

**ИНФОРМАЦИОННЫЙ БЮЛЛЕТЕНЬ
АССОЦИАЦИИ**

“ИСТОРИЯ И КОМПЬЮТЕР”

№ 31, СЕНТЯБРЬ 2003

**МОСКВА
2003**

РЕПОЗИТОРИЙ БГПУ

СОДЕРЖАНИЕ

КОНФЕРЕНЦИИ, СЕМИНАРЫ

<i>Владимиров В.Н., Гарскова И.М.</i> VIII конференция Ассоциации «История и компьютер»: историческая информатика в начале XXI века	3
Фоторепортаж с VII конференции АИК	9
Фоторепортаж с XV Международной конференции АНС	30
XV Международная конференция ассоциации "History and Computing" <i>Володин А.Ю.</i>	33
<i>Валетов Т.Я.</i>	39
<i>Пиотух Н.В.</i>	42
<i>Сидорцов В.Н., Балыкина Е.Н.</i> Педагогические аспекты исторической информатики – 2003: научный семинар в Гродно	45

ИСТОРИЯ В ИНТЕРНЕТЕ

<i>Сибиряков И.В.</i> Проблемы гражданской войны в сети Интернет ...	50
<i>Морозов Д.Б.</i> Тематические сайты по истории техники (на примере интернет-ресурса "СВВП"	55

РУНЕТ: ИСТОРИЯ И СОВРЕМЕННОСТЬ

<i>Иванов Д.Г.</i> Информационный сектор российского интернета: 1990–2002 гг. (хронологическая таблица)	56
<i>Иванов Д.Г.</i> Специфика аудитории российского интернета	72
<i>Иванов Д.Г.</i> Федеральная целевая программа "Электронная Россия": попытка реализации долгосрочного проекта по информатизации органов государственной власти	80

КОМПЬЮТЕРИЗАЦИЯ УЧЕБНОГО ПРОЦЕССА

<i>Джалишвили З.О., Кириллов А.В.</i> У истоков информатизации отечественного гуманитарного образования	87
<i>Вержбицкий В.В., Власова Ю.Ю.</i> Социальная среда информатизации образования в Российской Федерации	94
<i>Можяева Г.В.</i> Интернет для историков: о модели построения тематического образовательного портала	110
<i>Балыкина Е.Н., Макаров А.В.</i> «Проектирование и разработка тестовых сред по циклу социально-гуманитарных дисциплин в образовательных системах» (учебная программа)	113
<i>Володин А.Ю.</i> Мультимедийное преподавание статистики	129
<i>Лозицкий В.Л.</i> Из опыта создания компьютерной обучающей программы "Бакалавр"	138
<i>Миницкий Н.И., Толочко А.Ф.</i> Информационные технологии структурного анализа исторического текста	151
<i>Вин Ю.Я., Гриднева А.Ю., Кондратьев Д.Е., Тихонова О.В.</i> Концепция семантического поля исторического источника: основные положения	166
<i>Груздев Д.В., Журбин И.В.</i> Интерполяция и экстраполяция при пространственном моделировании археологических объектов	177

БАЗЫ ДАННЫХ

<i>Пушков Л.В., Пушков В.П., Завьялов С.М., Гришина З.В.</i> Выпускники Московского университета и формирование интеллектуального потенциала дореволюционной России (первые итоги создания базы данных)	184
<i>Чайченко Н.С.</i> К исследованию социальных характеристик оппозиционно настроенных студентов Московского университета по данным Агентурного отдела Московского охранного отделения, 1902–1916 гг. (база данных и ее анализ)	193

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ СТРУКТУРНОГО АНАЛИЗА ИСТОРИЧЕСКОГО ТЕКСТА

Прагматическая цель нашего исследования – предложить решение некоторых вопросов проблемы сложности учебного текста на основе применения информационных технологий. Существование этой проблемы имеет долговременный и многоплановый характер. К настоящему времени сформировались ее следующие аспекты: теоретико-содержательный, концептуально-методологический, логико-понятийный, лингвистический, информационный. Для историков традиционно доминирующими были теоретико-содержательный и концептуально-методологический факторы. В последнее время все большее внимание привлекает решение названной проблемы на методологической основе лингвистического анализа с использованием компьютерных технологий.

Но, если будут исследованы все названные стороны, то и в данном случае мы будем далеки от окончательного решения проблемы трудности текста. Перечисленные аспекты относятся к рациональному осмыслению текста. В стороне остается проблема иррационального или бессознательного восприятия учебной информации. Отнюдь не отвергая и не умаляя рациональных методов ее усвоения, а, опираясь на принцип дополнительности, мы предлагаем метод структурного анализа текста, в котором учитывается элемент бессознательного восприятия текста. Он основан на применении психолингвистического и математического моделирования с помощью информационных технологий.

На наш взгляд, потребность такого подхода обусловлена состоянием современной учебной литературы по истории. Создалась ситуация, которая сейчас требует значительных усилий и коренной перестройки всей научно-методической и организационно-издательской работы. В частности, несмотря на существующую систему контроля за качеством учебников, не на должном уровне остаются язык и стиль учебной литературы. Учебники не соответствуют существующим стандартам и превышают объем допустимой учебной нагрузки. Ранее эти задачи решались за счет многократной переработки и экспериментальной проверки учебной литературы в школах. Это удлиняло время процесса изготовления и выпуска учебной литературы. Реалии сегодняшнего дня в Республике Беларусь совершенно иные: резко увеличилось количество типов учебных заведений, идет сложный процесс перехода на 12-летнее обучение в средней школе, меняются формы обучения в высшей школе, много трудностей в экономике и т.д.

В этой связи применение старых архаичных методов оценки качества учебной литературы, когда вся оценка этого качества производится только

на основе использования человеческого фактора, проблемы не решает, и нарекания на низкое качество учебной литературы продолжают.

Это порождает необходимость создания специальных технологий, способствующих росту качественного уровня учебников, активизации научных исследований в этой области. В частности, существует актуальная потребность в разработке диагностических показателей качества, ориентированных на объективные, воспроизводимые методы контроля, содержащих количественные критерии его оценки; проведении систематической и регулярной процедуры сбора данных по важным образовательным аспектам (мониторинга) и т.п. К сожалению, эти работы должным образом не развернуты, а в дискуссиях современных авторов учебников по истории доминируют проблемы концептуально-содержательного характера.

Решение проблемы восприятия текста учебника связано со следующими противоречиями: 1) растущим разрывом между объективно увеличивающимся объемом научного знания и тем объемом, который может быть усвоен учащимися; 2) применением сложных логических конструкций знания и игнорированием учета возрастного уровня мышления; 3) применением слишком сложных конструкций предложений в учебном тексте. 4) отсутствием согласования между вербально-логической и знаково-символьной информацией. В итоге следует констатировать, что сегодняшние учебники по истории порой превышают допустимые объемы, им характерны логическая и стилистическая сложность, психологический дискомфорт в восприятии текста.

Условия для решения существующих проблем имеются. Теоретические предпосылки для решения проблем текста учебной литературы содержатся в исследовательских работах по теории учебника, а также в области технологий обработки текста и представления знания. Здесь сложность заключается в многочисленности литературы и выборе наиболее эффективных технологий, проверке самих технологий, которые были разработаны еще в 60-х годах прошлого века. Существующие ранее и разработанные в самое последнее время технологии обработки текста и представления учебного знания уже сегодня позволяют приступить к позитивному решению названных проблем.

Подобная технология может применяться к текстам как естественнонаучного, так и гуманитарного знания. В то же время нами учитывается предметная специфика гуманитарного и, в частности, исторического знания (предметом обработки могут стать любые тексты). Существует довольно значительное количество литературы по психолингвистике. Работы в этом направлении велись нами и ранее¹, но для данного случая применена иная, оригинальная технология, о которой и пойдет речь далее.

¹ Миницкий Н.И. Равков А.В. Звуко-цветовые ассоциации поэмы Драконтия «Хвала господе» («Сотворение мира») // Информационный бюллетень ассоциации «История и компьютер», апрель 1994, № 10, М., С. 65–67; Миницкий Н.И. Равков А.В. Колористика знака и текста (компьютеризованный ана-

* * *

В настоящее время остро встал вопрос о создании технологии обработки учебного текста и конкретного инструментария (математического и физического). И ранее для решения проблем, связанных с различного рода диагностикой, создавался специальный технологический инструментарий. Его возможности сегодня исчерпаны. Новые средства не могут быть созданы на основе той техники, которая существовала ранее. Только появление современных информационных технологий позволяет ставить и решать задачи диагностики учебника практически в полном объеме.

Кроме того, в других отраслях науки создались благоприятные предпосылки для позитивного решения проблемы текста учебника. К примеру, уже достаточно давно психологами установлено отражение в речи особенностей личности. В лингвистике имеются фундаментальные результаты в исследовании структуры текста. Диапазон исследований весьма широк и включает весь возможный спектр: от анализа структуры одного предложения до представлений о тексте как картины мира. Несмотря на значительные успехи в отдельных отраслях науки, пока еще не сложилась технология, адаптированная к потребностям гуманитарного образования. Тем не менее, современное развитие информационных технологий уже сегодня позволяет реально ставить и решать задачу использования скрытой информативности и экспрессивности языка в практических целях учебного книгоиздания.

Известно, что каждый язык характеризуется вполне определенной статистической (стандартной) структурой, которая не исчерпывается вероятностью появления отдельных букв. В языке всегда существуют дополнительные закономерности, что приводит к уменьшению степени неопределенности (энтропии) одной буквы. Анализ этой неопределенности показывает, что в письменном тексте содержится много "лишней" информации. Количество этой "лишней" информации характеризуется избыточностью языка. Избыточность – это величина, на которую удлиняются сообщения на данном языке по сравнению с минимальной длиной, необходимой для отображения той же информации. Подобный показатель является весьма важной статистической характеристикой языка.

Имеющиеся данные о статистической структуре различных языков позволяют считать, что избыточность наиболее распространенных европейских языков превышает 50%. Конкретные показатели избыточности национальных учебных изданий точно неизвестны, но они намного больше стандартной статистической нормы русского языка. К сожалению, следует

лиз) // Информационный бюллетень ассоциации «История и компьютер», №14, апрель 1995, М. С.124–126; Миницкий Н.И., Толочко Н.А. Компьютерный анализ звуковой метафоры поэтического текста // Информационный бюллетень ассоциации "История и компьютер". – Мн., №21, март 1997. С. 111–112.

отменить, что авторы учебников весьма редко обращались к исследованиям в области логической или семантической структуры языка изложения. В основном их подходы ограничивались интуитивной оценкой, либо сугубо частным анализом. Серьезнейшим образом в этом направлении начали работать специалисты информационных технологий, но решали они при этом сугубо частную задачу качественной передачи информации по каналам связи.

Сложилась достаточно характерная для современной эпохи ситуация, когда необходимо решить сложную научно-техническую задачу на стыке совершенно разных направлений деятельности: лингвистики, психологии, педагогики и информационных технологий. Здесь необходим полидисциплинарный подход, ибо специалистам гуманитарного профиля, самим эту задачу принципиально не решить. В системе учебного книгоиздания практическое решение этой проблемы возможно за счет поэтапной и постоянной работы творческой группы, состоящих из специалистов разного профиля. Можно начать с частных задач, несколько ограничивающих на первом этапе точность исследования, но позволяющих дать достаточную для практических целей оценку. Ниже будет показан один из возможных вариантов такого подхода.

В качестве первого этапа в сложившихся условиях наиболее предпочтительным по экономическим и научным соображениям является создание математического инструментария, т.е. математической модели статистической структуры языка на основе вероятности появления отдельных букв. Предварительных наработок на уровне постановки задачи для этого вполне достаточно. Необходимо, однако, установить правила соответствия, связывающие реальные объекты исследования (печатные издания) с определенными математическими объектами и отношениями, базирующимися на основе теории вероятности, математической статистики, векторных пространств и других областей математики. При исследовании проблемы следует выделить и учесть все существенные связи и одновременно отвлечься от несущественных. Следовательно, анализу подвергается не само явление во всей его сложности, а его упрощенная идеальная модель. Критерием правильности принятой модели являются результаты сравнения выводов теории с данными практики, опыта с проверкой на физической модели и затем испытания в реальных условиях.

В решаемой задаче разработка принципиально новой методики анализа текста на предмет выявления его скрытой информативности и экспрессивности осуществляется на основе использования относительных частот появления отдельных букв. Разработанная нами математическая модель исследования текстового материала позволяет выявить функцию возмущений (отклонений) частоты появлений отдельных букв алфавита на заданной выборке относительно этих же характеристик на генеральной совокупности русского языка (стандартной структуре).

Разработанный инструментарий (математическая модель) позволяет нивелировать текст и даже провести тестирование авторского коллектива на предмет психолингвистической совместимости. По психолингвистическим характеристикам нетрудно подсчитать общий процент отклонения рассматриваемого текста от статистической (стандартной) структуры языка. Установлено, что человек комфортно себя чувствует в естественной среде обитания, в том числе и языковой. Отклонения от этой среды автоматически ведут к отклонению от обычного состояния обучения, выражающемуся в удлинении времени усвоения материала, повышенной возбудимости либо утомляемости учеников. Проверка этой идеи на старых советских учебниках, получивших положительную оценку и прошедших многократное переиздание получила подтверждение. Отклонение от статистической структуры не превысило 5% и распределилось примерно равномерно по всем буквам (гласным).

В связи со сложностью процедур обработки, когда на первом этапе проведения работ используется полуавтоматический режим и исследуется много вариантов, создание дорогостоящего программного продукта нецелесообразно. В этих условиях ограничение исследований только гласными оказалось достаточным для практических целей, что в последующем не исключает полных исследований по всему алфавиту. На основе полученной модели исследования текстового материала по отклонениям частоты появлений отдельных букв алфавита можно в принципе давать рекомендации потенциальным авторам.

Умение писать изначально с незначительными отклонениями от статистической структуры языка, делает всех авторов близкими по стилю изложения и легкости восприятия. Если же это условие выполнить невозможно, то возникает необходимость разработки методов проведения оценки степени авторской совместимости. Степень отклонений от собственной лингвистической характеристики можно проверить на малых объемах, например, на одной странице текста.

В этом случае, возможно использование различных алгоритмов распознавания образов. Наиболее удобным при этом оказалось использование функций правдоподобия. Полученный коэффициент правдоподобия может стать вторым фактором, позволяющим использовать скрытую информативность используемого текста. Анализ случайных текстов это положение подтверждает, что позволяет эффективно производить корректировку стиля. Следует особо отметить, что процент отклонения от статистической структуры языка и коэффициент правдоподобия – принципиально разные показатели. Если первый определяется на основе генеральной совокупности относительно языка в целом, то второй является частным относительно конкретного автора, группы авторов, образца-эталона и т.п. Опираясь на оба фактора, мы узнаем общую закономерность чисто лингвистической характеристики текста.

Эта общая закономерность проявляется как частный случай в феномене неосознанного психологического восприятия цвета звукобукв. Данный

аспект позволяет создать новый вариант технологии обработки текста. Само явление колористического восприятия текста как смыслонесущей категории было известно еще со времени древних цивилизаций. Еще в Древней Греции Платон утверждал, что звуки, также как и предметы, характеризуются такими признаками как форма, объем, динамика и цвет. Несколько в другом аспекте эту же идею рассматривал Ломоносов, который утверждал: "В русском языке, как кажется, частое повторение письма А способствовать может к изображению великолепия, великого пространства, глубины и вышины; учащением писем Е, И, Ю – к изображению нежности, ласкательства, плачевных или малых вещей; через Я познать можно приятность, увеселение, нежность и склонность; через О, У, Ы – показать страшные и сильные вещи: гнев, зависть, боль и печаль...".

Как видим, идея фоносемантизма не нова, она является лишь частью глубоких исследований конца второго тысячелетия по теории, истории и значению цвета в жизни человека. И здесь имеется огромное количество литературы, которое мы в силу ее известности приводить не будем. Сошлемся лишь на основательный труд Л.Н.Мироновой «Учение о цвете», где излагается история знаний о цвете от первобытных и примитивных культур до нашего времени².

Феномен человеческого восприятия звуков в цвете (цветовой слух), являющийся психологической характеристикой звуков (букв), позволяет в нашем случае определить статус цвета как самостоятельную смыслонесущую категорию. Интегральная обработка различных материалов звуко-цветовых соответствий, учет того, что лингвисты считают гласные А, О, Е, И основными, а физики главными считают соответствующие им цвета: красный, желтый, зеленый, синий, позволяет составить графики звуко-цветовых соответствий.

Методика звуко-цветового анализа текста основана на частоте встречаемости звуков в речи (тексте). Как установили психологи, стандартную статистическую структуру мы в эмоционально-психологическом плане воспринимаем ровно и спокойно. Если звуки находятся в пределах нормы, то они не несут дополнительной нагрузки и, как бы не замечаются нами, а их значимость остается скрытой. Но если доля каких-то звуков заметно превышает норму, то доминирующим цветом анализируемого текста будет цвет звуко-букв, количественно превышающий норму своей частотности. Текст – это сложная и упорядоченная многоуровневая структура. Восприятие его тоже характеризуется многоуровневостью представления. Сюда входят сознательное осмысление и бессознательное восприятие, которое в психолингвистике рассматривается не как самостоятельная психическая реальность, противостоящая сознанию, а как низлежащий уровень сознания, характеризующийся меньшей расчлененностью и рефлексивностью.

² Миронова Л.Н. Учение о цвете. –Мн.: Выш. шк., 1993. 463 с.

Следует заметить, что в последнее время понимание значимости бессознательного в восприятии действительности возрастает.

Как уже указывалось, воспринимая окраску звуков, мы рефлекслируем интуитивно, подсознательно. Но это не умаляет значимость интеллектуальной рефлексии. Текст, окрашенный в определенные цвета, вызывает ряд эмоций на подсознательном уровне помимо сознательного его осмысления. Эмоции соответственно определяют возможные формы поведения субъекта, направленность его в принятии решения. В связи с этим и тексты в зависимости от назначения имеют разную цветовую окраску. Ранее нами отмечалось, что текст, имеющий минимальное отклонение от статистической структуры языка, наиболее пригоден в качестве образца – эталона для учебной литературы. В этом случае окраска звуков не превышает привычной нормы и не вызывает дополнительных эмоций на подсознательном уровне.

Иное дело поэзия. Художник слова стремится как можно полнее, ярче, живее выразить у слушателя и читателя нужные впечатления, переживания, размышления. Достигается эта цель всеми языковыми средствами, главное из которых – смысловая сторона языка с тонкой и бесконечно разнообразной игрой оттенков значения слов и их сочетаний. В стихотворении важны все аспекты формы: и ритм, и рифма, и композиция. И, конечно же, звуки – живая плоть стиха. Поэт со свойственной его таланту глубиной тонко чувствует содержательность звуков и точно оперирует ею, создавая звуковую ткань стиха, творя музыку речи. Эта потенциальная сила во многом проявляется через нелингвистические характеристики звуков, одной из которых является их цветовая окраска. В результате эмоциональная и цветовая экспрессия поэтического текста потенциально значительно выше прозаического.

Трактовка воздействия одного и того же цвета на человека в различных источниках и у разных авторов выглядит весьма разнообразно и противоречиво. Но более глубокий анализ говорит о следующем. В зависимости от ситуации и целей авторы отражают близкую им часть проблемы с определенной эмоциональной оценкой. В самом деле, лечение цветом, влияние цвета на физическом, психическом, метафизическом уровнях, ассоциативное воздействие, музыка и цвет – все эти факторы делают актуальной ту или иную сторону многогранного колористического восприятия мира. Для учебного книгоиздания должны быть проведены свои исследования, после чего выработаны критерии воздействия цвета на учащегося. По результатам исследований различных авторов представляется целесообразным использовать следующие исходные интегральные психологические значения цвета.

О чем говорит красный цвет? Он вызывает наиболее сильную физиологическую реакцию. Эмоции в данном случае могут быть, как положительные, так и отрицательные, в любом случае реакция на красный цвет – реакция возбуждения. Этот цвет по психологическому воздействию беспо-

койный, при длительном воздействии слишком долгое возбуждение переходит в раздраженное утомление.

Желтый цвет воздействует легко и возбуждающе, привлекает внимание и сигнализирует о чем-то новом. Иногда этот цвет может вызвать раздражение и утомляет. Вам не приходилось встречать человека, из которого так и брызжет оптимизм, так что через некоторое время вы его видеть не можете? То же самое происходит и с желтым цветом. Если в зеленом преобладает желтый, то он возбуждает, а если синий – то успокаивает. Зеленый – это прочность, надежность, долговечность, благополучие и уверенность.

В окружении синего вы чувствуете себя в гармонично-ненапряженном состоянии. Синий цвет считается деловым, профессиональным и авторитетным. Однако если синий излишне доминирует, то он может быть подавляющим и даже депрессивным.

Фиолетовый цвет подчеркивает незаурядность и оригинальность. Сочетает в себе энергию красного и элегантность синего.

Звуко-цветовой анализ этих характеристик с использованием фактора различного психолингвистического воздействия цвета на человека показывает высокую степень совпадения их с содержанием материала. Значимость имеющихся и полученных результатов психолингвистического анализа дают все основания полагать, что они найдут применение в сфере учебного книгоиздания. Созданный инструментарий, по сути дела, может послужить действенным и эффективным технологическим средством реализации издательских норм и рекомендаций. Кроме того, возможен учет этой технологии в издании литературы для обучаемых с различного рода девиациями в психофизиологическом развитии.

Обобщая изложенное, можно высказать следующее общее прогностическое предположение. Применение современных высоких технологий создает возможности для эффективного использования скрытых потенциальных возможностей текста (скрытой информативности и экспрессивности языка), что крайне важно в системе учебного книгоиздания и уменьшения степени субъективности экспертиз учебной литературы.

Разработанный инструментарий (математическая модель) психолингвистического анализа текста (его первая версия) уже сегодня позволяет решать ряд следующих задач системы учебного книгоиздания.

- 1) тестирование авторского коллектива на предмет психолингвистической совместимости;
- 2) определение психолингвистических характеристик авторских текстов и выдача рекомендаций по их корректировке;
- 3) проведение редактирования текста в направлении максимального приближения его к стандартной статистической структуре языка изложения с целью создания естественной языковой среды для учащегося, как наиболее оптимальной.