

С. О. Савчук
Институт русского языка им. В. В. Виноградова РАН
(Россия, Москва)
savsvetlana@mail.ru

МУЛЬТИМЕДИЙНЫЙ ПОЭТИЧЕСКИЙ КОРПУС: ОБЩАЯ СТРУКТУРА И ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ¹

Научное изучение звучащей стихотворной речи имеет в России давнюю традицию. В начале XX в. много внимания уделяли разработке теории этой проблемы представители ОПОЯЗа и ученые, близкие к кругу формалистов (В. Б. Шкловский, Б. М. Эйхенбаум, В. М. Жирмунский, Б. В. Томашевский, Ю. Н. Тынянов, С. И. Бернштейн, Л. П. Якубинский и др.). Звучащий стих в 1920-е годы стал объектом междисциплинарных исследований в Институте живого слова. С 1923 г. основная работа была сосредоточена в Комиссии по изучению художественной речи (КИХР) при Институте истории искусств. В рамках работы комиссии под руководством С. И. Бернштейна собрана уникальная коллекция записей чтения ведущими поэтами, известными чтецами-декламаторами и актерами, начато использование этих записей для изучения особенностей поэтических текстов и приемов авторской декламации. Подробнее об истории КИХРа и собранной коллекции см., например, [Звучащая художественная речь 2018], [Золотухин, Шмидт 2014]. Однако в силу объективных причин лингвистическим исследованиям в области теории звучащего стиха суждено было остаться на периферии филологической науки.

В настоящее время интерес к изучению звучащей стихотворной речи вновь усилился [Князева 2012], [Живое слово 2012]. В связи с этим проект создания мультимедийного поэтического корпуса, которому посвящен доклад, представляется очень своевременным. Его можно рассматривать как новый инструмент для исследования звучащего поэтического слова. Идея включения стихотворного материала в мультимедийный модуль в составе НКРЯ была предложена Е. А. Гришиной при разработке Мультимедийного параллельного корпуса (МультиПАРКа) [E. Grishina et al. 2012]. В чем отличие создаваемого современного корпуса от коллекций фонодокументов, обычно используемых для инструментального анализа поэтической речи?

Материал. Корпус будет включать не только аудиозаписи, но и видеозаписи выступлений, что стало возможно благодаря современному уровню развития техники. Это позволит проводить анализ не только звучащей стороны речи, но и анализировать видеоряд – жесты, мимику, позы, которые соотносятся со звучащей стихотворной речью и являются составной частью исполнения (представления).

Состав корпуса. По составу авторов корпус предполагается сделать максимально открытым и ограниченным объемом мультимедийного материала. На первом этапе мы предполагаем использовать открытые источники,

¹ Работа выполнена при поддержке программы Президиума РАН «Памятники материальной и духовной культуры в современной информационной среде».

представленные в сети Интернет, в дальнейшем надеемся привлечь архивные материалы из Гостелерадиофонда, Госфильмофонда, музейных коллекций; при сборе современного материала – пополнять собственные видео- и аудиотеки. По типу исполнителей в корпус будут включены записи трех видов: 1) авторское чтение стихов, 2) профессиональное чтение, исполненное мастерами художественного слова, к которым мы относим актеров, чтецов, дикторов радио и телевидения, 3) непрофессиональное чтение, исполненное людьми разных возрастов и профессий, которые не обучались специальным приемам декламации. В пилотный проект корпуса, представленный в настоящем докладе, вошли произведения А. А. Ахматовой, А. С. Пушкина, С. А. Есенина в исполнении разных чтецов.

Организация материала и поисковый интерфейс. Материал в корпусе организован на принципах выравнивания клипов и текстов, которые были реализованы в МультиПАРКе. Одной канонической версии текста стихотворения, опубликованной в авторитетных изданиях, сопоставлены разные исполнительские версии стихотворения в аудио- и видеозаписях. Каждая запись разделена на клипы продолжительностью от 5 до 20 сек.; соответствующая текстовая расшифровка, а также каноническая версия разделена на фрагменты (в среднем соответствующие одной строфе); клипы и тексты выровнены между собой. Таким образом, по запросу на какое-либо слово или конструкцию будет выдаваться содержащий эту единицу фрагмент текста в нескольких вариантах исполнения. При этом каждый текстовый фрагмент будет сопровождаться видео- или аудиоклипом, который можно прослушать и сохранить с целью дальнейшего анализа с помощью внешних программ.

Возможности корпуса. Такая организация материала и поисковый интерфейс дают возможность сопоставительного изучения стихотворного текста. Возможность сравнивать письменный текст с его звучанием, сопоставлять разные интерпретации звучащего стиха представляет интерес для изучения фонетических и просодических явлений, открывает большие перспективы для стиховедческих исследований, исследований поэтического синтаксиса и семантики и др.

ЛИТЕРАТУРА

Живое слово: логос – голос – движение – жест: Сборник статей и материалов / Сост., отв. ред. В. В. Фещенко. М., 2015.

Звучащая художественная речь: Работы Кабинета изучения художественной речи (1923–1930) / Сост.: В. Золотухин, В. Шмидт. М., 2018.

Золотухин В., Шмидт В. «Хрипели в КИХРе голоса...»: Опыт реконструкции исследований звучащей художественной речи в Государственном институте истории искусств (1920–1930) // НЛЮ. 2014. № 5. С. 236–253.

Князева Е. М. Авторский орфоэпический выбор как языковой фактор актуализации ритма в звучащем стихотворном тексте // Лингвистика и

методика преподавания иностранных языков и культур. 2012. Вып. 4. (электронное издание). С. 100–115.

E. Grishina, S. Savchuk, D. Sichinava. Multimodal Parallel Russian Corpus (MultiPARC): Main Tasks and General Structure // LREC 2012 Workshop on Best Practices for Speech Corpora in Linguistic Research. Istanbul, Turkey, May 21, 2012. ELRA, pp. 13–16.