

Употребление А. Блоком различных слов с недостаточно конкретным, обобщенным или неопределенным значением, несомненно, связано с крайней абстрактностью, недосказанностью, иррациональностью взглядов самого поэта в то время.

Таким образом, в школьном изучении лирики Блока педагогу необходимо уделять особое внимание вопросам мастерства поэта, рассмотрению характерных черт его поэтики. Такой подход формирует у учащихся интерес к художнику слова, обеспечивает более глубокое постижение его лирики, эстетически грамотный анализ его стихов.

Литература

1. Блок А. А. Полное собрание сочинений и писем: В 20 т. / РАН. Ин-т рус. лит. (Пушкин. дом). М., 1997–. Т. 1: Стихотворения. Книга первая (1898 – 1904). 1997.
2. Кожин А. Н. Лексический повтор в стихотворных текстах А. Блока // Образное слово Александра Блока. М., 1980. С. 56–89.
3. Некрасова Е. А. Виды экспрессивной тональности в стихотворных произведениях Александра Блока // Образное слово Александра Блока. М., 1980. С. 89–122.
4. Чуковский К. И. Александр Блок // Блок А. А. Стихотворения и поэмы. М., 1978. С. 3–26.

ЭМОТИВНАЯ КЛАССИФИКАЦИЯ ПРЯМОЙ РЕЧИ ПО СЛОВАМ АВТОРА В РУССКИХ И БЕЛОРУССКИХ ТЕКСТАХ

Ю. С. Бородина

Синтез речи находит применение в различных сферах, например, при создании автоответчиков, естественно-языковых интерфейсов, озвучивании информационных сообщений в транспорте и т. д. Кроме того, технологии синтеза речи могут использоваться для машинного создания аудиокниг. Обычно это требует серьезной работы дикторов и актеров, но с помощью существующих наработок в сфере синтеза речи процесс может быть автоматизирован.

Разработанный в лаборатории распознавания и синтеза речи Объединенного института проблем информатики НАН Беларуси синтезатор может «читать книги» [1, с. 269]. Тем не менее, процесс качественного синтеза речи еще не закончен, и остается несколько существенных проблем, которые до сих пор не были решены, и одна из них – это проблема интонации.

В пьесах авторы упрощают работу актеров с помощью ремарок, которые подсказывают необходимую интонацию. В прозе при прямой речи присутствуют слова автора, способные выполнять эту же функцию. Так, приведенный ниже текст, если он будет озвучен актером в аудиокниге,

никогда не будет прочитан с нейтральной интонацией: эмоции говорящего видны через знаки препинания, суть самой фразы, и, не в последнюю очередь, через «подсказки» в словах автора.

– При чем тут вековые устои?! – неожиданно крикнула Лида. – Я не маленькая. И все знаю. Якира арестовали, и вы подло бросили жену с ребенком. Вы... Вы... Вы – неинтересный человек!

Так, для качественного создания аудиокниг с помощью синтезатора речи нужно учитывать эмоции, заложенные в репликах героев и словах автора.

В сфере искусственного интеллекта последние 20 лет активно велись исследования по синтезу и распознаванию эмоций в голосе [2]. Работки в этой области достигли первых успехов: был разработан Emotional Markup Language [3], на рынке и в открытом доступе появились первые эмоциональные синтезаторы [4; 5]. Наличие этих разработок позволяет перенести задачу эмотивного синтеза речи из фантастической плоскости в реальную.

Данная статья знакомит с первичными результатами работы, цель которой заключалась в том, чтобы найти в фразах диалогов идентификаторы эмотивной полярности (позитивный, негативный, нейтральный) и предложить возможные средства их формализации, а также протестировать их на небольшом рабочем корпусе для дальнейшего синтеза эмоций при автоматическом создании аудиокниг. Отметим, что на данном этапе работы не ставится задача описания и распознавания эмоций при сарказме, иронии, шутке и др.

Исследование ведется одновременно для белорусского и русского языков. Для белорусского языка в качестве материала были выбраны повести О. Рукаля «На службе князя Радзівіла», В. Короткевича «Дзікае паляванне караля Стаха» и «Чорны замак Альшанскі». Для русского языка материалом послужили повесть С. Довлатова «Компромисс» и роман В. Пикуля «На задворках Великой Империи». При помощи грамматик по извлечению из текста прямой речи вместе со словами автора для обоих языков были составлены конкордансы объемом в 250 реплик.

Для каждой из реплик русского и белорусского языков была указана эмотивная полярность, идентификатор, по которому она определяется и рекомендации для создания алгоритма в программе NooJ.

Например, для фразы: *– Ну вось, – зямянін дабрадушна ўсьміхнуўся скрозь вусы, – нарэшце пазналі адзін аднаго. Давай абдымемся* эмотивная полярность была определена как позитивная по словам-идентификаторам *добрадушно* и *улыбнулся*, которые в свою очередь являются наречием и глаголом. Предложение для алгоритма описывается как ADVERB_pos + VERB_pos и представляет собой комбинацию слов

из списков ADVERB_pos (позитивно тональные наречия) и VERB_pos (позитивно тональные глаголы).

После обработки всего материала выяснилось, что есть два основных способа создания позитивной или негативной эмотивной тональности в словах автора: 1) маркированные глаголы, которые сами по себе создают необходимую тональность, например, *улыбнулся, процедил, выдавил*; 2) употребление нейтральных по тональности глаголов с эмотивно маркированными наречиями (*беззлобно, счастливо, сумна, сцюдзена, суха, добродушно*) или сочетаниями существительного с предлогом (*с улыбкой, с трудом*), деепричастными оборотами (*тяжело дыша, улыбаясь*) и другими обстоятельствами образа действия, описывающими способ произнесения реплики (*ненавидящим шепотом, исполненной муки голосом*). Также встречаются случаи комбинирования первого и второго способов.

На основании вышеописанных способов были составлены следующие списки идентификаторов: VERBS_neg, VERBS_pos, VERBS_neut, ADVERBS_neg, ADVERBS_pos, NOUN_neg, NOUN_pos, COLLOCATION_neg, COLLOCATION_pos,.

Чтобы протестировать способность компьютера по спискам идентификаторов определять эмотивную тональность прямой речи со словами автора, были созданы тестовые синтаксические грамматики на базе ранее разработанных в лаборатории для определения прямой речи с помощью международной лингвистической программы NooJ [6, с. 29; 7]. Грамматики содержат в себе описание структуры слов автора, набор знаков препинания, которые формально описывают прямую речь и слова автора в ней, а также сами списки идентификаторов, т. е. слов, которые должны встретиться в тексте в окружении знаков препинания, присущих словам автора. На рис. 1 можно видеть фрагмент одной из грамматик.

Грамматики для русского и белорусского языков позитивной и негативной и нейтральной эмотивной полярности были применены к материалам, и программа нашла все определенные экспертом предложения, например: – *Спортсмен будет, – улыбается главный врач Михкель Теппе; – Довлатов, – исполненным муки голосом произнес Туронок, – Довлатов, я вас уволю; – Что? – В ужасе закричали воины* и т. д.

Таким образом, были собраны коллекции идентификаторов эмотивной полярности текста для русского и белорусского языков, вынесены рекомендации по их формализации и разработаны тестовые грамматики по нахождению в корпусе эмотивно полярных фраз прямой речи со словами автора.

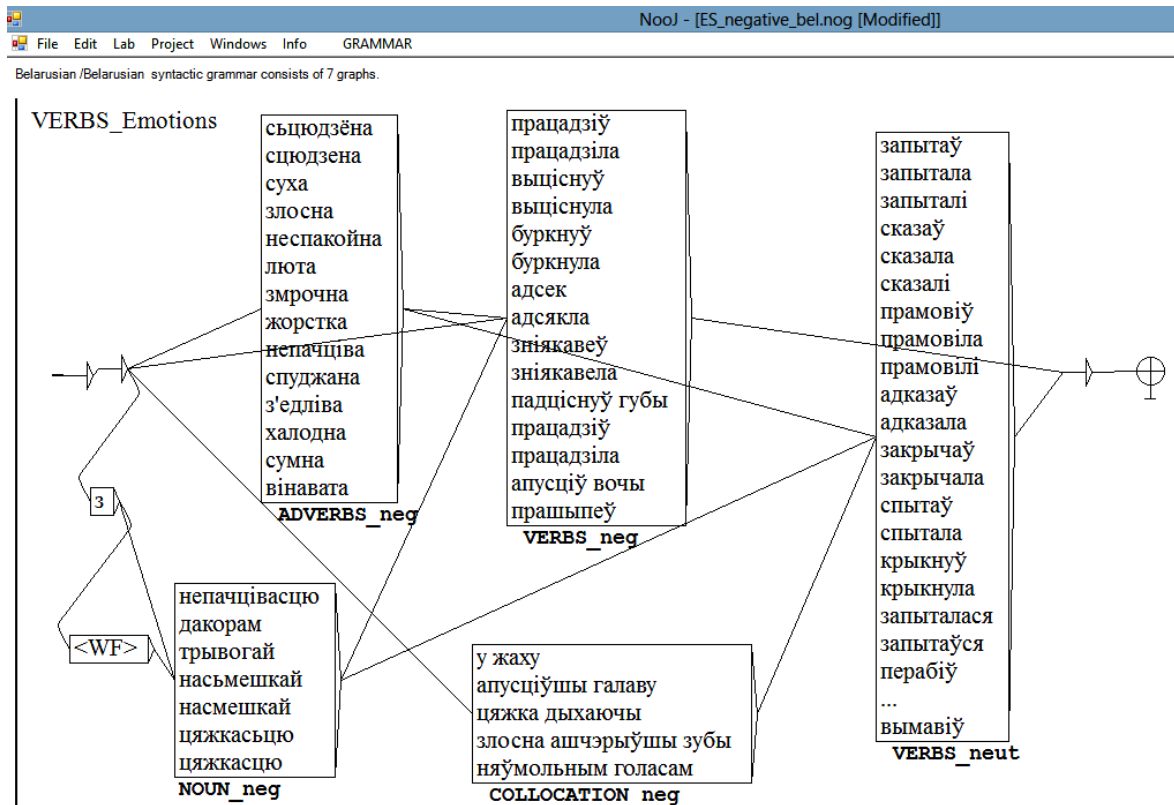


Рис. 1. Вид подграфа VERBS_Emotions для негативной тональности белорусского языка

Тема требует дальнейшей разработки, поэтому будут увеличены текстовые корпуса и, соответственно, списки идентификаторов эмоций. Кроме того, определение эмоций будет вестись не только по словам автора, но и по прямой речи.

Литература

1. Гецэвіч Ю. С., Окрут Т. І., Лабанаў Б. М. Аўтаматызацыя шматгаласавога стварэння аўдыёкніг на беларускай мове з дапамогай сінтэзатараў маўлення па тэксце // Развитие информатизации и государственной системы научно-технической информации (РИНТИ-2013): Доклады XII Международной конференции (Минск, 20 ноября 2013 г.). Минск, 2013. С. 269–276.
2. Schröder M. Emotional Speech Synthesis – A Review // Proc. EUROSPEECH 2001 Scandinavia, 7th European Conference on Speech Communication and Technology, 2nd INTERSPEECH Event, Aalborg, Denmark, September 3–7, 2001. Vol. 1. P. 561–564.
3. Schröder M., Baggia P., Burkhardt F., Pelachaud C., Peter C., Zovato E. EmotionML – an upcoming standard for representing emotions and related states // Proc. Affective Computing and Intelligent Interaction. Memphis, 2001. P. 316–325.
4. Aylett M., Pidcock C. The CereVoice Characterful Speech Synthesiser SDK // Intelligent Virtual Agents, 2007. P. 413–414.
5. Burkhardt F. Emofilt: the Simulation of Emotional Speech by Prosody-Transformation // Interspeech'2005 – Eurospeech, 9th European Conference on Speech Communication and Technology, Lisbon, Portugal, September 4–8, 2005. P. 509–512.

6. *Hetsevich Y.* Overview of Belarusian And Russian dictionaries and their adaptation for NooJ //Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the NooJ 2011 Intern. Conf. / Eds. K. Vučković, B. Bekavac, M Silberztein. Newcastle, 2012. P. 29–40.
7. Лингвистический процессор NooJ. Интернет-адрес: <http://nooj4nlp.net/pages/nooj.html>.

БІБЛІЯ Ф. СКАРЫНЫ: ЦАРКОЎНАСЛАВЯНСКАЯ ТРАДЫЦЫЯ І ГІСТОРЫЯ МОЎНАГА ФЕНОМЕНУ

К. П. Варонька

Рэлігійная пісьменнасць для кожнага нацыянальнага кантэксту з'яўляецца вельмі істотным і важным паказчыкам агульнай распрацаванасці і дасканаласці моўнай сістэмы. Існаванне рэлігійных тэкстаў адначасова ў межах некалькіх моўных структур часта вызначае сінкрэтычны характар перакладных ці адаптаваных варыянтаў Старога і Новага Запаветаў, а таксама іншых тэкстаў, што выкарыстоўваюцца падчас богаслужэння. Неканвенцыянальнасць знака, дыгლოსная моўная сітуацыя на пачатку XVI ст. стваралі значныя перашкоды для ўзнікнення Бібліі на старабеларускай мове. Тым больш, што корпус біблейскіх кніг таксама не быў жорстка акрэслены.

Царкоўнаславянская мова і царкоўнаславянскія ў прыватнасці вельмі доўгі час адназначна маркіравалі рэлігійны дыскурс. Гэта найперш было звязана з тым, што трывалая нарматыўная літаратурная форма была лепшай асновай для дэмакратычнага варыянту рэлігійнага тэксту, які вызначаўся б ідыялектным устанаўленнем адпаведнасцей паміж старажытнаславянскай мовай, беларускай народнай гаворкай, чэшскімі і польскімі формамі. Ф. Скарына і яго паслядоўнікі, найперш С. Будны і В. Цяпінскі, засвоілі царкоўнаславянскую традыцыю, але мова іх выданняў у рознай ступені вызначаецца дэмакратычнасцю формы. Граматычныя і лексічныя трансфармацыі, якія ўзніклі ў выданнях Ф. Скарыны, часткова ці цалкам з'яўляюцца ў выданнях яго паслядоўнікаў: прасочваецца ўжо новая традыцыя ці нават норма.

Мова рэлігійнай пісьменнасці, яе ўзаемаадносінны з літаратурнай нацыянальнай нормай, з кананічнай традыцыяй – найбольш агульныя пытанні, якія з'яўляліся прадметам даследавання ў многіх лінгвістычных працах. Адзначым, што большасць з іх адрозніваецца аспектным падыходам, які грунтуецца на асобным моўным феномене ці абмежаваным у часе культурна-гістарычным этапе. Цікавасць і важнасць найноўшых даследаванняў заключаецца ў вызначэнні алгарытму трансфармацыі лексікі царкоўнаславянскага паходжання ў рэлігійных тэкстах, пачынаючы з выданняў Ф. Скарыны і да сённяшніх перакладаў.