

Плюсы и минусы удаленной документации: сравнение опыта работы с языками Дагестана, Сибири и Дальнего Востока

Н. Б. Аралова^{5,6}, Е. М. Будянская¹, М.-Э. А. Винклер¹, Т. И. Давидюк^{1,2},
А. П. Евстигнеева¹, Ю. Б. Коряков^{1,3}, Д. Д. Мордашова¹, В. С. Харитонов¹

¹ Институт языкознания РАН, ² МГУ им. М. В. Ломоносова, ³ НИУ «Высшая школа экономики», ⁴ Westfälische Wilhelms-Universität Münster, ⁵ Christian-Albrechts-Universität zu Kiel

Под удаленной документацией (англ. *remote / distributed documentation*) понимается такой вид сбора и / или обработки данных, когда лингвист работает с участниками исследования дистанционно, не находясь непосредственно в месте проживания языкового сообщества. Данный доклад представляет собой попытку обобщить и проанализировать опыт удаленной работы с носителями языков нескольких ареалов.

Методология удаленных исследований стала особенно активно развиваться с 2020 года в связи с пандемией. Так, в сентябре 2021 г. Программа по документации исчезающих языков (англ. Endangered Languages Documentation Programme, ELDP) организовала серию мероприятий, посвященных различным аспектам дистанционной работы. Также в 2020–2022 гг. были опубликованы методологически ориентированные статьи об опыте удаленных документационных проектов (см. [Leemann et al. 2020; Rice 2021; Rozhanskiy 2021: 33–34; Williams et al. 2021; Grzech, Shaca 2022]). Тем не менее, на настоящий момент нам неизвестны работы, которые фокусировались бы на специфике дистанционной работы с языками России.

В докладе мы представим опыт работы с носителями анклухского диалекта цудахарско-сирхинского языка, татского (язык татов-мусульман из сс. Джалган, Нижний Джалган и Митаги-Казмаляр), мансийского, уильтинского (орокского), якутского, эвенского и северноюкагирского языков. С точки зрения используемых инструментов, удаленная работа с информантом может выглядеть по-разному в зависимости от ряда факторов. С одной стороны, это ограничения со стороны носителей: уровень компьютерной грамотности, доступность необходимой для работы техники и ее характеристики, доступ к сети Интернет (который в том числе зависит от ситуации в регионе в целом), существование письменности на языке и грамотность носителя. Другими факторами являются цели исследователей: тип задачи — сбор нового материала (запись текстов, словаря или материалов для разговорника) или анализ уже собранных материалов. Если задачей является сбор нового материала, то методология также зависит от типа материала, который предстоит собрать. Кроме того, мы покажем, что в некоторых случаях одна и та же задача может успешно решаться разными способами.

Так, в рамках каждого из наших проектов проводится удаленная работа с записанными текстами (транскрипция и перевод), однако эта задача решается по-разному. В первой группе случаев (анклухский диалект цудахарско-сирхинского языка, татский, эвенский, мансийский) носители транскрибируют и переводят на русский фрагменты записи в текстовом редакторе, затем отправляют лингвистам свою работу, а задача лингвиста заключается в переносе материала в ELAN. Случай татского языка представляет особый интерес, так как у татского в Дагестане нет официальной письменности, и носитель разработал собственную систему записи языка. В случае с мансийским дистанционные технологии использовались не только для расшифровки полученных данных, но и для сбора новых. Во второй группе случаев (якутский, северноюкагирский) носители были обучены пользоваться программой ELAN и работали в ней самостоятельно. В третьем случае (уильтинский) лингвисты

предварительно расшифровывают тексты в ELAN, а затем корректируют запись и переводят совместно с носителями, используя демонстрацию экрана в Zoom.

В докладе мы представим сравнительный обзор методов, используемых в наших проектах, и более подробно остановимся на преимуществах и ограничениях каждого из них по сравнению с традиционной полевой работой, а также сопоставим наш опыт с международным опытом, описанным в литературе.

Литература

Grzech, Shaca 2022 — Grzech, Karolina, Selena Tisalema Shaca. 2022. Building trust on Zoom: A workflow for language documentation via videoconferencing software. *Language Documentation & Conservation*, 16, p. 79–97.

Leemann et al. 2020 — Leemann, Adrian, Péter Jeszenszky, Carina Steiner, Melanie Studerus & Jan Messerli. 2020. Linguistic fieldwork in a pandemic: Supervised data collection combining smartphone recordings and videoconferencing. *Linguistics Vanguard*, 6(s3).

Rice 2021 — Rice, Alexander. 2021. Using YouTube as the Primary Transcription and Translation Platform for Remote Corpus Work. *Language Documentation & Conservation*, 15, p. 514–550.

Rozhanskiy 2021 — Rozhanskiy, Fedor. 2021. A Manual of Linguistic Fieldwork. Tartu. URL: https://copius.eu/files/fieldwork_2021.pdf.

Williams et al. 2021 — Williams, Nicholas, W. D. L. Silva, Laura McPherson & Jeff Good. 2021. COVID-19 and documentary linguistics: Some ways forward. *Language Documentation and Description*, 20, p. 359–377.