключать полностью логическую возможность «нет».

7. Модификация формального аппарата на основе нечеткой логики

Тут и приходит на помощь нечеткая логика. Вместо обычных числовых переменных она предоставляет нам то, что ее создатель, Лотфи Заде, назвал «лингвистической переменной» — такую переменную, различные значения которой выражены в терминах реального человеческого восприятия (типа «много», «мало» и т.п.), а не в числах¹³. При желании, однако, эти величины можно аппроксимировать действительными числами. В частности, если у нас речь идет о долях («большая часть», «меньшая часть» и т.д.), то это будут числа в интервале от 0 до 1 (где твердое и четкое «нет» принимается за 0, такое же «да» за 1, а все остальные, то есть реально встречающиеся значения, занимают весь остальной континуум).

¹² См. подробно мою статью 2015 года, где совсем за пределами комментария остаются только две важных подробности: упоминание Кипра и Крита как промежуточных остановок на пути Кирилла из Александрии в Фессалонику. География агиографических легенд неспособна упоминать неважные для нее места, поэтому из критико-агиографических соображений и эти элементы текста нельзя считать случайными. Однако должен признать, что значение их в географии СЛ мне пока неясно.

¹³ Лучшим введением в основы нечеткой логики остается небольшая монография Лотфи Заде: *Zadeh L.A. The Concept of a Linguistic Variable and Its Application to Approximate Reasoning, I // Information Sciences. 1975. Vol. 8(3). P. 199—249; idem. The Concept of a Linguistic Variable and Its Application to Approximate Reasoning, II // Information Sciences. 1975. Vol. 8(4). P. 301—357; idem. The Concept of a Linguistic Variable and Its Application to Approximate Reasoning, III // Information Sciences. 1975. Vol. 9(1). P. 43—80.* Рус. пер.: *Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений / Пер. Н.И. Ринго. Математика. Новое в зарубежной науке. Вып. 3. М.: Мир, 1976.*

Оценка того, насколько результат данного наблюдения говорит в пользу одной из двух альтернативных гипотез, — это значение лингвистической переменной¹⁴.

Таким образом, параметр m в формуле биномиального распределения вероятностей (1) приобретает смысл лингвистической переменной и может быть, при желании, аппроксимирован суммой дробных чисел, каждое из которых находится в интервале от 0 до 1 (в обычной же теории вероятностей он представляет собой сумму единиц, то есть некоторое натуральное число).

В нашем случае m , став лингвистической переменной, будет образовываться как сумма не единиц, а лингвистических переменных m_i , каждая из которых может быть аппроксимирована дробным числом в интервале от 0 до 1.

$$(3) \quad m = \sum_{i=1}^n m_i$$

Нечеткие количественные представления основаны на бинарных оппозициях: например, «много» — это то, что не «мало»; «куча» (из античного парадокса кучи, не имеющего решения в классической логике) — это то, что уже не является счетным числом песчинок... Точность таких заведомо нечетких чисел зависит от подразумеваемой шкалы: например, имеется ли в виду только двучленная оппозиция «много—мало» или же более сложное деление, подразумевающее также «не очень много» и «не очень мало»... Такую подразумеваемую шкалу принято называть терм-множеством для значений лингвистической переменной; оно включает те значения, которое такая переменная может принимать внутри заданного для нее интервала (сам же этот интервал значений называется множеством-носителем).

Для практических целей часто бывает достаточно взять пять значений (образующихся при делении всего интервала возможных значений сначала на два, а потом каждой половины — еще раз на два): «(почти) всё», «большой частью», «средне / поровну», «меньшей частью», «(почти) ничего». Эта шкала грубо аппроксимируется численными значениями, соответственно: 1; 0,75; 0,50; 0,25; 0 (для последующих арифметических действий с этими величинами важно будет помнить, что «0» в данном случае — это не 0, а некая положительная величина, неуловимо — в условиях наших наблюдений — превышающая значение 0).

Чем хороша такая пятичленная шкала: она, с одной стороны, является достаточно общей, чтобы разные эксперты не сильно спорили по поводу отнесения тех или иных наблюдаемых значений лингвистической переменной к одному из ее термов, а, с другой стороны, она обеспечивает достаточную селективность для решения с ее помощью многих практических задач. В частности, такая шкала обеспечит и нам вполне ощутимые различия между правдоподобиями альтернативных гипотез.

Как только что было упомянуто, отнесение данного наблюдаемого значения к определенному терму терм-множества значений лингвистической переменной осуществляется на основе экспертной оценки. Учитывая собственную пристрастность в качестве эксперта в своем же собственном деле, я постараюсь сознательно завышать все результаты наблюдений в пользу славянской гипотезы (как отчасти уже видно из того, что

¹⁴ Предлагаемый ниже подход выдержан в общем духе работ Лотфи Заде по нечеткой теории вероятности и смежным проблемам. См., в частности, *Zadeh L.A. Toward a Perception-Based Theory of Probabilistic Reasoning with Imprecise Probabilities // Journal of Statistical Planning and Inference. 2002. Vol. 105. P. 233—264.*

саму славянскую гипотезу я рассматриваю в ее идеальном и практически неосуществимом виде). Но поскольку я буду обсуждать каждое наблюдение отдельно, желающие смогут подставить в те же формулы другие значения лингвистической переменной и получить свою собственную оценку.

Конкретно процедуру оценки значения m для индивидуального наблюдения лучше пояснить на примере. Напомню, что в «четких» вычислениях эта величина всегда равна сумме единиц и нулей, т.е. результат каждого данного наблюдения — либо «за», либо «против» выбранной гипотезы. В нашем случае все результаты будут лишь отчасти «за», а отчасти и «против». Это лучше пояснить на примере, относящемся к СЛ.

В тексте СЛ трижды (что много для небольшого текста) употреблена идиома «спускаться в» в значении «входить (а вовсе не спускаться) в населенный пункт». Это типичный семитизм. Можно сказать, что его наличие, да еще систематическое (трижды), идеально согласуется с гипотезой о семитском (т.е. сирийском) оригинале. Используя аппроксимацию термов дробными числами, мы можем записать $m_i^+ = 1$, но мы не можем записать просто $m_i = 1$, т.к. наличие подобных семитизмов возможно и в оригинальном славянском тексте, хотя вероятность их там мала; в случае же СЛ она даже и не так уж мала, поскольку текст построен на библейских аллюзиях и даже просто на библейском субстрате (новый Завет по Иеремии), и поэтому семитизмы как библицизмы там вполне ожидаемы. Можем (с некоторой щедростью в пользу славянской гипотезы) оценить $m_i^- = 0,5$. Получается, что искомое нами число m всегда складывается из двух компонент — степени благоприятствования данного факта для той гипотезы, по отношению к которой мы ведем сравнение (m_i^+), и степени благоприятствования данного факта для альтернативной гипотезы (m_i^-)¹⁵.

Каждая из двух величин, m_i^+ и m_i^- , может принимать любые значения в интервале от нечеткого «0» до нечеткой «1» и, вообще говоря, никак не зависит от другой. Их значения связывает только одно общее правило: они не могут быть похожими до неразличимости. Совпадение (в пределах нечеткости) для этих величин означает, что данное наблюдение одинаково в пользу той и другой гипотезы. Но это противоречит нашему начальному условию брать для рассмотрения только те наблюдения, в которых различие между гипотезами вполне уловимо. Кроме того, эти величины не могут быть одновременно равны нечеткому «0» еще и потому, что это противоречило бы нашему априорному условию, согласно которому наши две сопоставляемые гипотезы являются единственными гипотезами, о которых имеет смысл говорить. Если эмпирически мы наблюдаем для какого-

¹⁵ Как нетрудно заметить, введенное сейчас понятие «степени благоприятствования» соответствует тому, что обычно в нечеткой логике называется функцией принадлежности, определяемой для элементов нечеткого множества (в данном случае, множества наблюдений с положительным исходом). Функция принадлежности показывает, насколько данный элемент (данный результат наблюдения) принадлежит нечеткому множеству (множеству наблюдений с положительным исходом, которое у нас стало нечетким). Переход от нечеткой теории вероятностей к оценке байесовых правдоподобностей пропозиций осуществляется через отождествление последних с вероятностями того, что Л.А. Заде назвал «нечеткими событиями» (*fuzzy events*): Zadeh L.A. Probability Measures for Fuzzy Events // Journal of Mathematical Analysis and Applications. 1968. Vol. 23. P. 421—427. Идеология этого подхода сформулирована Джулианеллой Колетти и Романо Скоццафавой; см. особ.: Coletti G., Scozzafava R. Probabilistic Logic in a Coherent Setting (Trends in Logic. Vol. 15). Dordrecht: Kluwer, 2002; Scozzafava R. The Membership of a Fuzzy Set as Coherent Conditional Probability // On Fuzziness: A Homage to Lotfi A. Zadeh — Volume 2 / R. Seising et al. (eds.) (Studies in Fuzziness and Soft Computing. Vol. 299). Berlin—Heidelberg: Springer-Verl., 2013. P. 631—635.

то факта значение «0» сразу для m_i^+ и m_i^- , то это означает, что мы сделали ошибку в своем априорном суждении относительно заведомой неправдоподобности любой третьей гипотезы.

Уже из этого запрета на одинаковые величины для m_i^+ и m_i^- можно заметить, что эти значения лингвистической переменной ведут себя не так, как те дробные числа, которыми их можно аппроксимировать в некоторых случаях.

Принимая к рассмотрению только те значения m_i^+ и m_i^- , которые ясным образом отличимы друг от друга, мы сталкиваемся с необходимостью найти правила для вычисления по ним нечеткого числа m_i (которое также является значением лингвистической переменной, но уже полученной из сложения — правила которого мы должны будем установить — m_i^+ и m_i^-).

Эти правила можно установить из следующих соображений (подразумевающих для лингвистической переменной m_i такое терм-множество, которое устроено аналогично терм-множествам переменных m_i^+ и m_i^-).

Симметричная ситуация, когда результат наблюдения свидетельствует одинаково в пользу обеих гипотез, соответствовала бы $m_i = 0,5$, но это значение у нас запрещено, т.к. оно показывает, что наше наблюдение не обладает селективностью по отношению к сравниваемым гипотезам и поэтому не должно было приниматься во внимание вовсе. Значения m_i больше 0,5 означают, что результат наблюдения, скорее, в пользу первой гипотезы, а значения меньше 0,5 — что, скорее, в пользу второй. Безусловному перевесу первой гипотезы соответствует нечеткое значение $m_i = 1$, а такому же перевесу второй гипотезы — $m_i = 0$ (также нечеткому; т.е. этот «0» всегда чуть больше 0, а «1» всегда чуть меньше 1). Дальше естественно установить средние значения перевеса как условные (нечеткие), соответственно, 0,75 и 0,25, а для слабых и сильных значений перевеса выбрать, соответственно, 0,6 и 0,9 (в пользу первой гипотезы) и 0,4 и 0,1 (в пользу второй гипотезы).

Соответствующие правила сложения m_i^+ и m_i^- приводятся в Таблице 1. Таблица примерно соответствует расчету по формуле:

$$\frac{m_i^+ + (1 - m_i^-)}{2}$$

с тем, однако, отличием, что под «мало» («малая часть») мы все же понимаем, скорее, некое меньшее количество, чем 0,25, а под «много» («большая часть») — скорее, большее, чем 0,75. Поэтому удобнее воспользоваться не какой-либо уточненной формулой, а таблицей. За этим внешне малым отличием имеется большой теоретический бэкграунд, обсуждать который подробно мы не будем; если коротко, то он состоит в том, что для прогнозов в нечеткой логике больше, нежели концепция нечетких множеств Заде, подходит особая концепция интуитивистских нечетких множеств¹⁶.

¹⁶ Строго говоря, отказавшись следовать слепо за формулой, оперирующей функцией принадлежности, мы перешли от обычных нечетких множеств к интуитивистским нечетким множествам, в которых, наряду с функцией принадлежности $\mu(x)$, выражающей степень принадлежности элемента x нечеткому множеству X , вводится «функция непринадлежности» $\nu(x)$, что позволяет присвоить x так называемый интуитивистский индекс $\pi(x) = 1 - \mu(x) - \nu(x)$, который служит мерой неопределенности самой принадлежности элемента x множеству X . Этот подход разработан Красимиром Атанасовым: *Atanassov K.T. Intuitionistic Fuzzy Sets // Fuzzy Sets and Systems. 1986. Vol. 20. P. 87—96; подробно см.: Atanassov K.T. Intuitionistic Fuzzy Sets. Theory and Applications (Studies in fuzziness and soft computing. Vol. 35). Berlin—Heidelberg: Springer-Verl. 1999, и, особенно: idem. On Intuitionistic Fuzzy Sets Theory (Studies in fuzziness and soft computing. Vol. 238). Berlin—Heidelberg: Springer-Verl., 2012.*

Согласно этим правилам, в рассмотренном выше примере ($m_i^+ = 1$, $m_i^- = 0,5$) $m_i = 0,75$. Здесь и ниже наши расчеты ведутся при выборе сирийской гипотезы в качестве первой, а славянской в качестве второй, т.е. «не-сирийской» (мы могли бы поступить и противоположным образом: тогда надо было бы m_i^+ обозначать результаты для славянской гипотезы, а m_i^- — для сирийской)

Таблица 1.

$m_i^- \backslash m_i^+$	0	0,25	0,5	0,75	1
0	–	0,6	0,75	0,9	1
0,25	0,4	–	0,6	0,75	0,9
0,5	0,25	0,4	–	0,6	0,75
0,75	0,1	0,25	0,4	–	0,6
1	0	0,1	0,25	0,4	–

Читатель, знакомый со стандартными примерами из теории вероятностей, уже мог заметить, что мы тут попытались формализовать нечеткий аналог бросания монеты, у которого может быть только два исхода: орел или решка (один из них условно называется «успешным», а другой «неуспешным»); тогда m подсчитывается как количество успешных бросков. Наша «нечеткая» монета ведет себя похоже, но не вполне так¹⁷. Она всегда заваливается набок, но никогда не лежит и не стоит. Даже если кажется, что выпал орел (или решка), то на самом деле это лишь наш глаз не может отличить настоящее положение монеты от горизонтального (поэтому мы записываем в счет m то ли ноль, то ли единицу, но в действительности это нечеткие величины, *похожие до неотличимости* на 0 и 1). В действительности монета лежит под углом, хотя и очень малым. У нас такая монета, которая вообще не может выпасть иначе, чем под углом. Гораздо чаще этот угол настолько велик, что заметен невооруженным глазом, и для его значений легко подобрать подходящее термножество. Но, как всякая порядочная монета, наша монета тоже никогда не выпадает на ребро, причем, даже на нечеткое ребро. Если состояния, практически неотличимые от горизонтальных (обоих), для нее возможны, то вертикальное состояние невозможно как таковое, даже если оно, строго говоря, не является вертикальным, а лишь похоже на вертикальное до неотличимости. На ребро не падают даже нечеткие монеты – так что даже в мире нечеткости *Ordnung muß sein* («совесть тоже надо иметь»).

8. Применение формального аппарата на основе нечеткой логики

Попытаемся теперь сравнить славянскую и сирийскую гипотезы, оценивая потенциально зависимые от этих гипотез факты в трех оставшихся (после исключения лексической) областях: фонетико-лексикологической, контекстуальной и текстологической.

В фонетико-лексикологической части: оценивается реконструкция трех сегментов, которые, согласно славянской гипотезе, являются искаженными славянскими словами, а, согласно сирийской гипотезе — транслитерациями сирийских слов. Всего наблюдений в этой области три, соответственно количеству сегментов.

¹⁷ Благодарю за помощь в описании этой монеты О.В. Митренину.

Для первого сегмента (същице): доступное в идеализированной славянской гипотезе слово со значением «сучки» подходит идеально, что дает $m_i^- = 1$; в сирийской реконструкции такого идеального совпадения нет, но все же соответствие неплохое, что дает $m_i^+ = 0,75$; получаем $m_1 = 0,4$.

Для второго сегмента (скокине): доступное в идеализированной славянской гипотезе слово «кокин» (славянское, а не греческое слово, означающее красную нить или шнурок, причем, оформленное по правилам не церковнославянской, а македонской грамматики; предложено Загребиным) дает $m_2^- = 1$; в сирийском соответствии также является точным, что дает $m_2^+ = 1$. Для этого случая нельзя будет рассчитать m_2 , и он просто исключается из дальнейшего сопоставления. Косвенно он все равно будет учитываться — на контекстуальном уровне. Таким образом, наше наблюдение номер 2 оказалось неудачным, то есть неселективным. Бывает.

Для третьего сегмента (соуголе) даже идеализированная славянская гипотеза дает не очень хороший результат: все разночтения с «г» предполагают сильное искажение текста в рукописи, а разночтение с «п», хотя и предполагает не такое сильное искажение, имеет само очень малую вероятность быть аутентичным. Поэтому, в лучшем случае, имеем здесь $m_3^- = 0,5$. В то же время, для сирийской гипотезы имеем здесь $m_3^+ = 1$. Окончательно получаем $m_3 = 0,75$.

В контекстуальной части: оценивается, во первых, соответствие реконструкции темного места контексту легенды и, во-вторых, наличие в других местах текста каких-либо признаков, свидетельствующих в пользу того или иного языка оригинала.

Для оценки соответствия контексту легенды реконструированной части нужно принять во внимание, как минимум, следующие четыре параметра, которые мы пронумеруем, продолжая нумерацию наших прежних наблюдений.

4. Соответствие сброшенных предметов буквам: тут надо учитывать, что Кирилл, по легенде, считает сами сброшенные на него предметы, а не буквы, на них нанесенные (как реконструируют смысл этой сцены те, кто видят в сброшенных предметах «дощечки»). Впрочем, в рамках обобщенной славянской гипотезы доступно толкование «сучки», и тогда тут проблемы не возникает. Но зато возникает другая проблема — объяснить, почему буквы оказались представлены какими-то сучками. Поэтому у сторонников славянской гипотезы тут, в любом случае, не все гладко, хотя и не все плохо: $m_4^- = 0,5$. Сирийская гипотеза предлагает видеть в сброшенных предметах писчие перья, и тут связь с буквами очевидна. Поэтому $m_4^+ = 1$. Получаем $m_4 = 0,75$.

5. Символизм нового Завета: о нем подумали те сторонники славянской гипотезы, которые предлагают видеть в сброшенных предметах две дощечки, аналогичные скрижалям Моисея. Это довольно хорошо, но не идеально, т.к. новый Завет тем и отличается от ветхого Моисеева, что он пишется не на скрижалях, а на сердце. Специальные перья для письма на сердце предполагаются в сирийской гипотезе. Поэтому $m_5^- = 0,75$, $m_5^+ = 1$, $m_5 = 0,6$.

6. Образ птицы (и желательно именно ворона): в рамках славянской гипотезы он не имеет жесткой связи с остальным сюжетом, хотя и никак не мешает его остальным элементам. Это означает $m_6^- = 0,5$. В сирийской гипотезе, где появляются птичьи перья для письма, причем, очень вероятно, что именно ворона, — $m_6^+ = 1$. Получаем $m_6 = 0,75$.

7. Проникновение сброшенных предметов в тело: этот сюжетный ход не особенно совместим с деревянными дощечками, но кое-как совместим с сучками. Поэтому для идеализированной славянской гипотезы можно принять $m_7^- = 0,25$. С перьями для письма он, разумеется, совместим идеально, и поэтому $m_7^+ = 1$. Получаем $m_7 = 0,9$.

Следующие три наблюдения относятся к тексту СЛ в целом, а не специально к темному месту.

8. Синтаксис: чрезвычайно много союзов «и» в начале фраз, причем видно, как редакторы старались от части из них избавляться. Это не просто нехарактерно для оригинальных текстов на славянских языках, но нехарактерно даже и в качестве библицизма. Я бы вообще сказал, что в оригинальном славянском тексте такой синтаксис невозможен, но сторонники славянской гипотезы будут слишком сильно протестовать против такого радикального вывода, и поэтому я сформулирую мягче: $m_8^- = 0,25$ (сделаю вид, что изредка такой синтаксис возможен — хотя, на самом деле, едва ли можно найти хотя бы один близкий пример). Что касается сирийского языка, то в нем подобный синтаксис не только возможен, но и нормативен, поэтому $m_8^+ = 1$. Получаем $m_8 = 0,9$.

9. Уже разобранный выше пример с идиомой «сходить (в город и т.п.)» в значении «входить»: $m_9^- = 0,5$; $m_9^+ = 1$; $m_9 = 0,75$.

10. Идиома «ни гласа», когда по-славянски естественнее было бы сказать «ни слова» («я не слышал *ни гласа* о Болгарской земле»). В славянских языках «глас» в значении «сведения, слух» употребляется, но не особенно часто, тогда как в сирийском весьма часто употребляемое слово имеет в качестве обычных значений и «речь, слово», и «глас». Ситуацию можно оценить как $m_{10}^- = 0,25$; $m_{10}^+ = 0,75$; $m_{10} = 0,75$.

11. Последнее наблюдение относится к области **текстологии**: в чем причина «темноты» темного места? Славянская гипотеза тут оказывается во власти парадокса: чем яснее предлагаемое ей объяснение, тем труднее объяснить, почему такой текст был искажен еще на уровне общего архетипа всех рукописей. Приходится допускать, что текст в этом месте изначально содержал что-то темное и заумное — но это плохо вяжется с остальным, весьма ясным текстом легенды. Гипотеза о переводе с сирийского предполагает, что переводчик не смог правильно интерпретировать ключевое слово, означавшее в данном тексте перо для письма, сделанное из птичьего пера, а поэтому не понял и всей синтагмы, построенной вокруг этого слова, оставив ее в транслитерации вместо перевода. Предположим с осторожностью (завышая, как мне представляется, значение для славянской гипотезы и занижая для сирийской), что тут $m_{11}^- = 0,5$; $m_{11}^+ = 0,75$; $m_{11} = 0,6$.

Итак, мы попытались сделать 11 наблюдений, из которых нам удалось 10. Получаем для них $n = 10$, а для m — по формуле (3) $m = 0,4 + 0,75 + 0,75 + 0,6 + 0,75 + 0,9 + 0,9 + 0,75 + 0,75 + 0,6 = 7,15$.

9. Оценка апостериорного компонента соотношения правдоподобий гипотез

Напомню, что мы будем исходить из того, что $q + r = 1$. Будем рассматривать результаты вычислений для все тех же значений $r = 0,8; 0,9; 0,95$ ($q = 0,2; 0,1; 0,05$), что соответствует разным предположениям о когерентности текста СЛ, но не зависит от сопоставляемых гипотез. За первую гипотезу мы уже приняли сирийскую, поэтому именно сирийская гипотеза принимается за h_q , а славянская — за h_r . (Эта процедура выбора «первой» гипотезы из двух никак не влияет на результат, а влияет лишь на то, будет ли конечное отношение правдоподобностей гипотез большим целым числом или малым дробным).

Получаем следующие результаты для трех разных оценок консистентности текста (0,8; 0,9; 0,95):

$$(4.1) \quad \frac{P[e^{10} | h_{0.8} \cdot b \cdot c^{10}]}{P[e^{10} | h_{0.2} \cdot b \cdot c^{10}]} = \frac{0.8^{7.15} (1-0.8)^{10-7.15}}{0.2^{7.15} (1-0.2)^{10-7.15}} = 388.023 \approx 388$$

$$(4.2) \quad \frac{P[e^{10} | h_{0.9} \cdot b \cdot c^{10}]}{P[e^{10} | h_{0.1} \cdot b \cdot c^{10}]} = \frac{0.9^{7.15} (1-0.9)^{10-7.15}}{0.1^{7.15} (1-0.1)^{10-7.15}} = 12683.607 \approx 12684$$

$$(4.3) \quad \frac{P[e^{10} | h_{0.95} \cdot b \cdot c^{10}]}{P[e^{10} | h_{0.05} \cdot b \cdot c^{10}]} = \frac{0.95^{7.15} (1-0.95)^{10-7.15}}{0.05^{7.15} (1-0.05)^{10-7.15}} = 315239.394 \approx 315239$$

10. Обсуждение результатов: теоретическая сторона

Сейчас самое время поговорить о том, что же мы, в итоге, навывчисляли, и к чему это можно применить. Прежде чем обратиться к нашим практическим задачам, скажем теперь корректнее о том, откуда взялись формулы, которыми мы пользовались выше, а также о том, какие еще существуют формулы, чтобы мы смогли понять, как можно применить полученные результаты.

Основой всех наших вычислений является теорема Байеса, примененная к понятию правдоподобности гипотезы; само это понятие было определено выше согласно выражению (1). Теорема Байеса применительно к правдоподобности гипотезы h_i имеет вид:

$$(5) \quad P[h_i | b \cdot c^n \cdot e^n] = \frac{P[e^n | h_i \cdot b \cdot c^n] \times P[h_i | b]}{P[e^n | b \cdot c^n]} \times \frac{P[c^n | h_i \cdot b]}{P[c^n | b]}$$

Последний ее член учитывает влияние гипотезы на сами условия наблюдений или экспериментов. Вероятно, этот член должен учитываться, например, тогда, когда мы строим гипотезы об аутентичности тех или иных рукописных чтений, но доступные нам варианты статистически репрезентативно тяготеют к определенным субархетипам. Если же гипотеза не влияет на условия наблюдений или ее влиянием можно пренебречь, то последний член равен единице и поэтому может быть пропущен при записи (5).

Членом $P[e^n | b \cdot c^n]$ определяется вероятность данного исхода наблюдения при данных условиях и при том же бэкграунде b , если бы никаких особых гипотез по его поводу не выдвигалось. Когда мы дальше будем сравнивать отношения правдоподобностей разных гипотез, то от соответствующих членов нам удастся избавиться, а поэтому практическим вычислением этой величины нам заниматься нет нужды.

Ключевое значение, помимо уже знакомого нам члена $P[e^n | h_i \cdot b \cdot c^n]$ (апостериорной правдоподобности), имеет член $P[h_i | b]$ — априорная правдоподобность: насколько правдоподобной представляется данная гипотеза до того, как мы начали делать наши наблюдения. Сюда входят не только чисто психологические компоненты, связанные с психологическим устройством научного сообщества (в котором, как показал Томас Кун, радикально меняющие прежнюю научную картину данные обычно отвергаются до тех пор, пока не сменится поколение ведущих ученых¹⁸). Отчасти это и представления научного сообщества о имеющих отношение к оцениваемой гипотезе данных из смежных дисциплин. Например, в случае гипотез о СЛ — это данные об истории христианства у славян и истории

¹⁸ *Kuhn Th. S.* The Structure of Scientific Revolutions. 50th Anniversary edition. Chicago: University of Chicago Press, 2012 (много раз издавалась в рус. пер.).

агиографических традиций, связанных с созданием славянской письменности. Эти данные не могут быть безразличными для создания убедительной теории о происхождении СЛ, но мы никаких наблюдений в этих областях не ставили, и поэтому имеем для них только априорные оценки.

Из теоремы Байеса, то есть выражения (5), видно, что при сравнении правдоподобности разных гипотез (при условии одинаковых бэкграундных знаний, отсутствия влияния выбора гипотезы на условия наблюдений, одинаковых методов интерпретации фактов) решающая роль будет принадлежать члену $P[e^n | h_i \cdot b \cdot c^n] \times P[h_i | b]$. Выше мы оценивали лишь первую часть этого выражения, используя формулу (2). Но полная формула для определения соотношения правдоподобности разных гипотез учитывает и второй член (при этом я предполагаю, что член, связанный в влиянием выбора гипотезы на условия, учитывать не надо, т.к. такого влияния нет совсем или нет практически). Это не что иное, как снова теорема Байеса, но записанная в форме соотношения:

$$(6) \quad \frac{P[h_j | b \cdot c^n \cdot e^n]}{P[h_i | b \cdot c^n \cdot e^n]} = \frac{P[e^n | h_j \cdot b \cdot c^n]}{P[e^n | h_i \cdot b \cdot c^n]} \times \frac{P[h_j | b]}{P[h_i | b]}$$

Для пробабилистически независимых наблюдений (как мы приняли для наших наблюдений над СЛ) выражение (6) может быть записано в следующей форме:

$$(7) \quad \frac{P[h_j | b \cdot c^n \cdot e^n]}{P[h_i | b \cdot c^n \cdot e^n]} = \frac{P[h_j | b]}{P[h_i | b]} \times \prod_{k=1}^n \frac{P[e^k | h_j \cdot b \cdot c^k]}{P[e^k | h_i \cdot b \cdot c^k]}$$

Выражение (7) показывает, что (при условии пробабилистической независимости наблюдений) гипотезы могут сопоставляться по результатам каждого из наблюдений в отдельности. Выше мы воспользовались этим свойством наших данных, чтобы исключить ряд наблюдений из расчетов заранее — в тех случаях, где соотношение правдоподобности исхода наблюдений в пользу каждой из гипотез было неотлично от единицы.

Из выражений (6) и (7) видно, почему высокое соотношение апостериорных правдоподобностей в пользу сирийской гипотезы еще не означает автоматически, что она выигрывает у славянской в общем зачете. Совершенно очевидно, что в соотношении априорных вероятностей может выигрывать только славянская гипотеза. Поэтому окончательная оценка отношения правдоподобий в целом будет зависеть от соотношения порядков величин, с одной стороны, соотношения априорных правдоподобностей, а, с другой стороны, апостериорных.

11. Обсуждение результатов: практическая сторона

Теперь мы должны будем оценить наши результаты (4.1-3) в связи с тем, что можно сказать об априорном соотношении вероятностей сирийской и славянской гипотез. Конечно, могут найтись такие ученые, которые считают априорную вероятность сирийской гипотезы неотличимой от нуля, и с ними тут обсуждать просто нечего. Их априорная убежденность «помножит на ноль» (здесь в буквальном смысле) любые результаты наблюдений.

В соотношении априорных вероятностей двух наших гипотез хотелось бы выделить составляющую, которая хотя бы очень ориентировочно может быть оценена апостериорно. Для этого можно предложить несколько разных подходов, дополняющих друг друга.

Подход филологический. Будем оценивать вероятность оказаться переводом с сирийского (прямым, минуя греческий посредник) для любого *нам известного* произведения, написанного на славянском языке до XV в. включительно (времени древнейшей рукописи СЛ). Представление об общем объеме этой письменности может быть получено на основе различных справочников, но, к сожалению, общей базы данных, аналогичной *TLG*, пока нет. На мой взгляд, речь может идти, максимум, о нескольких тысячах произведений. Во всяком случае, количество независимых произведений (не редакций), учтенных в сумме в *Clavis Patrum Graecorum* и в *Bibliotheca Hagiographica Graeca* составляет что-то около 10 000 (скорее, меньше). Но это все произведения греческой патристики (независимо от того, на каких языках они сохранились) до VIII в. включительно плюс все произведения агиографии на греческом языке, включая послевизантийское время. Славянская письменность указанного выше периода едва ли располагает и парой тысяч произведений (при том, что мы включаем в это число и всю древнерусскую литературу, т.к. в ней нередко сохранялись южнославянские памятники, утраченные в письменности южных славян).

Далее сделаем два следующих альтернативных предположения: [1] либо, как я попытался показать в других работах, из этого числа не менее четырех произведений, не считая СЛ, являются прямыми переводами с сирийского¹⁹; это предположение можно ослабить, заменив «не менее четырех», скажем, на «не менее двух»; [2] либо среди этих произведений переводом с сирийского может быть, максимум, одно (а я в своих прочих исследованиях неправ).

Для первого предположения мы имели бы правдоподобность появления в славянской письменности указанного периода прямого перевода с сирийского порядка нескольких единиц во втором десятичном знаке (результат деления «не менее двух» на «не более 2000»), то есть правдоподобность измерялась бы в сотых долях. Для второго предположения она была бы на порядок меньше, то есть измерялась бы в тысячных долях.

Различие между [1] и [2] принципиально для оценки результата (4.1). Напомню, что это результат такого расчета, в котором предполагается, что СЛ написана неумело и переписывалась небрежно.

Если принять [1], то (4.1) интерпретируется как приблизительная равновероятность славянской и сирийской гипотезы, а если принять [2], то как проигрыш сирийской гипотезы славянской, хотя и не с разгромным счетом.

Конечно, если бы в (4.1) у нас была бы заложена не идеализированная и в реальности невозможная славянская гипотеза, а любая из реально предложенных, то результаты и там были бы резко в пользу сирийской гипотезы. Но таким расчетом мы бы не решили поставленной в начале задачи — оценивать не конкретные предложения толкователей СЛ как оригинального славянского произведения, а потенциал той научной парадигмы, в которой они думают.

Подход исторический: СЛ, как и всякая легенда, не может быть вне той исторической реальности, которой она порождена. Это реальность до-кирилломефодиевского христианства у славян, которая каким-то образом представлена в СЛ, чей заголовок в не так давно (1966) обнародованной рукописи *S* эксплицитно обозначает сюжет легенды как крещение славинии. Но идеализированная славянская гипотеза, в отличие от большинства реальных, и не настаивает на после-кирилломефодиевском происхождении легенды. В последнее время такой взгляд высказал Иван Добрев (2003). Также и христианство второго хана Первого

¹⁹ *Lourié B.* Direct Translations into Slavonic from Syriac: a Preliminary List // *FS Mario Capaldo / C. Diddy et al* (eds.) (forthcoming).

Болгарского царства, Тервеля (695–721), хотя и получает в последние 20 лет признание ученых, ничего не говорит в пользу сирийского происхождения СЛ. Большинство исторических фактов о раннем христианстве у славян по отношению к нашим гипотезам нейтральны.

Резко в пользу сирийской гипотезы говорит только один исторический эпизод. Он зафиксирован с идеальной достоверностью (в протоколах Шестого Вселенского собора, XVI заседание от 9 августа 681 г.), но так как он только один, на нем нельзя выстраивать статистику (эпизодов селективных в пользу славянского происхождения СЛ пока неизвестно вообще, поэтому у нас тут вся выборка была бы из одного-единственного наблюдения). Все это заседание собора было посвящено выслушиванию предложения об унии между сторонниками одной и двух волей во Христе. Выступал плохо говоривший по-гречески сирийский пресвитер Константин из Апамеи. Выступая, он жаловался, что если бы в самом начале собора (сентябрь 680 г.) его пустили бы выступить, как он просил, то сейчас не было бы недавнего (лето 681 г.) поражения в войне с «Болгарией». Театром военных действий в упомянутой войне была византийская территория, где только что была учреждена фема Фракия. Ее новоназначенный глава (некто вроде генерал-губернатора) патрикий Феодор и был тем человеком, как это отмечено в протоколах собора, которому, в конце концов, отцы собора уступили, согласившись послушать Константина Апамейского (это не помешало им его в итоге анафематствовать).

Этот исторический эпизод относится к той самой войне, которая в 681 г. привела к созданию Первого Болгарского царства на Дунае и к вовлечению в орбиту его влияния славиний к югу от него, на территории Византии. Другим эпизодом той же самой войны является трехлетняя (676—678) осада Фессалоник болгарами и славянами, ставшая одним из центральных сюжетных узлов СЛ. Случай Константина Апамейского доказывает нам непосредственно, что византийское гражданское и военное начальство (то и другое в лице патрикия Феодора) считало влияние сирийцев некоего не византийского вероисповедания значимым для славян и болгар того места и того времени, о котором говорит СЛ. Это, в свою очередь, доказывает нам — опосредованно, но не менее достоверно, т.к. компетентности патрикия Феодора целесообразно доверять, — что какая-то эффективно работавшая сирийская миссия среди славян и болгар Фракии (в первоначальных границах одноименной фемы, т.е. включая будущую фему Македония) имела место на самом деле.

Возвращаясь к результату наших расчетов (4.1), мы можем отметить, что исторические данные способны сдвинуть наши оценки правдоподобности в пользу сирийской гипотезы в том случае, если без их учета у нас была ситуация, близкая к равновесию (см. условие [1] выше). Но если сирийская гипотеза проигрывала славянской (условие [2]), то исторические соображения могут, в лучшем случае, помочь приблизить общую оценку к равновесной.

Подход критико-агиографический, в отличие от чисто исторического подхода, мне кажется крайне многообещающим. Он предполагает полный анализ агиографических легенд, связанных с разными культами святых и разными версиями легенд о создании письменности в раннем славянском христианстве, где можно выделить значительный сирийский субстрат, но об этом я постараюсь написать отдельное исследование.

Разумеется, возможны и **другие подходы**, так или иначе способные доставить нам данные для наблюдений, селективных по отношению к нашим альтернативным гипотезам о СЛ. Но уже и сказанного выше достаточно, чтобы представить себе, за счет чего может меняться та составляющая общей сравнительной правдоподобности двух наших гипотез, которая в нашем исследовании рассматривалась как априорная.

Теперь применим эти знания к оценке следующего результата — (4.2). Этот расчет сделан в предположении, что СЛ составлена профессиональным и ответственным автором (то есть таким автором, который не составлял свой текст из небрежно подогнанных фрагментов более ранних источников, и сам при этом обладал достаточными навыками художественного изложения) и переписывалась со среднестатистической аккуратностью, а не как-то особенно неграмотно или небрежно.

При оценке (4.2) на фоне того, что мы выше описывали как результаты филологического подхода к оценке соотношения априорных вероятностей, мы видим уверенную победу сирийской гипотезы. Даже неблагоприятное для сирийской гипотезы предположение [2] (будто вероятность для славянского произведения оказаться переводом с сирийского должна измеряться в тысячных, а не сотых долях) не позволяет уничтожить апостериорный 12-тысячный перевес сирийской гипотезы. Дальнейшие априорные соображения (хотя бы эпизод с Константином Апамейским) дополнительно увеличивают этот разрыв в пользу сирийской гипотезы.

Третий результат наших расчетов, (4.3), выполнен исходя из предположения о крайне высокой степени упорядоченности как текста, так и рукописей СЛ. Такое предположение не претендует на реалистичность, но было полезно для того, чтобы посмотреть, насколько быстро возрастает селективность наших наблюдений в зависимости от выбираемого предположения о связности текста. После отметки 0,9 счет быстро переходит на сотни тысяч раз.

Итак, по моему представлению, ближе всего к реальности оценка наших наблюдений по расчету (4.2), что дает в итоге не разгромный, но уверенный счет в пользу сирийской гипотезы.

Сирийская гипотеза сможет считаться доказанной в том случае, если — за счет каких-либо других наблюдений и априорных, но разумных соображений — мы сможем увеличить разрыв между правдоподобиями альтернативных гипотез хотя бы до десятков тысяч раз. Как отчасти видно из только что изложенных соображений, вероятности в тысячных долях — это самые малые вероятности, которыми еще нельзя пренебрегать.

В.М. Лурье

Солунская легенда:

оригинальное славянское произведение или перевод с сирийского?

Количественный подход к оценке правдоподобности альтернативных гипотез

Для количественного сравнения двух гипотез о происхождении «Солунской легенды» использован метод индуктивной логики, опирающийся на расчет отношения байесовых правдоподобий альтернативных гипотез. Расчет, в свою очередь, базируется на методах нечетких теорий вероятности и прогнозирования. Метод удобен для количественной оценки правдоподобия альтернативных гипотез в любой гуманитарной области, где количество всех подлежащих рассмотрению гипотез можно свести к двум.

Ключевые слова: «Солунская легенда», первоначальная славянская письменность, славянская агиография, индуктивная логика, правдоподобность гипотезы, теорема Байеса, интуитивные нечеткие множества.

Basil Lourié

The Thessalonican Legend:

Original Work in Slavonic or Translation from Syriac?

A Quantitative Approach to Evaluation of the Likelihoods of the Alternative Hypotheses

The two hypotheses on the original language of the *Thessalonican Legend* are compared in a quantitative way using the procedures of Bayesian inductive logic. The ratio of likelihoods is calculated as that of fuzzy likelihoods using the apparatus of the theory of (intuitive) fuzzy sets. The method could be convenient in humanities for quantitative evaluation when the number of alternative hypotheses is reducible to two.

Keywords: Thessalonican Legend, early Slavic writing, Slavic hagiography, Bayesian inductive logic, intuitive fuzzy sets, fuzzy likelihood.