

БАШКИРСКИЙ ЯЗЫК И ЛИТЕРАТУРА В УСЛОВИЯХ ГЛОБАЛИЗАЦИИ И ПОЛИЭТНИЧЕСКОЙ СРЕДЫ: ОПЫТ И ПЕРСПЕКТИВЫ

ISBN 978-5-91608-181-7

С. 39-46

https://doi.org/10.31833/conf_ihll2019.090

УДК: 81.322

Е. С. Аплонова

*Французский Национальный Центр Научных Исследований
7, rue Guy Môquet - BP 8, 94801 VILLEJUIF Париж, Франция
E-mail: aplooon@gmail.com*

О МЕТОДАХ ИЗУЧЕНИЯ СТРАТЕГИЙ ПЕРЕДАЧИ ЧУЖОЙ РЕЧИ В БАШКИРСКОМ ЯЗЫКЕ НА МАТЕРИАЛЕ АУДИО КОРПУСА¹

В статье описывается методология изучения конструкций передачи чужой речи в башкирском языке в рамках проекта “Discoursereportinginstorytelling”. Исследование проводится на материале двух устных аннотированных корпусов, размеченных с помощью специального шаблона для кросс-лингвистического изучения конструкций передачи чужой речи. Данные конструкции представляют особый интерес для комплексного подхода к изучению текстов, так как, с одной стороны, являются недостаточно изученными как отдельная лингвистическая конструкция в типологическом понимании, с другой стороны, предполагают сбор и анализ традиционных нарративов, которые являются ценным источником знаний о культуре носителей языков.

Ключевые слова: конструкции передачи чужой речи, корпусная лингвистика, устные корпуса, методология разметки текстов, мультимодальный анализ.

1. Введение

В последние десятилетия значительно возрос интерес к документированию языков, находящихся под угрозой исчезновения. Согласно [Haleetal, 1992], к основным последствиям исчезновения языка относятся потеря традиционных знаний и размывание этнической самобытности. Документация языков обычно сосредоточена на создании грамматик и словарей, тем не менее, носители исчезающих языков часто выражают беспокойство именно по поводу потери своих “традиционных способов говорить” без каких-либо заметных признаков зарождающейся языковой смерти, и задолго до того, как язык может считаться под угрозой исчезновения в соответствии с любым из критериев, установленных лингвистами [ЮНЕСКО 2003]. Мне довелось слышать множество подобных рассуждений от носителей башкирского языка, которые отмечали, что на сегодняшний день мало кто умеет рассказывать традиционные сказки, а те, кто умеет, вряд ли вспомнит сюжет в силу преклонного возраста. Так происходит в результате проникновения современных культурных развлекательных практик в повседневную жизнь (телевидение, радиопередачи), под давлением которых традиционные формы передачи знаний, включая рассказывания сказок, становятся

¹Исследование осуществляется при поддержке European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 758232).

менее необходимыми и, как следствие, забываются. Потеря этой традиции зачастую сопровождается потерей языковых стратегий, с помощью которых строились традиционные нарративы. Иногда лингвистические стратегии, свойственные устной традиции могут стать частью литературного языка. Например, идеофоны – слова со специальными синтаксическими свойствами, которые используются для более яркого повествования, могут проникнуть в литературные жанры как стилистические маркеры, которые призваны передать характер речи некоторых персонажей [Nikitina 2012c]. Однако чаще всего лингвистические стратегии, которые характеризуют старую традицию, не переживают переход к новым текстовым формам.

Одной из главных лингвистических особенностей традиционных нарративов являются конструкции передачи речи персонажей и/или автора. Изучение таких конструкций проводится в рамках проекта “Discoursereportinginstorytelling” (LLAKAN, Франция), членом которого является автор. Мы исследуем связь между конкретными дискурсивными практиками и лингвистическими стратегиями, на которые они опираются, на материале устных корпусов тюркских языков России (башкирский, чувашский) и языков Западной Африки (гиньянга, джоола, догон, лимба, муан, мундаби, тикар и т.д.). Выбор этих двух регионов обусловлен целью сравнить лингвистические стратегии, свойственные традиционным нарративам из географически и исторически не связанных регионов. Ожидается, что сравнение позволит пролить свет на межкультурные изменения в выборе дискурсивных стратегий в традициях с различными практиками рассказывания историй. Башкирские традиционные нарративы, с одной стороны, относительно хорошо документированы², хотя, как правило, недоступны нерусскоязычным исследователям. Одной из целей проекта является обеспечение широкого доступа к таким данным и их интеграция в современные сравнительные лингвистические и антропологические исследования.

Данная статья является кратким описанием методологии проекта “Discoursereportinginstorytelling” в работе с устными корпусами на примере башкирского корпуса. Статья организована следующим образом. Во второй части обсуждаются используемые корпуса башкирского языка, процесс сбора данных и их разметка. Третья часть посвящена конструкциям передачи чужой речи: после краткого типологического обзора обсуждаются предварительные задачи корпусного исследования башкирского материала.

2. Методология

2.1. Корпуса

В рамках нашего проекта, исследование стратегий передачи чужой речи в башкирском языке проводится на материале двух устных корпусов: Устный корпус башкирского языка дер. Рахметово и с. Баимово [Овсянникова et al, 2019] и Корпус традиционных нарративов, собранный автором летом 2019 г. Тексты обоих корпусов синхронизированы со звуком в программе Elan и оглосированы. Используемая система транслитерации представляет собой расширенную латиницу. Соответствия между башкирской орфографией и используемой системой записи доступны на сайте Устного корпуса. Общая продолжительность текстов в двух корпусах составляет около 10 часов. На графиках представлена информация о жанровом составе двух корпусов и их социолингвистические характеристики.

² См. детальный обзор литературы по башкирскому сказковедению в диссертации Хусаиновой Г. Р. [2007].

График 1. Жанровый состав

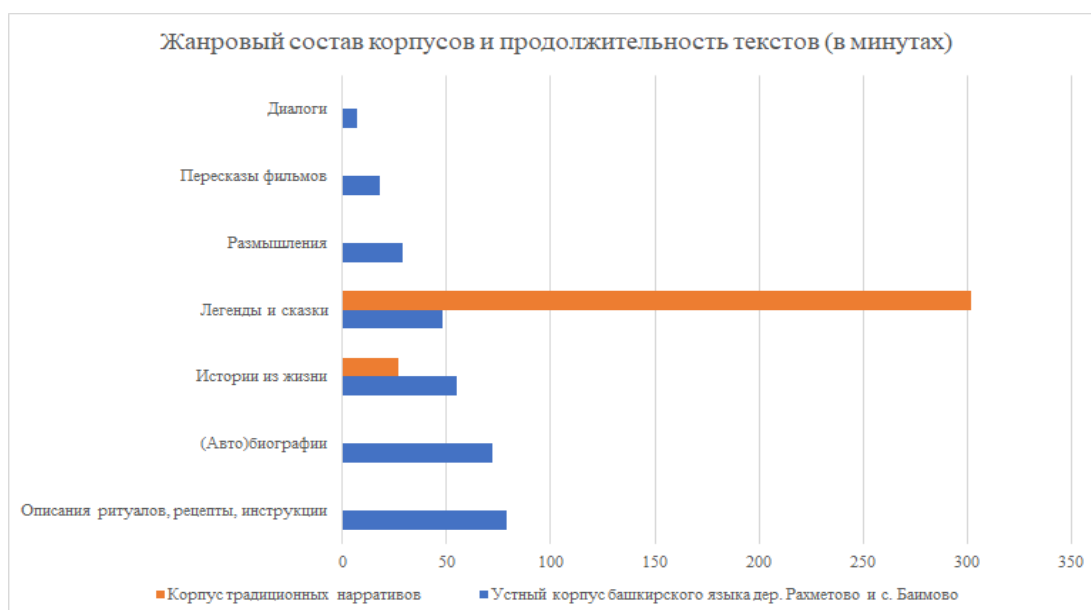
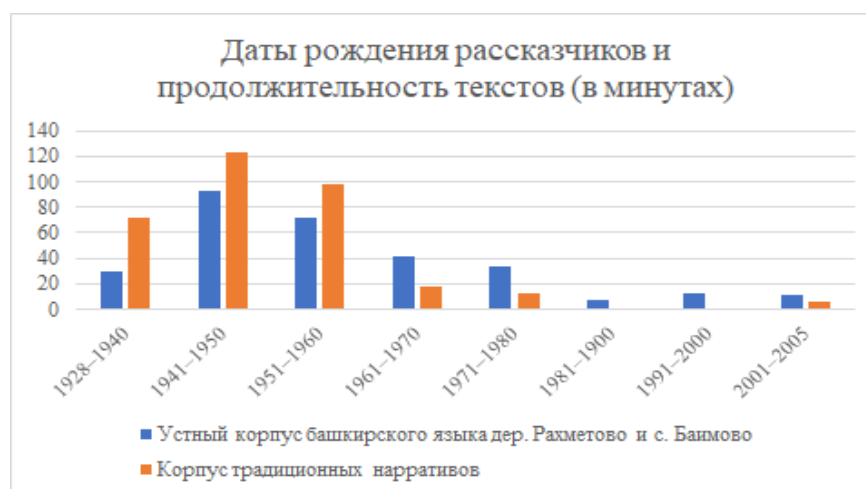


График 2. Социолингвистические характеристики



Как видно из графиков, жанровый состав Корпуса нарративов менее разнообразен и большее количество текстов рассказано людьми старшего поколения. Я намеренно ограничивала рассказчиков в выборе жанра, так как проект “Discoursereportinginstorytelling” фокусируется именно на сказках как на основном жанре материала исследования. Что касается возраста рассказчиков, одной из целей проекта является документация рассказов людей старшего возраста, так как зачастую они являются единственными, кто умеет рассказывать “как раньше”. Однако перспектива сравнить стратегии передачи речи в текстах разных жанров и авторства носителей разных возрастов на материале двух корпусов также кажется нам интересной.

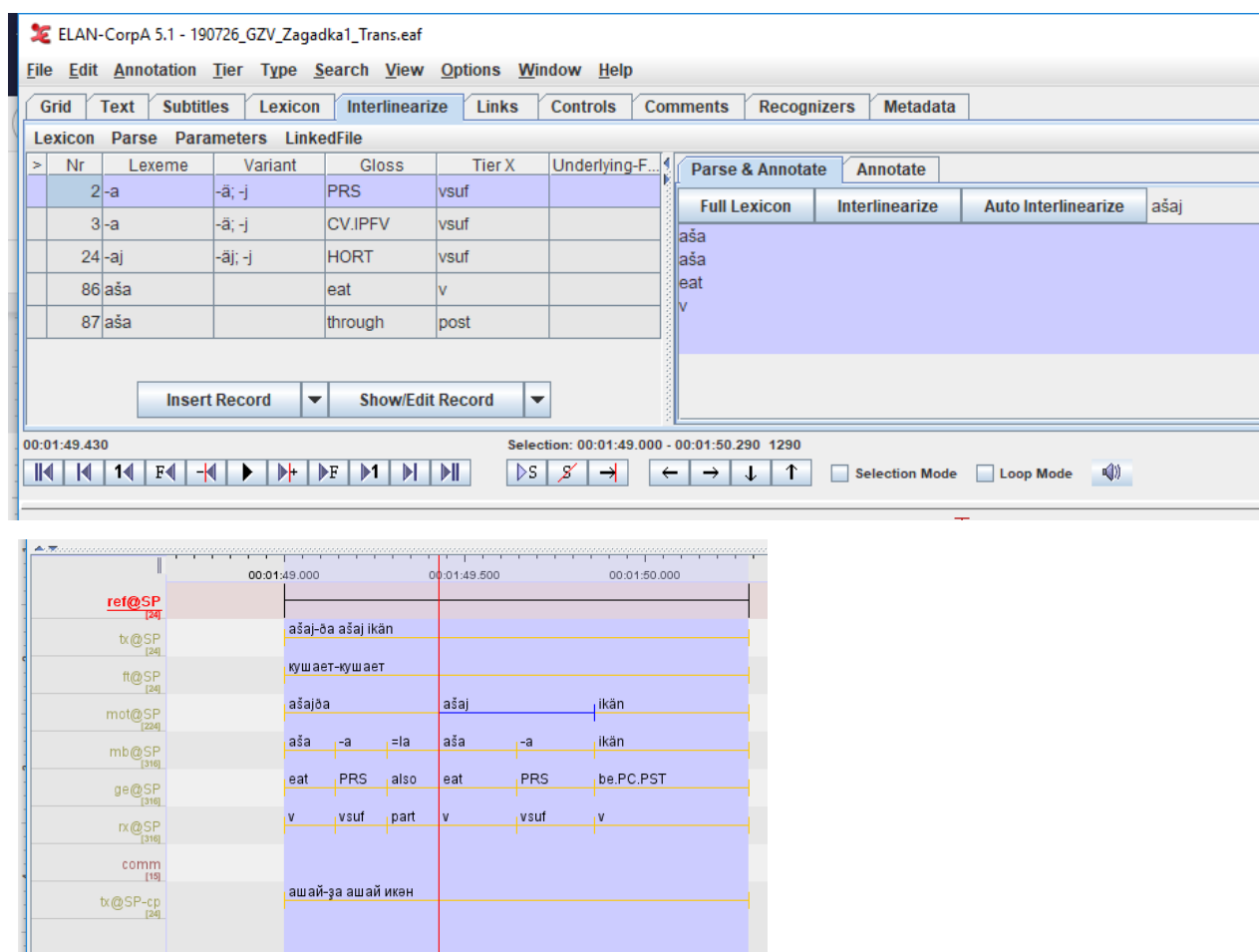
2.2. Сбор данных

Тексты Устного корпуса были записаны в 2011–2017 гг. в Абзелиловском районе (Рахметово, Баимово). Тексты Корпуса нарративов были записаны летом 2019 г. в Архангельском (Архангельское, Абзаново, Тавакачево, Кигзи, Муллакаево, Тереклы, Усаклы), Белорецком (Зуяково) и Абзелиловском районах (Баимово, Аюсазово, Еникеево). В целом, тексты обоих корпусов близки к литературному башкирскому языку, но имеют некоторые диалектные черты на уровне фонетики и морфонологии. Помимо аудиозаписей, тексты Корпуса нарративов записаны на видео.

2.3. Разметка данных

Тексты Устного корпуса были отгlossированы в программе Toolbox, после чего экспортированы в программу Elan для синхронизации с аудио файлами. Тексты Корпуса нарративов аннотируются в новой версии программы Elan – ElanCorpA [Chanard, 2015], которая позволяет glossировать файлы прямо в программе, без использования Toolbox³. На Рисунке 1 представлен скриншот встроенного в ElanCorpA парсера и процесс сегментации токенизированного аннотированного отрезка.

Рисунок 1. Glossирование в Elan CorpA



Важно отметить, что в двух корпусах приняты разные принципы сегментации аудио файлов при транскрибировании (слой tx): в Устном корпусе отдельными сегментами являются синтаксические единицы (предложения), тогда как Корпус нарративов сегментирован на просодические единицы (законченные интонационные отрезки). Однако это различие нейтрализуется при использовании специального шаблона, о котором будет сказано ниже.

Отгlossированные тексты обоих корпусов размечаются с помощью специального шаблона, созданного для кросслингвистического изучения конструкций передачи чужой речи⁴ [Nikitinaetal, 2019]. Шаблон состоит из четырёх дополнительных слоёв, в которых размечаются просодические единицы, в которых представлены конструкции передачи чужой речи (слой gr); семантический тип конструкции передачи чужой речи (слой tur); синтаксическая конструкция, используемая для выражения передачи чужой речи (слой qt);

³ Для более подробного описания новых функций ElanCorpA, см. http://llacan.vjf.cnrs.fr/res_ELAN-CorpA.php.

⁴ Подробная инструкция по использованию шаблона доступна по ссылке: <https://github.com/KatyaAplonova/ReportedDiscourseProject>

отсылки к участникам нарратива (слой par). На Рисунке 2 представлен пример разметки башкирского текста.

Рисунок 2. Пример разметки башкирского текста с помощью шаблона



3. Конструкции передачи чужой речи

3.1. Краткая типология стратегий передачи чужой речи в языках мира

Конструкции передачи чужой речи широко изучались в рамках различных лингвистических традиций. В недавней статье Спронка С. и Никитиной Т. В. [2019] обсуждаются наиболее известные исследования этих конструкций с целью формулировки рабочего определения конструкции передачи чужой речи (КПЧР)⁵ как самостоятельной синтаксической области. В этом разделе будут кратко представлены основные тезисы авторов данной статьи. Согласно Спронку и Никитиной (2019), КПЧР (1 и 2), имеют несколько особенностей, которые заметно отличают их от других типов конструкций.

- (1) *Вася сказал: «Смотри, дождь пошёл!».*
- (2) *Вася говорит, что пошёл дождь.*

Авторы формулируют ряд наблюдений о синтаксисе КПЧР, описывающих различные аспекты ее формального выражения и семантики в разных языках.

Обычно КПЧР имеют биклаузальную структуру, которая не вписывается в общепринятые трактовки ни подчинения, ни сочинения. Авторы предлагают обозначения R и M для составляющих КПЧР. R (reported) обозначает переданное сообщение (*Смотри, дождь пошёл!* в (1) и *пошёл дождь* в (2)), M (matrix) обозначает матрицу, вводящую R (*Вася сказал* в (1) и *Вася говорит, что* в (2)). R и M могут быть выражены двумя клаузами, но в некоторых случаях M может быть выражен морфемой или даже не выраженным вовсе, то есть маркирование M в КПЧР может быть факультативно.

Составляющие R и M имеют разную интерпретацию: R зачастую основывается на семантике, тогда как M – на прагматике. Так в (3а и 3б), дождь пойдёт в один и тот же день, однако в (3а) в R отсылка к моменту времени производится с помощью дейктического наречия “завтра”, тогда как в (3б) – с помощью “сегодня”. При этом, в M подобного сдвига в семантике дейктического наречия “вчера” не наблюдается.

⁵ Необходимо сделать оговорку о переоценки дихотомии прямой-косвенной речи. Данное различие справедливо лишь для небольшого количества западноевропейских языков. Эванс [2013] интерпретирует это противопоставление не как две дискретные категории, а как шкалу, где на одном полюсе находится “каноническая прямая речь”, а на другом – “каноническая косвенная речь”.

(3а) *Вчера Вася сказал: “Завтра пойдёт дождь”.*

(3б) *Вчера Вася сказал, что сегодня пойдёт дождь.*

В КПЧР представлен ряд сдвигов, такие как индексикальный [Shklovsky&Sudo 2013] и модальный (Evans 2006, 2013). Индексикальным сдвигом называется наличие внутренней отсылки в конструкциях с синтаксическим подчинением. В (4) глагольная форма в зависимой клаузе *àlózi* имеет маркер первого лица единственного числа, отсылающего к говорящему⁶.

Каримоджонг (Нилотские, Уганда)

(4) *àbù papà tolim ebè àlózi ijèz morotó*
AUX отец говорить что 1SG.идти.NPST 3SG Морото

‘Отец сказал, что поедет в Морото (букв.: Отеці сказал, что яі поеду в Морото)’

Модальным сдвигом называется биперспективизация эпистемических модальных маркеров. В некоторых языках наличие модального маркера в R меняет значение M. В (5) *goanna* ‘варан’ маркирован эвиденциальной частицей со значением неуверенности, сфера действия которой распространяется на *Mngamaranyalangkun* ‘я думал’, в результате чего мы понимаем, что говорящий был неправ и голова варана, на самом деле, находилась в другом месте.

Унгариныин (Уоррорран, Австралия)

(5) [*goanna-karranga*] *nga-ma-ra nya-langkunkunogoanna*
варан-может.быть 1SG-делать-PST 3SG-голова nw-DIST

‘Я думал, что там, возможно, голова варана’. (но я ошибся) [Spronck 2015b: 178]

Наконец, КПЧР чаще, чем другие конструкции сопровождается жестами и просодическими сигналами, то есть обладает мультимодальными свойствами, которым необходимо уделять должное внимание при описании.

3.2. Башкирские данные (краткое описание задач исследования)

КПЧР в башкирском языке обычно выражается следующим образом: R является сентенциальным актантом M, который содержит показатель *tip*, который происходит от деепричастной формы глагола *ti-* ‘говорить’.

(7) *Azamat jamyər jaw-a tip äjt-te.*
Азамат дождь падать-IPFV сказать.CV говорить-PST

‘Азамат сказал, что идёт дождь’. [Князев 2017]

В грамматике Юлдашева (1981) показатель *tip* рассматривается как маркер, вводящий прямую или косвенную речь.

На материале двух корпусов, описанных во Введении будет проанализирована структура M и R. Предварительно могут быть сформулированы следующие вопросы: всегда ли *tip* является обязательным элементом M и может ли говорящий быть выражен после R (если да, чем может быть обусловлена такая структура). Индексикальному сдвигу в башкирском посвящена недавняя статья Князева М. Ю. [Князев, 2017], однако модальный сдвиг в башкирских КПЧР не был систематически изучен, что также планируется осуществить в рамках проекта. Помимо этого, в Корпусе нарративов все тексты записаны на видео, что открывает новые возможности для изучения мультимодальности КПЧР.

5. Заключение

Данная статья кратко описывает методологию проекта “Discoursereportinginstorytelling”, который находится на стыке лингвистики, антропологии и фольклористики. Мы надеемся, что подобный комплексный подход поможет описать КПЧР на материале разных языков, что дополнит кросслингвистический концепт КПЧР. Сравнительные перспективы проекта имеют два направления. Во-первых, будут проведены сравнения двух групп языков из разных регионов (Западная Африка и Россия). Во вторых, КПЧР будут проанализированы внутри группы тюркских языков России (башкирский и чувашский).

⁶ В русском языке индексикальный сдвиг невозможен, ср. **Васяі сказал, что яі пойду в парк.*

Помимо лингвистических задач, проект призван сделать обширные данные по языкам России более доступным исследователям за пределами России. Кроме того, записанные традиционные нарративы могут стать материалом исследования других наук, таких как антропология, фольклористика и текстология.

Список глосс

AUX – вспомогательный глагол
CV – деепричастие
DIST – маркер дальнего дейксиса

IPFV – имперфектив
NPST – непрошедшее
PST – прошедшее

Литература

- Овсянникова, М. Сай, С., Аплонова, Е., Сметина, А., Сокур, Е. Устный корпус башкирского языка дер. Рахметова и с. Баимово. СПб.: ИЛИ РАН; М.: Международная лаборатория языковой конвергенции, НИУ ВШЭ. (Доступно онлайн по адресу: https://linghub.ru/oral_bashkir_corpus/, дата обращения 20.09.2019).
- Князев, М. Ю. Конструкции передачи чужой речи в башкирском языке. Труды ИЛИ РАН. XIII. № 1. 2017. С. 71–109.
- Хусаинова, Г. Р. Башкирские волшебные сказки: поэтика и текстология. Диссертации на соискание ученой степени доктора филологических наук. Москва, 2017.
- Юлдашев А. А. (отв. ред.). Грамматика современного башкирского литературного языка. Москва: Наука, 1981.
- Chanard, C. ELAN-Corpa (version 5). 2018.
- Evans, N. View with a view: Towards a typology of multiple perspective constructions. In Rebecca T. Cover & Yuni Kim (eds.), Proceedings of the thirty-first annual meeting of the Berkeley Linguistics Society. Berkeley: Berkeley Linguistics Society. 2006. P. 93–120.
- Evans, N. Some problems in the typology of quotation: A canonical approach. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), Canonical morphology and syntax. Oxford: Oxford University Press. 2013. P. 66–98.
- Hale, K., M. Krauss, L. J. Watahomigie, A. Y. Yamamoto, C. Craig, L. M. Jeanne & N. C. England. Endangered languages. Language 68. 1992. P. 1-42.
- Nikitina, T. Russian verboids: A case study in expressive vocabulary. Linguistics 50(2). 2012c. P. 165-89.
- Nikitina, T., A. Hantgan, C. Chanard. Reported speech annotation template for ELAN. Discourse Reporting in African Storytelling repository. 2019.
- Shklovsky, K., Sudo, Y. The syntax of monsters. Linguistic Inquiry 45, 2013. P. 381–402.
- Spronck, S., Nikitina, T. Reported speech forms a dedicated syntactic domain. Linguistic Typology 23(1). 2019. P.119–159.
- Spronck, S. Refracting views: How to construct complex perspective in reported speech and thought in Ungarinyin. STUF – Language Typology and Universals 68(2). 2015b. P. 165–185.
- UNESCO. Language vitality and endangerment. 2013. (Доступно онлайн по адресу: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/pdf/Language_vitality_and_endangerment_EN.pdf, дата обращения 20.09.2019).

Е. S. Aplonova

*The French National Centre for Scientific Research
UMR 8135 du CNRS 7, rue Guy Môquet - BP 8, 94801 VILLEJUIF Paris, France
E-mail: aplooon@gmail.com*

METHODOLOGICAL NOTES ABOUT CORPUS STUDY OF REPORTED SPEECH CONSTRUCTIONS IN BASHKIR USING ORAL CORPUS REFERENCES

The article describes the methodology for studying the structures of the transmission of foreign speech in the Bashkir language in the framework of the project “Discoursereportinginstorytelling”. The study is based on the material of two oral annotated buildings marked up with the help of a special template

for cross-linguistic study of the structures of the transmission of other people's speech. These constructions are of particular interest for an integrated approach to the study of texts, since, on the one hand, they are not sufficiently studied as a separate linguistic construct in a typological sense, on the other hand, they involve the collection and analysis of traditional narratives, which are a valuable source of knowledge about the culture of native speakers .

Keywords: constructions of the transmission of someone else's speech, corpus linguistics, oral corpus, text markup methodology, multimodal analysis.

References

- Ovsyannikova, M., Sai, S., Aplonov, E., Smetana, A., Sokur, E. Oral case Bashkir language vill. Rakhmetova and Baimovo. SPb.: OR RAS; M.: International laboratory of language convergence, HSE. (Available online at: https://linghub.ru/oral_bashkir_corpus/, accessed 20.09.2019)
- Knyazev, M. Yu. Constructions of foreign speech transmission in the Bashkir language. Proceedings OR WOUNDS. XIII. No. 1. 2017. P. 71-109.
- Khusainova, G. R. Bashkir fairy tales: poetics and textual studies. Dissertations for the degree of doctor of Philology. Moscow. 2017.
- Yuldashev A. A. (resp. ed.). Grammar of the modern Bashkir literary language. Moscow: Nauka, 1981
- Chanard, C. ELAN-CorPA (version 5). 2018.
- Evans, N. View with a view: Towards a typology of multiple perspective constructions. In Rebecca T. Cover & Yuni Kim (eds.), Proceedings of the thirty-first annual meeting of the Berkeley Linguistics Society. Berkeley: Berkeley Linguistics Society. 2006. P. 93–120.
- Evans, N. Some problems in the typology of quotation: A canonical approach. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), Canonical morphology and syntax. Oxford: Oxford University Press. 2013. P. 66–98.
- Hale, K., M. Krauss, L. J. Watahomigie, A. Y. Yamamoto, C. Craig, L. M. Jeanne & N. C. England. Endangered languages. Language 68. 1992. P. 1-42.
- Nikitina, T. Russian verboids: A case study in expressive vocabulary. Linguistics 50(2). 2012c. P. 165-89.
- Nikitina, T., A. Hantgan, C. Chanard. Reported speech annotation template for ELAN. Discourse Reporting in African Storytelling repository. 2019.
- Shklovsky, K., Sudo, Y. The syntax of monsters. Linguistic Inquiry 45, 2013. P. 381–402.
- Spronck, S., Nikitina, T. Reported speech forms a dedicated syntactic domain. Linguistic Typology 23(1). 2019. P.119–159
- Spronck, S. Refracting views: How to construct complex perspective in reported speech and thought in Ungarinyin. STUF – Language Typology and Universals 68(2). 2015b. P. 165–185.
- UNESCO. Language vitality and endangerment. 2013. (Available online at: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/pdf/Language_vitality_and_endangerment_EN.pdf, accessed date 20.09.2019).