

ОТЗЫВ РУКОВОДИТЕЛЯ
на выпускную квалификационную работу обучающегося СПбГУ
Солоновича Егора Евгеньевича

по теме «Синтаксические модели языка бамана в контексте создания синтаксического корпуса языка»

Тема выпускной квалификационной работы Е. Е. Солоновича была сформулирована в ходе работы международного семинара по созданию синтаксически аннотированных электронных корпусов языков манден (манинка и бамана). Главная цель его работы — определение стратегии при выборе программного обеспечения, которое соответствовало бы далекоидущим целям синтаксической аннотации текстов бамана, и при этом позволяло бы сохранить все уровни морфологической аннотации формата морфологически размеченного Справочного корпуса бамана.

Е. Е. Солонович провёл исследование в соответствии с поставленной задачей: опираясь на знания грамматики языка бамана, полученные в ходе обучения на кафедре африканистики, он систематически изучил различные программы, имеющиеся в открытом доступе, и сравнил их возможности. Для этого ему потребовалось также приложить серьёзные усилия для повышения собственной компетенции в области информатики. Дополнительным требованием к программному обеспечению была его хорошая совместимость с форматом, разработанным в рамках открытого проекта Universal Dependencies, имеющего целью стандартизировать синтаксическую аннотацию разных языков мира, поэтому требовалось обратить особое внимание на изучение особенностей этого формата (описанию которого и посвящена значительная часть ВКР). В работе также представлены результаты совместной работы участников семинара по созданию системы конвертации формата Стандартного корпуса бамана в формат, принятый в рамках проекта Universal Dependencies.

В результате сравнения различных программ Е. Е. Солонович приходит к выводу, что сформулированным требованиям, в той или иной степени, отвечают по своим характеристикам две: DGAnnotator и WebAnno. Однако первый из них, DGAnnotator, не прошёл этап тестирования — как выяснилось, он не позволяет сохранять все строки аннотации Справочного корпуса бамана, и, к тому же, он приспособлен только для ручного синтаксического аннотирования, что делает его применение неперспективным.

Напротив, программа WebAnno лишена этих недостатков, и её использование для синтаксической аннотации баманских текстов имеет смысл.

Исследование Е. Е. Солоновича может, на первый взгляд, показаться достаточно скромным, однако мне хотелось бы обратить внимание на то, что оно позволяет сделать важный шаг в корпусной лингвистике в применении к языкам манден. Можно сказать, что главное достижение — это даже не столько выбор конкретной программы для синтаксической аннотации баманских текстов, сколько формулирование базовых требований и подходов в этой области, достаточно новой для кафедры африканистики.

Помимо этого, следует отметить, что исследование Е. Е. Солоновича являет собой положительный пример работы в команде: его цели формулировались в ходе обсуждений на семинаре, а результаты будут самым непосредственным образом использованы в ходе дальнейшей работы по созданию синтаксически аннотированных корпусов языков манден.

По моему мнению, данная квалификационная работа заслуживает отличной оценки.

Научный руководитель



д.филол.н., профессор

Выдрин Валентин Феодосьевич
Санкт-Петербургский Государственный Университет