

НЕСТАНДАРТНЫЙ ПОЭТИЧЕСКИЙ ТЕКСТ И ЕГО МОДЕЛИРОВАНИЕ С ПОМОЩЬЮ КОМПЬЮТЕРА

Поэтический текст представляет собой структуру большой сложности, гибкий и сложно устроенный художественный механизм. На поэтический текст накладываются дополнительные по отношению к языку ограничения: требование соблюдать определенные метроритмические нормы, организованность на фонологическом, лексическом и идейно-композиционном уровнях. Все это делает поэтический текст значительно более «несвободным», чем прозаический текст или обычная разговорная речь, и превращает его в сложную иерархическую и динамическую систему, обладающую достаточно высокой степенью формализации. Как известно, существуют стихотворные тексты, в которых нет рифмы. Нерифмованные стихотворения очень разнообразны. Одной из разновидностей стихов без рифмы являются белые стихи. Белый стих — это название метрических (стопных) стихов. Структурную роль рифмы здесь играют конечные слоги стихотворной строки, начиная с последнего ударного слога, а также ритмические окончания стихотворной строки. Для белого стиха характерна бедная строфика, так как строфическое разнообразие в стопном стихе определяется разнообразной системой рифмовки. Однако отсутствие рифмы не лишает белый стих поэтических достоинств. Основные компоненты стиха — ритм, образность языка, размер и др. в белом стихе сохраняются. Тем не менее, его можно считать нестандартным поэтическим текстом.

С целью более глубокого изучения высокоинтеллектуальной речемыслительной деятельности человека, а так же законов стихообразования процесс порождения стихотворений можно определенным образом формализовать. В настоящей статье предлагается вариант формальной модели процесса порождения англоязычного нестандартного поэтического текста, который может стать основой для моделирования творческой деятельности человека. Материалом для исследования послужили семнадцать 5-строчных стихотворений американского поэта Майкла Макклинтока [1] и двенадцать 7-строчных стихотворений разных англоязычных авторов [2].

В результате анализа статического содержания всех стихотворных текстов компьютером с опорой на один из статистических методов были выделены главные опорные слова, задающие их тему. Второстепенные слова не выделялись, так как размер текстов небольшой. Например, основное содержание представленного ниже 5-строчного нестандартного поэтического текста отражено в табл. 1:

*Carrying the sun, the clouds,
The mountains easily –
A small stream
Wandering unnamed
In this wild place.*

Таблица 1

Основное содержание нестандартного 5-строчного поэтического текста

№ п/п	Тип опорного слова	Опорное слово
1	ГОС1	<i>sun</i>
2	ГОС2	<i>clouds</i>
3	ГОС3	<i>mountains</i>
4	ГОС4	<i>stream</i>
5	ГОС5	<i>place</i>

Исходя из значений опорных слов текста, можно сделать вывод о том, что тема этого текста — природа.

Отметим, что необходимо учитывать два аспекта темы текста: статический и динамический. Статический аспект текста представлен формальным образом с помощью главных и второстепенных опорных слов, которые несут основную семантическую нагрузку в тексте, обеспечивая становление и развитие его темы. Поскольку порождение текста носит процессуальный характер, динамический аспект содержания текста может быть описан в виде логико-семантических формул. Но так как исследованные англоязычные нестандартные поэтические тексты достаточно малы по размеру (5–7 строк), логико-семантические формулы, отражающие динамический аспект развития их темы, не выделялись.

Формальное представление семантико-синтаксической организации темы осуществлялось с помощью искусственного языка СЕМСИНТ. Для более точного отражения в семантико-синтаксической формуле каждого предложения элементов конкретной ситуации были разработаны довольно объемные семантические подклассы слов знаменательных и служебных частей речи. Рассмотренное выше англоязычное 5-строчное стихотворение на языке СЕМСИНТ будет выглядеть следующим образом:

AAD<S0001.1+L0005+N0006/N0006*+L0005+N0006.1/N0006.1*+L0005+N0006.1/N0006.1+D0002*>AS2<L0006+J0003+N0006/N0006N0006*S0001.1+S0007.2*>ALK<K0001+P0002+J0005+N0015/N0015*>,

где AAD — семантическая функция, указывающая на способ; S0001.1 — причастие, обозначающее движение; L0005 — скреп, указывающий на определенность, известность; N0006 — имя существительное, обозначающее природу; N0006.1 — имя существительное во множественном числе, обозначающее природу; D0002 — наречие образа действия; AS2 — семантическая функция, обозначающая понятие; L0006 — скреп, указывающий на что-то новое; J0003 — имя прилагательное, обозначающее размер; S0007.2 — причастие, обозначающее состояние; ALK — семантическая функция, обозначающая место; K0001 — предлог, указывающий на местонахождение предмета; P0002 — указательное местоимение; J0005 — имя прилагательное, обозначающее состояние; N0015 — имя существительное, указывающее на место.

Списки главных опорных слов, семантико-синтаксические формулы, а также словарь, представляющий собой совокупность семантических

подклассов лексических единиц всех стихотворений, составили лингвистическую базу знаний системы. Лингвистическая база знаний послужила основой для построения формальной модели системы автоматического порождения англоязычных нестандартных поэтических текстов. Предложенная модель является вероятностно-алгоритмической, так как базируется на вероятностном выборе семантико-синтаксических формул и на определенных лексических, синтаксических и семантических правилах. На рис.1 представлен принципиальный алгоритм порождения англоязычного нестандартного поэтического текста.

Рассмотрим основные особенности данного алгоритма. Работа по порождению текста начинается с вывода на экран компьютера списков лексических единиц, входящих в таблицы основного статического содержания проанализированных поэтических текстов. Пользователь выбирает один из списков, то есть задает компьютеру тему порождаемого стихотворения — набор главных опорных слов, которые отражают главных субъектов, действия, место и время ситуаций, описываемых в будущем тексте (блок 1). Предложим, что из всех предложенных тем будущего стихотворения пользователь выбрал тему, заданную последовательностью главных опорных слов *mausoleum, willows, brooch, hair*. Далее начинается работа по организации и структурированию содержания порождаемого текста. Код каждой темы разворачивается в семантико-синтаксическую формулу строфы, случайным образом выбираемую компьютером из лингвистической базы знаний (блок 2). Например, для темы под кодом T001 будет выбрана следующая семантико-синтаксическая структура:

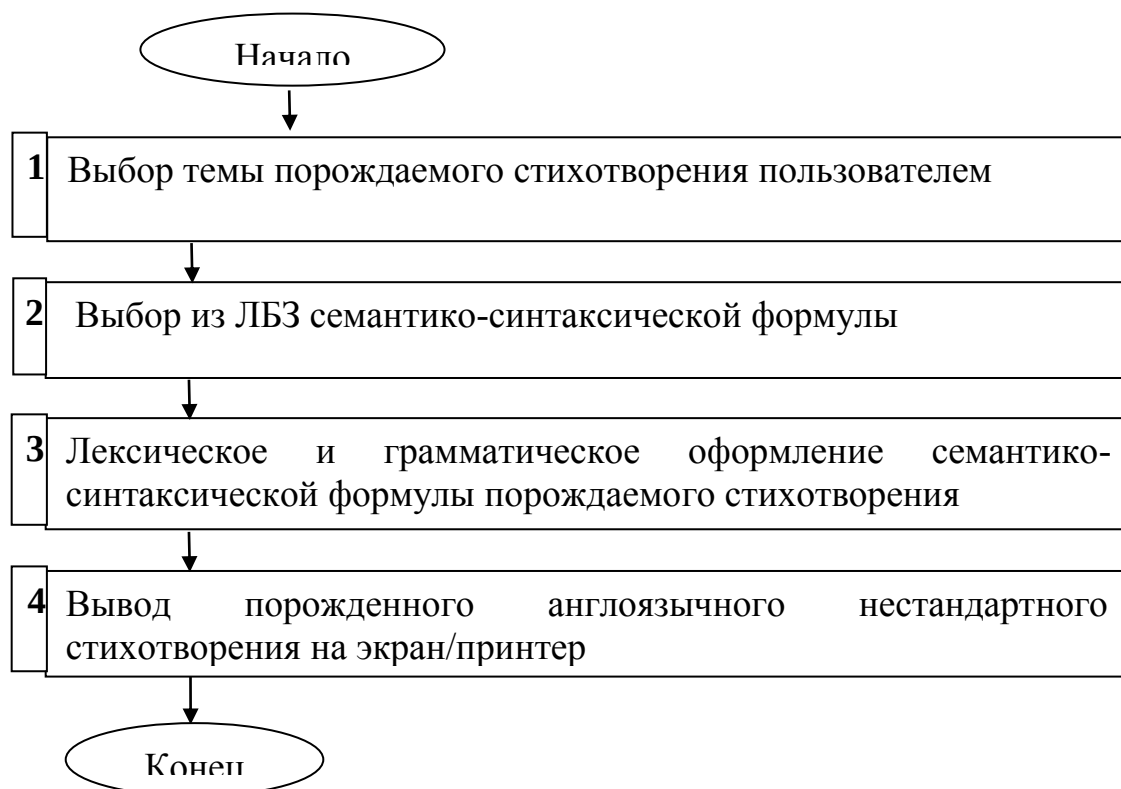


Рис. 1. Принципиальный алгоритм порождения англоязычного нестандартного поэтического текста

AO<L0005+N0015/N0015*+L0002+S0003.1+N0003.1/N0003.1*>ALK<D0001+L0005+J0001+N0004/N0004*>AAD<D0002*+P0001+R2<V0003/11.1>+K0001+P0001*>AD1<P0001.1+R2<V0006/7.2>+K0005+N0005/N0005*>

Лексическое наполнение каждой семантико-синтаксической формулы порождаемого текста (блок 3) проводится с учетом следующих условий:

1) вероятностных, связанных со случайным выбором имен существительных, имен прилагательных, глаголов и наречий из списка лексических единиц, относящихся к определенному семантическому подклассу;

2) грамматическое оформление предложений осуществляется по строго детерминированным правилам, опирающимся на синтаксические значения семантических функций, лежащих в основе семантико-синтаксического языка СЕМСИНТ, а также на коды грамматических форм имен существительных, глаголов, имен прилагательных, наречий, представленных в перечнях семантических подклассов лексических единиц англоязычных нестандартных поэтических текстов. Например, опираясь на представленную выше семантико-синтаксическую формулу (тема T001), компьютер может породить следующий текст нестандартного стихотворения:

*The library and laughing cedars
Inside the young brooch
Really it gathers on you
They are made of face*

Результатом реализации блока 4 является вывод порожденного англоязычного нестандартного поэтического текста на экран компьютера или принтер.

Априори можно предположить, что в порождаемых компьютером текстах возможны семантические несоответствия между некоторыми лексическими единицами в конкретно взятом контексте. Это можно объяснить тем, что вариативность отдельных лексических единиц, заданная в перечнях семантических подклассов, в каком-то предложении может быть недопустима, например, при использовании глагола с предлогом, с которым он не употребляется. Для устранения таких ошибок необходимо усовершенствовать систему порождения путем включения в лингвистическую базу знаний таблиц семантической сочетаемости лексических единиц, а в процесс порождения текста компьютером — точные правила их выбора.

ЛИТЕРАТУРА

1. McClintock, M. Meals at Midnight / M. McClintock. – Balti-more: MD, 2008.
2. [Электронный ресурс]. – Режим доступа: <http://sem44.narod.ru/uitmen/uitmen03.htm>. – Дата доступа: 25.03.2013.