УДК 159.95

ИССЛЕДОВАНИЕ ОЦЕНКИ ПОНИМАНИЯ НАРРАТИВНЫХ И ЭКСПОЗИТОРНЫХ ТЕКСТОВ С ПРИМЕНЕНИЕМ ЛАТЕНТНОГО СЕМАНТИЧЕСКОГО АНАЛИЗА

С.В. Курицин, В.М. Воронин (Екатеринбург)

Аннотация. Предлагается использование латентного семантического анализа (LSA) в качестве метода компьютерной оценки понимания текста. Проводится сравнительное исследование оценок понимания нарративных и экспозиторных текстов, полученных с помощью LSA и экспертов.

Ключевые слова: латентный семантический анализ; понимание текста; компьютерное тестирование; когнитивное моделирование; обработка дискурса; нарративный текст; экспозиторный текст.

Обучение пониманию текстов является одной из самых важных задач, поставленных перед современным образованием, которое направлено на развитие способностей мышления, выработку практических навыков, изучение процедур и технологий, формирование базовых компетенций. Умение адекватно воспринимать, осмыслять и в результате понимать прочитанное является важнейшим компонентом образовательного процесса.

Проблема оценивания результатов понимания до настоящего времени является малоизученной [1]. Это положение отмечается и в Концепции образовательной области «Филология» (2000): «До сих пор не выработаны научно обоснованные критерии оценки знаний, умений и навыков учащихся…» Несмотря на появление систем оценивания, основанных на критериях правильного выполнения текстового задания (ЕГЭ и др.), массовую школьную практику оценивания результатов понимания текста можно охарактеризовать как рутинную.

Компьютерное тестирование знаний становится все более актуальной и широко распространенной технологией оценки качества знаний обучающихся. Наряду с такими достоинствами, как относительная простота технической реализации, высокая степень автоматизации и минимизация затрат времени на проведение процедуры тестирования, опыт практического использования этой технологии позволяет говорить о следующих проблемах:

– в большинстве из широко распространенных компьютерных систем тестирования используются вопросы, основанные на прямом сравнении ответа с заранее заданным вариантом правильного ответа. Такие тесты подходят для проверки фактологических знаний и понимания концептуальных связей в предметной области, косвенной проверки практических навыков решения задач в определенной предметной области. При этом недоступны для оценивания дискурсивные аспекты знания, связанные со способностью тестируемого практически демонстрировать свои знания и умения в рассуждениях, дискуссиях, ответах на вопросы собеседников;

– практически невозможно проводить автоматическое тестирование творческих способностей обучающихся, например в рамках гуманитарных специальностей;

– наличие правильного варианта ответа на вопрос не исключает возможность простого угадывания или нахождения правильного ответа по принципу исключения.

Решение отмеченных проблем в настоящее время связывается с компьютерной лингвистикой и технологиями искусственного интеллекта.

На наш взгляд, продуктивным является привлечение метода латентно-семантического анализа (LSA) [6, 8] для преодоления ограничений тестового контроля по выборочному методу. Этот метод позволяет извлекать контекстно-зависимые значения слов при помощи статистической обработки больших наборов текстовых данных, и в его основу заложены принципы анализа главных компонентов, применяемого в создании искусственных нейронных сетей. Совокупность всех контекстов, в которых встречается и не встречается данное слово, задает множество обоюдных ограничений, которые дают возможность определить похожесть смысловых значений слов и множеств слов. Он позволяет моделировать отдельные когнитивные и психолингвистические процессы у человека, и реализация его возможна на современных персональных компьютерах.

Для решения поставленной задачи было необходимо разработать компьютерную программу – имплементацию алгоритма LSA и протестировать ее для русского языка, а также создать необходимый для работы программы корпус русского языка, репрезентирующий общие знания учащихся старших классов.

Экспериментальная апробация LSA в качестве инструмента оценки понимания текстов проводилась нами на примере свободно конструируемых развернутых ответов на понимание текстов двух видов – нарративного и экспозиторного.

Метод LSA (первоначально известный как Латентное семантическое индексирование (Latent Senatic Indexing, LSI) разрабатывался для решения задач поиска и извлечения информации (information retrieval), которое представляет собой выделение из большой базы данных документов небольшого количества документов, релевантных заданному запросу. Предшествующие подходы к решению этой задачи включали в себя поиск по ключевым словам (keyword-matching), удельный вес этих ключевых слов и векторную основу, изображающую наличие слов в документах. LSA распространил векторную основу на декомпозицию на сингулярные значения (Singular Value Decomposition (SVD)) [4] перестройки базы.

Хотя существуют некоторые вариации, но общий алгоритм работы LSA таков [7]:

1. Сбор большого массива (релевантного) текста и разделение его на «документы». В большинстве случаев каждый параграф обрабатывается как отдельный документ. Такой подход основан на том, что информация внутри параграфа имеет тенденцию быть логически связанной (когерентной) и последовательной.

2. Следующий этап – создание смежной матрицы документов и термов. Клетка в этой матрице соответствует документу *x* и терму *y* и содержит количество раз, которое *y* встречается в *x*. Терм определяется как слово, которое встречается более чем в одном документе, и при морфологическом поиске или другом морфологическом анализе не представляет собой попытки комбинировать различные формы того же слова.

Если есть *m* термов и *n* документов, то такая матрица может быть рассмотрена как репрезентация, в которой существует *m*-мерный вектор для каждого документа и *n*-мерный вектор для каждого терма.

3. Значение каждой клетки может быть уменьшено за счет эффекта от общих слов, которые встречаются во всем корпусе (т.е. наиболее часто встречаемых слов в общем массиве текстовой информации). Метод общего взвешивания – «логарифм энтропии» – базируется на теории информации (Information Theory), в которой значение повышается при получении информации.

4. SVD способствует осуществлению при помощи параметра *k*, который точно определяет желаемое число измерений. (В принципе SVD рассчитывается со всеми измерениями и создает три матрицы, которые при перемножении дают исходные данные, но соответствующее количество памяти, которое требуется для такой операции, фактически не реально. Вместо этого используют алгоритмы, оптимизированные для разреженного пространства данных и подсчитывают только наиболее значимые *k*-измерения матриц.) Результат описанного выше процесса – три матрицы. Одна имеет *k*-мерный вектор для каждого документа, другая – *k*-мерный вектор для каждого терма в корпусе, и в третьей – *k*-сингулярные значения. Первые две матрицы определяют два различных векторных пространства, которые также отличаются от пространства, определяемого исходной матрицей.

LSA начинает процесс изучения с определения частоты встречаемости слов в контекстах («документах»). LSA прочитывает текст в цифровой форме и определяет, когда встречаются слова в каждом сегменте теста и с какой частотой. Если слова соответствуют когнитивным единицам, то необходимо определить каждое слово как очень длинный вектор, содержащий вектор количества раз, когда слово появилось в каждом параграфе или документе. Но известно, что данное решение неудовлетворительно: причина, по которой пропозициональные репрезентации занимают первое место в исследованиях процессов понимания, такова, что слова не могут являться аналогами когнитивных единиц. Итак, вместо прямого определения слов в термах документов (и документов в термах слов) LSA заменяет семантическое приближение, что радикально уменьшает измерение пространства. Это делается с помощью хорошо известной техники декомпозиции матрицы на сингулярные значения. Теорема, взятая из алгебры, гласит, что любая квадратная матрица **M** может быть разложена на три матрицы:

$$\mathbf{M = A*D*A'},$$

где **A** и **A'** матрицы составляют собственный вектор (eigenvector) матрицы и **D** – диагональная матрица с собственными значениями (eigenvalues) (или сингулярными значениями) матрицы. В LSA нас интересуют не квадратные матрицы, а теорема, выведенная для неквадратных (non-square) матриц. Собственное значение последовательно в термах их величины или важности. Умножение трех матриц приводит к возврату к исходной матрице. LSA отбрасывает большинство собственных значений (и связанных с ним собственных векторов) и сохраняет только наибольшие, предъявляя 300 самых больших. Перемножение трех матриц, таким образом, уменьшает и не воспроизводит точно **M**, но приближает к оригинальной **M**. Таким образом, это является значительным преимуществом. Исходная матрица также содержит множество информации о всех деталях и случаях употребления слова. Путем отбрасывания всех этих деталей мы сохраняем только сущность значения каждого слова, это чистая семантическая структура, индифферентная к специфичным ситуациям. При таком конструировании семантического пространства, как правило в 300 измерений, каждое слово и документ исходной матрицы могут быть представлены как вектор. Более того, новые слова и документы могут быть вставлены в это пространство и просчитаны с любым исходным вектором. Есть различные способы сравнения векторов; рассмотрим в качестве примера один из них – он наиболее тесно связан с корреляцией: измерение связанности (relatedness) между векторами – нахождение косинуса между векторами в многомерном семантическом пространстве. Одинаковые векторы имеют косинус, равный 1, ортогональные векторы имеют косинус, равный 0, и противоположные векторы имеют косинус, равный –1. Например, «дерево» и «деревья» имеют cos = 0,85; «дерево» и «кошка» существенно независимы, cos = –0,01; «кошка» и «собака, преследующая кошку» – cos = 0,36 (по данным Kintsch [6]).

Макроединицы текста могут также быть представлены как векторы в LSA-пространстве. Действительно, как только текст был проанализирован по составляющим его словам и пропозициям, вектор, репрезентирующий текст в целом, – просто центроид составляющих векторов. Таким образом, макроструктура текста свободно создается на основе только что созданной микроструктуры текста (учитывая, что соответствующие мак-

роединицы ясно обозначены в тексте). Следовательно, макроединицы могут так же принимать участие в процессе активации знания, как слова или пропозиции.

Как отмечалось выше, для апробации метода LSA в качестве инструментария оценки понимания было необходимо создать корпус русского языка и программу – имплементацию алгоритма LSA. Русский корпус LSA, созданный для этого исследования, содержит документы, отражающие базовые знания среднестатистического российского школьника к 11-му классу. Он включает литературу, обязательную к прочтению в рамках школьной программы, учебники по всем школьным дисциплинам до 11-го класса, научно-популярные, художественные и фантастические произведения, энциклопедии, газетные статьи, сценарии фильмов, стенограммы специализированных интернет-форумов и т.д., т.е. ту информацию, которая, на наш взгляд, в достаточно большой степени репрезентирует базовые знания учеников 11-х классов. В целом корпус состоит из 71 267 документов (т.е. параграфов), включает 4 661 954 различных термина (без применения стемминга, т.е. без возврата словам их исходных форм). Размерность корпуса была определена опытным путем в 337 измерений.

Все вычисления мер LSA производились в специально разработанной программе, использующей имплементацию алгоритма LSA. Программа была написана на языке C для операционной системы Microsoft Windows и оптимизирована для параллельных и распределенных вычислений для процессоров компании Intel.

Методы исследования

В этом исследовании использовался LSA как процедура оценки семантического подобия между пересказом и исходным текстом. Для целей этого исследования способность LSA моделировать человеческие суждения о пересказе была проверена шестью различными способами. В литературе [2, 5, 9] различаются холистический (holistic) (H) и компонентный, или аналитический, методы (componential) (C). Принципиальные различия, которые существуют между холистическими и компонентными методами, основываются на том, как они оценивают пересказ. В то время как холистические методы обеспечивают полную оценку, основываясь на подобии пересказу исходного (глобального) текста, компонентные методы обеспечивают оценку, вычисляя подобие между множественными компонентами пересказа (например, между предложениями, когерентностью, дополнительными или главными темами).

Согласно Foltz и др. [2], каждый подход имеет свои преимущества. В то время как холистический метод может типично обеспечивать более точную меру полной итоговой качественной оценки, компонентный метод оценки может обеспечить более детализированные данные о том, какие компоненты пересказа были оценены лучше. В этом исследовании были выбраны шесть различных методов – четыре холистических и два компонентных.

Метод H1. *Пересказ – исходный текст*. Этот холистический метод заключается в сравнении пересказа каждого испытуемого с исходным текстом мерой косинуса между ними. Так, чем выше косинус между пересказом и текстом, тем лучше будет качество пересказа. Этот метод был применен E. Kintsch и др. [5] для задачи оценки пересказа в программе Summary Street.

Метод H2. *Пересказ – пересказ*. Этот метод заключается в анализе пересказов, написанных обучающимися, для установления подобия среди всех них. Каждому пересказу присваивается среднее значение его косинуса по сравнению с остальными пересказами, что означает, что пересказ, наиболее подобный остальной части пересказов (т.е. со всеми пересказами, данными обучающимися), получил бы самую высокую оценку, второй по мере подобия пересказ получил бы вторую наивысшую оценку и т.д. Landauer, Laham и др. [9] использовали подобный метод, но применяли вместо ранговой системы оценки матрицу расстояний (1-косинус). Матрица расстояний между всеми пересказами была «развернута» к единственному измерению (континууму), которое лучше всего восстанавливало все расстояния, и точка, соответствующая конкретному пересказу, в этом измерении бралась как мера его качества.

Метод H3. *Пересказ – экспертный пересказ*. Третий холистический метод заключается в оценке пересказов обучающихся путем сравнения их с эталонным пересказом, написанным экспертами-оценщиками. Таким образом, пересказ обучающегося, наиболее подобный экспертному пересказу, оценивается как наилучший, обладающий высшим качеством. В настоящем исследовании четыре пересказа, написанных экспертами-оценщиками, были выбраны как стандарт и оценка каждого пересказа обучающегося была вычислена как его косинус LSA со стандартом. Подобный метод использовался Landauer, Laham и др. [9] для пересказов обучающихся.

Метод H4. *Предградуированный пересказ – неградуированный пересказ*. Этот заключительный холистический метод заключается в предварительной оценке набора пересказов обучающихся экспертами-оценщиками. Набор пересказов сначала градуируется экспертами-оценщиками, затем вычисляется косинус между каждым неградуированным и каждым предградуированным пересказом, а каждому новому пересказу присваивается среднее значение косинусов небольшого набора (10) близко подобных пересказов. Главная сила этого метода – то, что он рассматривает человеческие суждения (оценки экспертов-оценщиков) как начальную фазу. Этот метод был применен Landauer, Laham и др. [9] для пересказов обучающихся.

Метод C1. *Пересказ – предложения исходного текста*. Этот компонентный метод заключается в исследо-

вании подобия пересказа обучающегося и каждого предложения в исходном тексте, который был прочитан. Вычисленный косинус, таким образом, среднее значение косинуса между пересказом обучающегося и всеми предложениями исходного текста.

Метод С2. *Пересказ – главное предложение исходного текста*. Этот последний компонентный метод очень схож с предыдущим. Он состоит в вычислении и усреднении косинусов между каждым предложением в пересказе обучающегося и рядом предложений исходного текста, которые эксперты посчитали наиболее важными. Этот метод был применен P. Foltz, D. Laham и T. Landauer [3] для пересказов обучающихся.

Характеристика выборки

В качестве испытуемых выступали 22 учащихся 11-х классов среднеобразовательной школы, а в качестве экспертов – 4 преподавателя (в том числе один кандидат психологических наук).

Процедура исследования

Задача испытуемых заключалась в передаче содержания текстов (нарративного и экспозиторного). Пересказы писались учениками от руки и затем перекодировались в электронную форму. Эксперты-оценщики оценивали пересказы по двум шкалам: по содержанию (от 0 до 4 баллов) и по когерентности (от 0 до 6 баллов).

В качестве нарративного текста использовался русский перевод рассказа «Circle Island» («Остров Круга») [10]. Этот рассказ состоит из 170 слов и требует определенного фонового знания для понимания. Данный текст был выбран в силу того, что является типичным нарративным текстом, поскольку изначально создавался для демонстрации пропозициональной схемы нарративного текста и психологического объяснения понимания такого текста.

В качестве экспозиторного текста была взята статья из энциклопедии, адаптированной для общих навыков чтения всех испытуемых, о деревьях в джунглях. Статья содержала 500 слов и также требовала предшествующего общего знания. Текст был выбран в качестве типичного экспозиторного текста, поскольку в нем присутствуют концептуализация знаний, причинно-следственные связи, специфическая терминология. К тому же текст аналогичен тем текстам, которые ученики изучают на уроках биологии в 11-м классе.

Результаты

Полученные данные прошли три ступени анализа. Во-первых, для каждого типа текста была оценена надежность Interrater оценок экспертов-оценщиков (согласованность оценок экспертов-оценщиков) и проведен корреляционный анализ мер косинуса LSA, полученных с помощью вышеупомянутых шести методов с оценками экспертов-оценщиков для каждого типа текста и каждого оцениваемого компонента (т.е. содержания и когерентности). Во-вторых, было произведено сравнение полученных корреляций с целью оценить относительную надежность методов для каждого типа текста и каждого компонента, используя дисперсионный анализ ANOVA (методы к тексту, к оценке). В-третьих, был проведен регрессионный анализ для оценки независимой пропорции различий оценок экспертов-оценщиков, объясняемой каждым методом.

Тест надежности Interrater оценок экспертов-оценщиков

Перед анализом того, является ли LSA надежным инструментом в оценке пересказа, было необходимо проверить надежность оценок экспертов-оценщиков. Для нарративного текста корреляции общих оценок варьировались от 0,79 до 0,84 (коэффициент Пирсона). Эти данные использовались как основание для сравнения корреляции между оценками экспертов-оценщиков и LSA. Тест надежности Interrater содержания варьировался от 0,81 до 0,86, а когерентности – от 0,66 до 0,75. Для экспозиторного текста надежность общих оценок экспертов-оценщиков варьировалась от 0,64 до 0,82 (коэффициент Пирсона). Тест надежности Interrater содержания варьировался от 0,53 до 0,81, а когерентности – от 0,58 до 0,79.

Анализ корреляций между косинусом LSA и оценками экспертов-оценщиков

В нарративном тексте корреляции между оценками LSA и оценками экспертов-оценщиков были просчитаны для каждого метода (табл. 1). Все корреляции были положительны и статистически значимы ($p < 0{,}001$). Для шести методов все корреляции между оценками экспертов-оценщиков и косинусами LSA были подобны, таким образом, все методы работали в сходной манере. В частности, для нарративного текста холистические методы сопоставимы с компонентными методами. Обнаруженные корреляции сопоставимы с обнаруженными в [5] для текстов о древних цивилизациях.

Корреляции между оценками LSA и оценками экспертов-оценщиков для экспозиторного текста показаны в табл. 2. Для первых пяти методов все корреляции были положительны и статистически значимы ($p < 0{,}01$). Метод *Пересказ – главное предложение исходного текста* (С2) показывает статистически незначимую корреляцию между оценкой третьего эксперта-оценщика и косинусом LSA. Для экспозиторного текста эти шесть методов не работают подобным образом, как для нарративного текста. При симулировании LSA оценок экспертов-оцен-

Таблица 1

Корреляционная матрица оценок пересказов LSA и экспертов-оценщиков для нарративного текста

Метод	Эксперт 1	Эксперт 2	Эксперт 3	Эксперт 4
H1 – *пересказ – исходный текст*	0,55***	0,54***	0,60***	0,47***
H2 – *пересказ – пересказ*	0,54***	0,55***	0,57***	0,49***
H3 – *пересказ – экспертный пересказ*	0,52***	0,52***	0,53***	0,46***
H4 – *предградуированный пересказ – неградуированный пересказ*	0,57***	0,50***	0,53***	0,50***
C1 – *пересказ – предложения исходного текста*	0,58***	0,56***	0,60***	0,50***
C2 – *пересказ – главное предложение исходного текста*	0,57***	0,55***	0,59***	0,48***

***p < 0,001.

Таблица 2

Корреляционная матрица оценок пересказов LSA и экспертов-оценщиков для экспозиторного текста

Метод	Эксперт 1	Эксперт 2	Эксперт 3	Эксперт 4
H1 – *пересказ – исходный текст*	0,40***	0,33***	0,37***	0,40***
H2 – *пересказ – пересказ*	0,42***	0,42***	0,31***	0,48***
H3 – *пересказ – экспертный пересказ*	0,56***	0,52***	0,41***	0,61***
H4 – *предградуированный пересказ – неградуированный пересказ*	0,52***	0,57***	0,48***	0,63***
C1 – *пересказ – предложения исходного текста*	0,27***	0,22**	0,22**	0,27***
C2 – *пересказ – главное предложение исходного текста*	0,21**	0,17*	0,14	0,22**

*p < 0,05; **p < 0,01; ***p < 0,001.

щиков некоторые методы более надежны по сравнению с другими. В целом во всех изученных случаях для экспозиторного текста холистические методы были более надежны, чем компонентные.

Дисперсионный анализ корреляционных данных

Чтобы сравнивать все корреляционные данные и сделать выводы о них, был выполнен дисперсионный анализ ANOVA (6 (Методы) на 2 (тип текста, нарративный и экспозиторный) на 2 (вид оценки, содержание и когерентность).

Результаты ANOVA показывают, что оценка, основанная на семантическом подобии, хорошо соответствовала человеческой оценке. Семантическое подобие исходит из сравнения пересказа с исходным текстом, с пересказом, сделанным экспертами, с частью пересказа, или, в случае компонентных методов, с предложениями исходного текста. Таким образом, семантические отношения становятся важным индикатором при оценке общего качества пересказа.

Регрессионный анализ

Для определения того, что выбранные методы оценивают сходным образом, и того, что они измеряют независимо, необходимо было выполнить регрессионный анализ, в котором была бы оценена та пропорция разницы суждений экспертов, которую каждый объясняет независимо.

Восемь пошаговых регрессионных моделей были выполнены на данных для каждого типа оценки того, как различные методы объясняют независимую пропорцию разницы в оценках содержания и оценках когерентности экспертов-оценщиков. Четыре пошаговые регрессионные модели были выполнены для прогнозирования содержания (одна для каждого эксперта-оценщика) и четыре модели – для прогнозирования когерентности. Независимые переменные были шестью методами, используемыми в данном исследовании.

Результаты показали устойчивый паттерн для двух типов текста. Метод H4 (*предградуированный пересказ – неградуированный пересказ*) присутствовал во всех регрессионных моделях. Кроме того, все 16 регрессионных моделей, за исключением одной, содержали другой метод в заключительном уравнении модели. Регрессионный анализ показал, что при комбинировании двух наиболее успешных методов процент объясняемой разницы для экспозиторных текстов может достичь процента объясняемой разницы для нарративных текстов. Или, другими словами, предсказание человеческих суждений при комбинировании несколько методов улучшилось для экспозиторных текстов, чем для нарративных.

Общие выводы

Результаты показали, что существуют различия в способе, которым методы ведут себя относительно нарративных и экспозиторных текстов. Таким образом, надежность LSA была выше для нарративного текста, чем

для экспозиторного, с подобием между оценками экспертов-оценщиков и косинусами LSA, являющимися больше для содержания, чем для когерентности.

Следующие выводы относятся к вопросу об относительной надежности различных методов, основанных на LSA, для вычисления качества пересказа. Во-первых, сравнение всех методов показало, что они все ведут себя одинаково хорошо для нарративных текстов, с корреляциями, подобными найденным Kintsch и др. [5]. Однако компонентные методы эксплицитно хуже выполнялись для экспозиторных текстов. Мы можем выразить это также, говоря, что методы, которые используют информацию исходного текста, чтобы оценить пересказ, были хуже. Методы H4 (*предградуированный пересказ – неградуированный пересказ*), H3 (*пересказ – экспертный пересказ*), H2 (*пересказ – пересказ*) были значительно лучше. Эти три метода используют для конечной оценки только информацию, содержащуюся в пересказе. Три худших метода используют информацию, основанную на исходном тексте. Кроме того, LSA в нарративном тексте коррелирует больше с оценками содержания экспертов-оценщиков, чем с оценками когерентности экспертов-оценщиков. Однако для экспозиторного текста мы обнаружили противоположные результаты. Эти различия могли появиться из-за того, как LSA оценивает содержание в нарративном тексте. Оценки когерентности и содержания экспозиторного текста и оценки когерентности нарративного текста фактически одинаковы. Фактически различия между когерентностью и содержанием не были статистически значимы.

Если эти данные показали, что холистические методы были более надежны, чем компонентные методы, то это также подтверждает идею о том, что LSA может обеспечить более точное измерение общего качества пересказа в противоположность предоставлению более определенной информации относительно того, какие компоненты пересказа могут быть оценены лучше. Эта точка зрения также предполагает, что LSA более чувствителен к оценке того, как семантическая информация обрабатывается в терминах концептуализации и абстракции, а не к оценке с помощью компонентных методов.

Результаты показали, что метод H4 (*предградуированный пересказ – неградуированный пересказ*), дополненный методом H3 (*пересказ – экспертный пересказ*), может использоваться для лучшего прогнозирования пропорции разницы в человеческих суждениях.

Литература

1. *Шаповал С.А.* Понимание текстов как результат решения учебных филологических задач: Автореф. дис. … канд. психол. наук. М., 2006.
2. *Foltz W., Gilliam S., Kendall S.* Supporting content-based feedback in on-line writing evaluation with LSA // Interactive Learning Environments. 2000. № 8. P. 111–128.
3. *Foltz P., Laham D., Landauer T.* Automated Essay Scoring: Applications to Educational Technology // Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 1999. Chesapeake, VA: AACE, 1999. P. 939–944.
4. *Golub G.H., Luk F.T., Overton M.L.* A block Lanczoz method for computing the singular values and corresponding singular vectors of a matrix // ACM Transactions on Mathematical Software. 1981. № 7. P. 149–169.
5. *Kintsch E., Steinhart D., Stahl G.* Developing Summarization Skills through the Use of LSA-Based Feedback // Interactive Learning Environments. 2000. № 8(2). P. 87–109.
6. *Kintsch W.* Comprehension: A paradigm for cognition. N.Y.: Cambridge University Press, 1998.
7. *Landauer T.K., Dumais S.T.* A solution to Plato's problem: the Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge // Psychological Review. 1997. № 104.
8. *Landauer T.K., Foltz P.W., Laham D.* Introduction to Latent Semantic Analysis // Discourse Processes. 1998. № 25. P. 259–284.
9. *Landauer T.K., Laham D., Rehder B., Schreiner M.E.* How well can passage meaning be derived without using word order? A comparison of Latent Semantic Analysis and humans // Proceedings of the 19th annual meeting of the Cognitive Science Society. Mahwah, NJ: Erlbaum, 1997. P. 412–417.
10. *Thorndyke P.W.* Cognitive structures in comprehension and memory of narrative discourse // Cognitive Psychology. 1977. № 9. P. 11–110.

APPLICATION OF LATENT SEMANTIC ANALYSIS FOR ASSESSMENT OF COMPREHENSION OF NARRATIVE AND EXPOSITORY TEXTS

Kuritsin S.V., Voronin V.M. (Ekaterinburg)

Summary. The paper describes an application of latent semantic analysis (LSA) as a method for computer assessment of text comprehension. Comparative research of assessments of comprehension of narrative and expository texts for LSA and experts is conducted.

Key words: latent semantic analysis; text comprehension; computer assessment; cognitive modeling; discourse processing; narrative text; expository text.