

## **Предложение о проведении Круглого стола по методам репрезентации текстов на малых языках.**

Отдел лингвокультурной экологии Института мировой культуры МГУ совместно с ТПУ предлагают провести в рамках конференции LENCA-3 круглый стол, посвященный проблемам глоссирования текстов на малых и/или бесписьменных языках.

В последние десять – пятнадцать лет была проведена значительная теоретическая работа по разработке методов документирования малых языков, в том числе, методов глоссирования текстов на малых языках. Достаточно полно результаты этой работы изложены в статьях Кристиана Лемана и в Лейпцигских правилах глоссирования. Тем не менее, следующий этап, а именно распространение предлагаемых методов среди дескриптивистов и стандартизация представления текстов на малых языках, пока еще очень далек от завершения.

В России, с ее значительным этнокультурным разнообразием, с одной стороны, остро стоит проблема документирования малых языков и ведется деятельность в этом направлении, а с другой, стандартизация представления текстов, в том числе стандартизация морфологического глоссирования текстов, находятся в зачаточном состоянии. Существует несколько изолированных традиций, внутри себя достаточно последовательных, которые, однако, в некоторых существенных моментах расходятся друг с другом и/или с Лейпцигскими правилами (в качестве примера приведем традицию репрезентации текстов в полевых исследованиях филфака МГУ). Стандартизация этих традиций требует совместных усилий полевых лингвистов разных регионов России. Предлагаемый Круглый стол, в рамках которого существующие стандарты и подходы будут разъясняться и оцениваться, может стать первым шагом в этом направлении, особенно благодаря тому, что он сможет объединить не только российских лингвистов как минимум трех различных регионов и традиций (Сибирь, Москва, Санкт-Петербург), но и представителей Института эволюционной антропологии Макса-Планка, где собственно разрабатывались Лейпцигские правила.

Важно также иметь в виду, что в ближайшей перспективе стоит задача создания электронных корпусов глоссированных текстов на малых языках РФ. Стандарт репрезентации текстов должен обеспечить не только сходное представление текстов в печатных публикациях, но и возможность единообразного ввода текстов в корпус, различной их обработки и поиска, в том числе поиска по грамматическим глоссам.

Проблематика, которую предполагается осветить в рамках Круглого стола, включает следующие моменты: инвентарь глосс морфологических категорий; инвентарь и правила использования разделителей; стандартизация транскрипции; репрезентация некоторых специальных типов текстов (например, фальстарты, репризы и т.п. в спонтанных текстах), технические проблемы представления глоссированных текстов средствами различных программ, доступных исследователю. Некоторые из перечисленных моментов очевидно требуют обсуждения.

Кроме того, одной из важных задач дискуссии может стать поиск баланса между стандартизацией методов представления текстов и эффективностью этих методов для представления данных конкретных языков. Иными словами, единый стандарт глоссирования, будучи достаточно жестким для того, чтобы обеспечивать сравнимость репрезентации текстов на разных языках, должен оставлять исследователю определенную свободу решений, необходимую для адекватного и эффективного представления языковых данных, а также его знаний о языке (часто недостаточных). Частными примерами этого являются:

- решение о необходимости или избыточности репрезентации просодических данных (акцентуации, фразовой просодии);
- затрудненность выбора между тем или иным типом разделителя (например, точка или дефис при фузии);
- отсутствие достаточных теоретических оснований или эмпирических данных для характеристики сегмента как отдельного фонетического слова или клитики.

Очевидно, что, как и в случае метаязыка теоретической лингвистики, поиск универсального метаязыка репрезентации текста должен осуществляться с привлечением как можно более широкого круга типологически различных языков и специалистов по этим языкам. Конференция LENCA-3 может стать форумом, как нельзя более удачно удовлетворяющим этим требованиям.

Язык проведения: русский, английский.

Продолжительность: ориентировочно, одна секция (будет дополнительно проясняться с организациями конференции и заинтересованными участниками).

Организаторы круглого стола:

А.Е. Кибрик (МГУ)

М.А. Даниэль (МГУ)

А.В. Архипов (МГУ)

А.Ю. Фильченко (ТПГУ)