

КОРПУСНЫЕ ТЕХНОЛОГИИ ПРИ АНАЛИЗЕ КИНОДИСКУРСА (НА МАТЕРИАЛЕ АВТОРСКОГО ЭЛЕКТРОННОГО КОРПУСА КИНО)

Шамова Н. А.

*Ивановский государственный университет, Иваново
E-mail: sh941@mail.ru*

В XXI веке происходит научно-технический прогресс и компьютеризация общества, в результате которой активно внедряются разнообразные технологии в лингвистические и лексикографические исследования. Все большую популярность завоевывают корпусные технологии. Их цель – оказать помощь в отборе, систематизации и анализе большого количества информации, благодаря наличию специальных функций. В последнее время наблюдается тенденция создания узких корпусов по различным областям знаний. В данной статье поднимается вопрос об эффективности применения корпусных технологии для анализа кинодискурса, а именно авторского электронного корпуса «cinema», созданного на платформе Sketch Engine.

Ключевые слова: *корпусные технологии, кинодискурс, Sketch Engine, авторский корпус.*

ВВЕДЕНИЕ

Корпусные технологии завоевывают все большую популярность в XXI веке среди отечественных и зарубежных ученых в связи с развитием информационных технологий и активным их использованием в научных исследованиях. Многие деятели науки подчеркивают большой потенциал и эффективность применения корпусных технологий, направленных на решение различных актуальных языковых проблем и задач. Особое признание среди корпусных технологий получают разнообразные базы данных, программы-конкордансы и платформы, основанные на корпусных данных (например, *DANTE, TextStat, AntConc, SKELL, Sketch Engine* и другие). В связи с активным развитием кинематографа и небольшим количеством научных работ по современному кинодискурсу нам представляется необходимым предложить новый эффективный способ анализа языка кинодискурса на основе авторского корпуса «cinema», сгенерированного на платформе Sketch Engine.

Применение авторского корпуса для анализа кинодискурса является актуальным исследованием, поскольку благодаря этому можно изучить особенности данного вида дискурса. Цель нашего исследования – разработка и краткое описание авторского электронного корпуса кино, а также рассмотрение назначения каждой функции отдельно с приведением наглядных примеров из корпуса.

ОСНОВНАЯ ЧАСТЬ

Обратимся к понятию *кинодискурс* и его составляющим. М. Динел проводит разграничение понятий «film discourse» и «cinematic discourse», утверждая, что они обозначают разное. По мнению ученого, последнее объединяет в себе набор кинематографических техник, которые изучаются главным образом за пределами лингвистики. Отсюда следует, что лингвистика изучает «film discourse». На наш взгляд, разграничение двух понятий, предложенное ученым, достаточно размыто и неясно [3].

Подчеркнем, что кинодискурс представляет собой особый вид дискурса, имеющий свои уникальные особенности и характерные черты. Г. Г. Слышкин, М. А. Ефремова справедливо отмечают, что языковая составляющая кинодискурса представлена знаками-символами, которые могут быть письменными (например, титры) или устными (например, звучащая речь актеров) [2]. В данной статье будут рассмотрены письменные знаки-символы в соответствии с приведенным выше делением, потому что в основу корпуса лег печатный журнал *American Cinematographer*, и, следовательно, нас будут интересовать тексты определенного периодического издания.

Корпусные технологии имеют массу преимуществ относительно ранее используемых ручных методов анализа информации. Кроме того, корпуса можно быстро пополнить новой информацией в случае необходимости. Благодаря грамотной обработке данных в корпусе, можно получить: контексты употребления слов и выражений, устойчивые словосочетания, наиболее частотные слова, а также едва уловимую разницу в семантике, известную лишь носителям языка.

С. А. Маник справедливо отмечает, что «корпусная лингвистика обеспечивает объективное представление языка, поскольку анализ корпусных сведений позволяет изучить практически любые языковые модели и структуры (лексические, грамматические, лексико-грамматические структуры, дискурсные модели, фонологические и морфологические обороты)» [1]. Мы разделяем позицию ученого, действительно, корпуса способствуют систематизации информации при изучении языка, тем более это важно, если речь идет об иностранном языке. Из вышеперечисленных С. А. Маник моделей и структур нас будут интересовать дискурсные модели в предметной области киноиндустрия.

Выбор *Sketch Engine* для создания собственного корпуса обусловлен тем, что данная платформа предоставляет возможность зарегистрированным пользователям создать корпус абсолютно бесплатно, а также использовать доступные опции для обработки информации. Программа обладает разнообразными функциями, а именно поиском, сортировкой и анализом

данных (*concordance*, *word list*, *word sketch*, *thesaurus*, *sketch-diff*). Следует отметить, что на базе данной платформы активно проводятся разнообразные лингвистические и лексикографические исследования многими учеными. Также это один из немногих ресурсов, который дополнительно дает возможность получить статистические данные. Представляется возможным далее описать преимущества и возможности корпуса.

Для создания корпуса «cinema» нами было выбрано периодическое издание *American Cinematographer*. Данный журнал представляет большое количество разнообразных текстов, в которых содержатся реальные примеры языкового использования. Источником создания корпуса послужило 34 выпуска данного журнала (2013–2015 годов). Отметим, что авторский корпус содержит 1051528 слов.

Рассмотрим назначение каждой функции отдельно и приведем наглядные примеры из авторского электронного корпуса «cinema»:

1) *Конкорданс (concordance)* позволяет увидеть окружение интересующего слова или фразы (контекст). Примерами данной опции могут служить следующие данные: «*gaffer*» («...he says he learned a great deal from gaffers, directors and other collaborators...»), «*epic*» («...made super 8 films, pixilated two minutes epics in which we would...»), «*filmmaker*» («...a team of camera assistants and other filmmakers who are out there working...»), «*stand-in*» («Tony Cottee talked himself out of a job, then watched his stand-in play a major role in the Everton tragi-comedy»), «*stuntman*» («The lorry driver was real, not a stuntman, because it was thought no one else would have the expertise to drive a long, heavy vehicle at a specific speed»), «*script writer*» («Hollywood script writer could have conjured up a story to compare with the one Michael Galwey has produced and starred in»). Исходя из контекста, нам становится легче установить грамматические особенности употребления слов, понять часть речи, а также выяснить значение (как в случае со словами «*stuntman*» и «*script writer*»). В вышеперечисленных примерах представлены профессии из сферы киноиндустрии. Отметим, что при необходимости можно увидеть более широкий контекст, а также информацию об источнике.

Используя данную функцию, можно получить также объяснение американского рейтинга фильмов благодаря контексту. Например: «*Rated PG-13 for some violence, sequences of sci-fi action, some suggestive content and partial nudity*». «*Rated PG for mild action and some rude humor*». «*Rated NC-17 for graphic disturbing content involving violence and sexuality, and a scene of brutality*». В данных примерах содержится непосредственно объяснение общепринятых сокращений. Наличие специфического рейтинга фильмов относится к особенностям кинодискурса.

Благодаря данной функции, мы можем увидеть, как используются разговорные выражения в контексте. Например, слово «*buzz*», «*Our members returned from the 20th edition of Camerimage buzzing about the many films*

screened in competition». Таким образом, мы выявили устойчивое выражение «*to buzz about a film*». На наш взгляд, знание данных слов важно, поскольку они представляют собой особый пласт лексики, речь социально или профессионально обособленной группы.

«Concordance» позволяет получить коллокации к словам. Например, к слову «*cinema*» будут следующие устойчивые выражения: «*digital cinema*», «*cinema camcorder*», «*pocket cinema camera*», «*cinema workflow*», «*cinema len*», «*cinema color gamut*», «*drive-in cinema*», «*silent cinema*», «*art house cinema*», «*outdoor cinema*», «*book on cinema*», «*cinema footage*», «*open-air cinema*», «*to keep current with cinema*», «*cinema's power*», «*impact of cinema*», «*trick of the cinema*».

К слову «*screen*»: «*scaling the screen*», «*size screen expands from...to*», «*screen version of*», «*image on screen*», «*...inch screen*», «*screen shot*», «*display on the screen*», «*capacitive touch screen*», «*to fade from the screen*», «*blank screen*», «*on the opening screen of*», «*to screen the movie*», «*screen time*», «*to appear on the screen*».

К слову «*film*»: «*film depicts the life of*», «*the film includes scenes from*», «*the film is a screen version of*», «*film is based on*», «*film celebrity*», «*film studio*», «*film script*». Наиболее часто используемыми являются следующие комбинации частей речи: существительное + существительное (noun + noun), глагол + существительное (verb + noun), прилагательное + существительное (adjective + noun). Знание грамматически или лексически устойчивых словосочетаний, в частности из сферы киноиндустрии, формирует естественную речь, поскольку данные языковые комбинации получены из речи носителей языка, а значит, они являются достоверными.

Нам также удалось выделить различные наименования для современных видов съемки. Например, со словом «*shot*» формируются следующие разновидности: «*over-the-shoulder shot*», «*high angle shot*», «*zoom shot*», «*extreme long shot*», «*point-of-view shot*», «*tight shot*», «*low angle shot*». Расширение традиционного списка видов съемки, на наш взгляд, связано с активным использованием современной техники в киноиндустрии.

2) *Список слов (word list)* предоставляет возможность сгруппировать слова по определенным критериям.

Нами были получены различные термины из сферы киноиндустрии: «*close-up*», «*best boy*», «*focus*», «*postproduction*», «*stunt*», «*episode*», «*producer*», «*cameraman*», «*dimmer*», «*viewfinder*», «*filmmaker*», «*aerial shot*», «*filtration*», «*zoom*», «*backup*», «*soundstage*», «*visual effect*», «*cut*», «*colorist*», «*matte*», «*gaffer*», «*blue-screen shot*», «*angle*», «*camera angle*», «*wrap*», «*arc shot*», «*boom*», «*blow-up*», «*lighter*», «*special effect*», «*footage*», «*stabilizer*», «*B-movie*», «*lobby*», «*color-correction*», «*flash*», «*caption*», «*cast*» и т. д. Следует отметить, что появление многих терминов в киноиндустрии также связано с

техническим прогрессом. Например, слово «*camcorder*» обозначает не просто камеру, а маленькую камеру, которая может легко помещаться в одной руке. Слово «*boom*» означает специальное длинное приспособление с микрофоном на конце, чтобы записывать речь актеров, оставаясь одновременно незаметным на экране для зрителей.

Следует акцентировать внимание на многообразии существующих аббревиатур: *SMPTE* (*Society of Motion Picture and Television Engineers*), *SDDS* (*Sony Dynamic Digital Sound*), *PVR* (*Personal Video Recording*), *PGA* (*Producer's Guide of America*), *OAV* (*Original Animated Video*), *VOD* (*Video on Demand*), *IMDB* (*Internet Movie DataBase*), *ICG* (*International Cinematographers Guild*), *HFC* (*Historians Film Committee*), *CGI* (*Computer-Generated Imagery*). Добавим, что в кинодискурсе аббревиатуры чаще используются именно в письменной, а не в устной речи, поскольку употребление аббревиатур на письме значительно экономит место, в то время как использование их в устной речи вызывает непонимание коммуникантов. Данный вывод сделан на основе анализа материала корпуса. Лишь незначительное количество аббревиатур использовалось в диалогах и интервью актеров и режиссеров, в то время как они активно употреблялись при обзоре кинофильмов или при обсуждении достоинств разнообразной кинотехники.

Благодаря корпусу мы выделили большое многообразие жанров фильмов. Проанализировав полученные жанры, мы заметили, что в XXI веке все большую популярность начинают набирать поджанры и гибриды жанров. Например, «*animated cartoon film*», «*caper film*», «*found footage*», «*musical/dance film*», «*biopic*», «*chick flicks*», «*courtroom drama*», «*detective-mystery*», «*disaster film*», «*guy movie*», «*road film*», «*sport film*», «*supernatural film*». Возможно, это обусловлено желанием создателей фильмов привлечь большее внимание зрителей к непривычным формам, а также желанием кинозрителей получить непривычное содержание. Отметим, что в настоящее время есть тенденция включать в название жанра фильма разнообразные технические характеристики. Примером могут служить такие жанры, как: «*short-length film*», «*full-length film*», «*three-dimensional film*», «*two-dimensional film*», «*stereoscopic 3D film*».

3) *Opisanie слов (word sketch)* позволяет увидеть коллокации, то есть наиболее частотно употребляемые словосочетания. Для удобства они выделяются программой серым цветом. Представлены некоторые грамматические категории, среди которых, например, «*object*», «*subject*», «*modifier*», «*prepositional object*» и ряд других. Рассмотрим их употребление на примерах. Существительное «*screen*» имеет следующие отношения. В качестве дополнения («*get*», «*capture*», «*plan*», «*do*», «*track*», «*grab*»), в роли существительного («*show*», «*follow*», «*require*», «*look*», «*have*»), в виде определения («*wide*», «*few*», «*dolly*», «*long*»), в отношении с союзом «и»

(«close up», «handheld», «scene», «camera»), в качестве предложного дополнения («background», «most», «couple», «number», «kind», «lot»). Существительное «camera» имеет следующие отношения. В виде дополнения («point», «use», «set», «face», «focus»), в роли существительного («roll», «follow», «catch», «record», «roll», «capture»), в отношении с союзом «и» («microphone», «equipment», «light», «tripod», «project»). Необходимо отметить, что знание данных связей очень важно при изучении языка. Основываясь на данной информации, можно лучше понять структуру иностранного языка и, следовательно, овладеть языком на более высоком уровне.

4) *Тезаурус (thesaurus)* позволяет получить синонимы и частотность их употребления. Благодаря предоставляемой платформой функции, можно разнообразить речь, используя слова одного синонимического ряда без потери значения. Частотность употребления выражается в числовом эквиваленте, что позволяет более объективно анализировать слова.

К слову «film» программа выдает следующие близкие по значению слова: «picture», «movie», «image», «scene», «production», «work», «project», «source», «cinema», «camerawork», «artwork», «motion picture», «pic», «the flicks». К слову «actor»: «player», «star», «performer», «speaker», «boy», «artist», «guy». Употребление данных слов позволяет сделать речь более выразительной.

5) *Различие слов (sketch-diff)* наглядно показывает существующую разницу или сходство между словами. Рассмотрим данную функцию на примере двух жанров «comedy» и «horror». К слову «comedy» программа выдает следующие наиболее употребимые прилагательные: «divine», «romantic», «musical», «slapstick», «witty», «amusing». К слову «horror»: «incredulous», «unmitigated», «gothic», «abject», «mock». Благодаря данной функции, можно увидеть сходство и различие значений двух слов, а также частотность их употребления.

ВЫВОДЫ

Таким образом, проанализировав полученный материал и подробно рассмотрев предлагаемые платформой «Sketch Engine» функции, мы можем сделать вывод, что созданный нами электронный корпус «cinema» является эффективным средством для анализа и выявления особенностей кинодискурса. Создание узких корпусов и анализ благодаря им различных видов дискурса является, на наш взгляд, перспективным направлением в лингвистике.

Список литературы

1. Маник С. А. Использование языковых корпусов в преподавании английского языка (на материале базы SketchEngine) / С. А. Маник // Образовательный процесс в университете: реалии и совершенствование. Сборник научно-методических статей. – Иваново, 2013. – С. 231–239.
2. Слышкин Г. Г. Кинотекст: опыт лингвокультурологического анализа / Г. Г. Слышкин, М. А. Ефремова. – М.: Водолей, Publishers, 2004. – 153 с.
3. Dynel M. Stranger than fiction? A few methodological notes on linguistic research in film discourse / M. Dynel // Brno Studies in English. – 2011. – Vol. 37, No 1. – P. 41–46.

CORPUS TECHNOLOGIES IN THE CINEMA DISCOURSE ANALYSIS (BASED ON THE AUTHOR'S CORPUS)

Shamova N.A.

In the XXI century there is a scientific and technological progress and a computerization of society. As a result new technologies are applied in linguistic and lexicographical researches. Corpus technologies become more and more popular. Their aim is to help in a selection, systematization and analysis of large amounts of information thanks to special functions. Recently there is a tendency of the creation of narrow corpora in different fields of knowledge. In the article a question about the effectiveness of the usage of corpus technologies in cinema discourse analysis is raised, namely the author's corpus «cinema» that has been developed on the platform «Sketch Engine».

Keywords: corpus technologies, cinema discourse, Sketch Engine, the author's corpus.