

Роль языковой технологии в сохранении и ревитализации языка

Тронд Тростерюд

13 марта 2011 г.

Содержание

1	Что такое языковая ревитализация?	1
2	Почему лингвистическая технология важна?	3
2.1	Вообще	3
2.2	Гиеллатекно	3
2.2.1	Лингвистическая технология в Тромсё	3
2.3	Клавиатура	5
2.4	Электронные словари	5
3	Килдинский язык	5
3.1	Орфография	5
3.2	Буквы	5
3.3	Клавиатура	5
3.4	Основная технология	8
3.5	Педагогические программы	8
3.6	Система проверки правописания	8
3.7	Машинный перевод	9
4	Какова ситуация сегодня?	9
5	Заключение	10

1 Что такое языковая ревитализация?

Половина языков в мире может исчезнуть. Для многих это благоприятная ситуация: меньшее количество языков означает меньше языковых барьеров.

Для меня, как лингвист, — это катастрофа. Язык - это то, что отличает людей от животных.

Каждый язык содержит уникальную кодифицированную информацию об окружающем мире. С каждым языком, который исчезает, мы теряем уникальное представление мира.

Но сохранить многоязычное общество не просто. Поддерживать язык, или даже оживлять его, невозможно только имея профессионального интереса к языкам. Мотивация непосредственно должно исходить от самого языкового сообщества.

Языки не исчезают сами собой. Но обычно происходят следующие процессы:

Проводился процесс модернизации от традиционной натуральной экономики к современной экономике, которая была основана на деньгах и совпала с укреплением государства. В новой интегрированной экономике было идеологическое давление на языковые меньшинства чтобы они заменили свой язык на язык большинства.

Многие носители языков меньшинств несут травмы из своего детства: они обучались в школах-интернатах, где они не понимали нового языка. И им запрещали говорить на единственном языке, который они знали - на родном языке.

Люди этого поколения не хотели, что их дети пострадали таким же образом. Они воспитывали своих детей на языке большинства, хотя они часто сами не очень хорошо на нем говорили. Следующее поколение выросло, слушая язык меньшинства, и возможно, разговаривая на нем со своими бабушками и дедушками.

Сегодня третье поколение становится взрослым. Мы называем это "феномен третьего поколения". Они не пережили те же самые травмирующие события, которые пережили их бабушки и дедушки. У них все" еще есть этническая идентичность.

Но они не говорят на языке своей этнической группы. Возможно, они понимают его. Вероятно, они способны очень хорошо произносить слова, но они не говорят.

Такая ситуация характерна для многих языковых меньшинств.

В процессе языкового сдвига была сделана категорическая ошибка. Ошибка состояла в том, что не было обращено внимание на возможность многоязычия. В многонациональном интегрированном обществе одного языка недостаточно. Даже два языка недостаточно.

Другая ошибка состоит в том, что язык меньшинства рассматривается как язык, который отличается от всех других языков. Язык не считается средством коммуникации, а символом для чего-то другое. Язык рассматривается как кое-что, что не может быть изучено, а только забыто.

И это одна из причин для того, почему преподавание языка меньшинства дает плохие результаты. Потому что уже изначально цель не состоит в изучении языка, а в выполнении символического акта.

Противоположно к этому является коммуникативное изучение языка. Вовсе возможно выучить язык даже без погружения в языковую среду. Это показывает пример того, что я называю методом "Интуриста".

Каждый норвежец, который побывал туристом в Советском Союзе знаком с этим явлением: Наши русские туристические гиды имели отличные знания норвежского языка, хотя они никогда не были в Норвегии.

Мотивация для языковой ревитализации является следующей: Языковая ревитализация - это вопрос идентичности. Для членов этнического меньшинства доступ к языку является частью обнаружения собственной идентичности.

Это не может быть сделано посторонними. Я, как норвежец, не могу восстановить язык кильдинских саами. Русские тоже не могут восстановить язык кильдинских саами.

Но как лингвист, я могу дать советы. Я могу рассказать данному языковому сообществу о событиях в других языковых сообществах, привести аналогичные примеры. И я могу рассказать языковому сообществу об опыте других языковых сообществ. И я могу рассказать вам, почему языковая технология важна, хотя самое важное - это начать использовать язык.

2 Почему лингвистическая технология важна?

2.1 Вообще

Что такое языковая технология? Технология - это ремесло создания инструментов. Языковая технология - это ремесло создания инструментов, которые необходимы для использования языка. Некоторые языковые технологии независимы от языка, другие связаны с определенными языками.

2.2 Гиеллатекно

2.2.1 Лингвистическая технология в Тромсё

В Tromsø у нас есть центр голландской лингвистической технологии. Начальную страницу показывают на числе 1. Иллюстрация 2 показывает, как наши анализаторы склоняют слова.

В компьютере такие автоматы

Чтобы показать пример для того, какую роль может играть языковая технология для языка меньшинства, я хотел бы показать Вам, чего мы достигли для северосаамского языка:

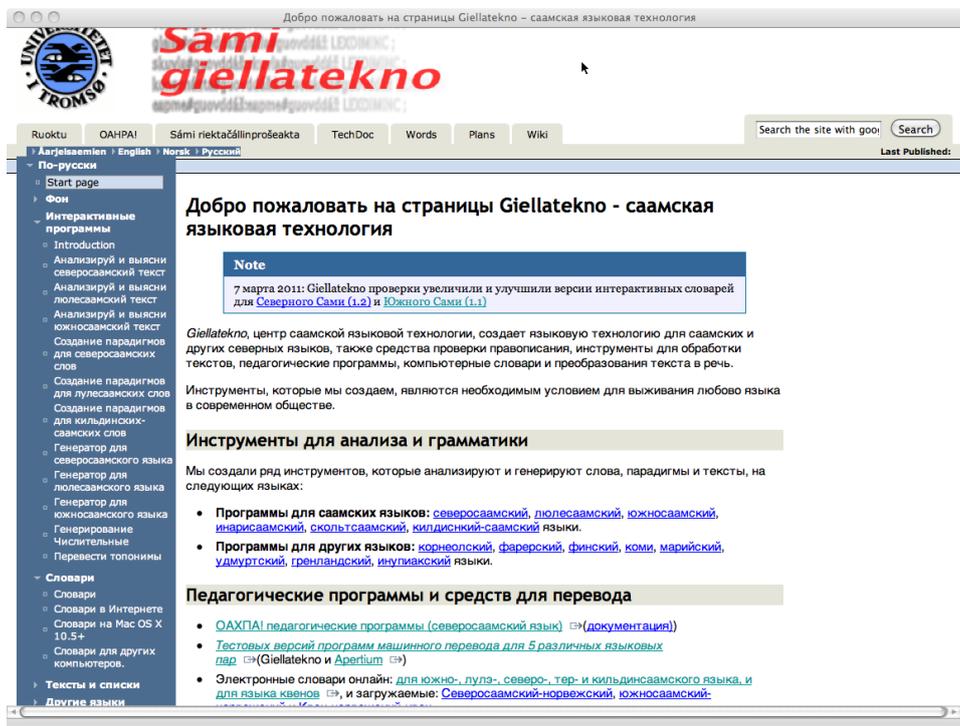


Рис. 1: Домашняя страница Гиеллатекно

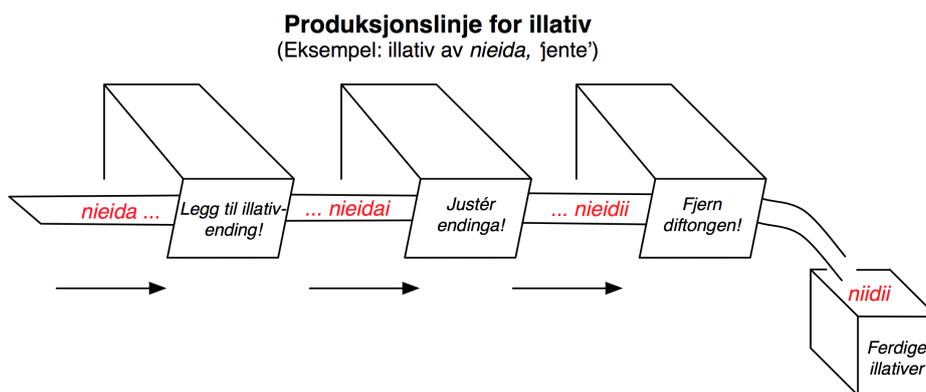


Рис. 2: Грамматика как морфологический завод

2.3 Клавиатура

Б Рис. ?? клавиатура для северосаамского языка.

2.4 Электронные словари

Наши электронные словари комбинируют словарь и морфологический анализ. См. 4 и 5.

Словарь также дает грамматическую информацию.

3 Килдинский язык

3.1 Орфография

Орфография - это спорная проблема для языка. Для кильдинского языка есть два основных позиций: использовать кириллицу, или использовать латинский алфавит. Сегодня используется два варианта кириллической орфографии: одна с двумя латинскими буквами (h, j), и одна только с кириллическими буквами. Тоакже есть интерес к изменению кириллицы на латинский алфавит.

Мой совет, когда дело доходит до орфографии — следующий: не превращайте это в главную полемику. Конечно важно иметь хорошую орфографию, но сама орфография не сможет сохранить язык. Борьба из-за орфографии может убить движение оживления, ревитализации языка.

Различие между двумя кириллическими версиями орфографии для кильдинского языка является настолько маленьким, что с помощью компьютерных программ легко конвертировать текст между ними (т.е. можно использовать оба варианта орфографии паралельно). Латинский алфавит не сделан для саамских языков: в саамских языках есть много согласных и палатализаций, точно так же как в русском языке.

3.2 Буквы

Unicode - это стандарт кодирования символов, позволяющий представить знаки практически всех письменных языков в цифровых документах.

Международный стандарт Unicode содержит все кильдинские буквы.

3.3 Клавиатура

В кильдинском языке есть много символов, которые отсутствуют в русском языке. Есть много разных расположений символов на клавиатурах для кильдинского

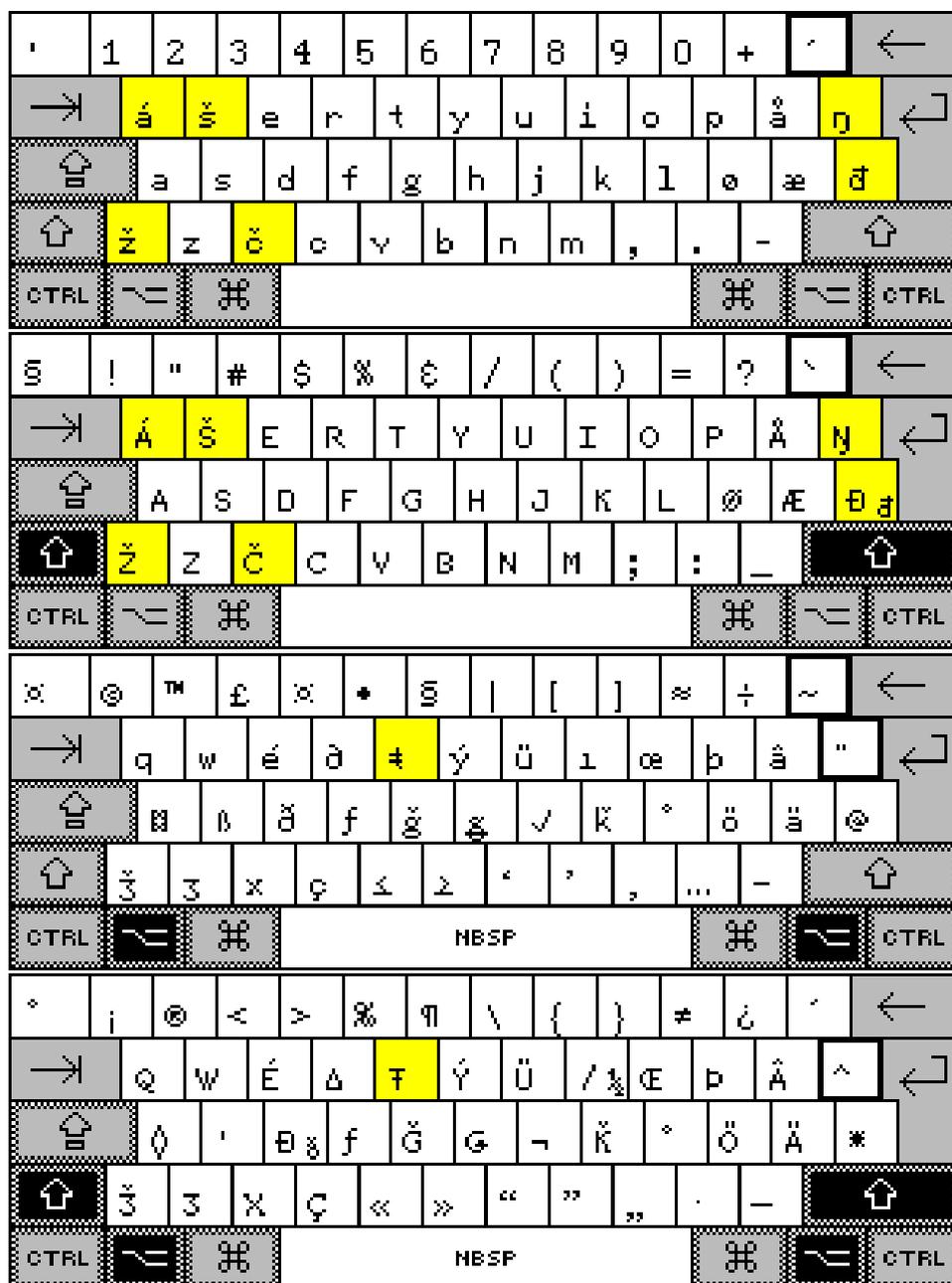


Рис. 3: Северсаамская клавиатура



Рис. 4: Как прочитать текст с электронным словарем



Рис. 5: Грамматика в словаре

языка, но ни одно из них в общественном использовании. Также эти клавиатуры имеют многие технические проблемы.

Сегодня создание хороших клавиатур с хорошим дизайном является самой важной задачей для языковой технологии кильдинского языка. Пока у нас нет хорошей клавиатуры и напечатать текст - это медленный процесс.

В Тромсё мы планируем сделать клавиатуры для коми и кильдинского языков. Мы будем оценивать все различные раскладки клавиатуры, которые существуют. Также мы сделаем статистические исследования букв, и пользователи кильдинского и коми языков будут проверять клавиатуры. Наша самая важная задача - это улучшить техническое решение для клавиатуры и процедуру инсталляции.

3.4 Основная технология

Все расширенные решения для языковой технологии основаны на лингвистическом анализе языка. Для кильдинских саамов это значит, что мы должны будем найти форму леммы любой формы слова.

Мы также должны быть в состоянии генерировать любую форму слова от основной формы. Для каждого существительного мы должны создать все падежные формы и для каждого глагола мы должны создать все формы спряжения.

3.5 Педагогические программы

С помощью программ для грамматического анализа мы также можем создать интерактивные педагогические программы. Эти изображения показывают некоторые примеры, см. 6, в которое компьютер преподает новые слова.

3.6 Система проверки правописания

Система проверки правописания является центральной частью современной продукции текстов. Ситуация для кильдинского языка очень серьезна. Языковая норма еще не до конца установлена.

Хотя многие говорят на кильдинском языке, только очень мало людей умеет писать на кильдинском языке. Программа проверки орфографии будет функционировать как демократизация языка: в будущем все носители языка будут иметь более простой доступ к языку.

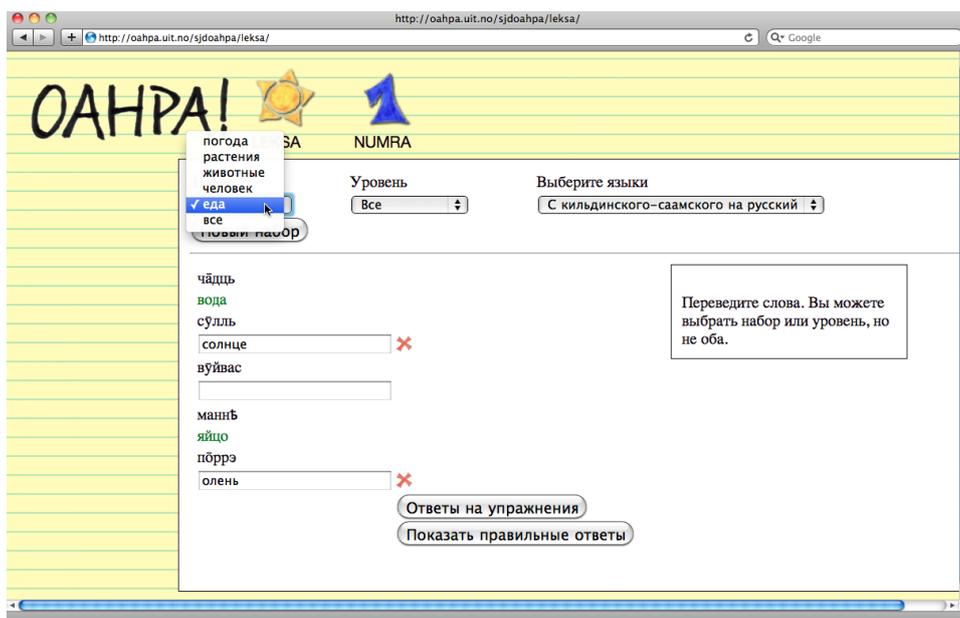


Рис. 6: OAHPI! Lexsa: Учится запас слов

3.7 Машинный перевод

В будущем, многоязычные интернет-страницы и многоязычные системы администрирования будут поддержаны автоматическими машинными программами для переводов. Этот доклад, например, написан с помощью машинного перевода.

Мы можем различить два типа машинного перевода: перевод, чтобы понимать текст, и перевод, чтобы произвести текст. Например: Вы читаете текст на английском языке, и Вам трудно его понимать. Вы используете машинный перевод и можете читать его на русском языке. Российский текст полный ошибками, но Вы можете понимать его.

Интернет-словари для нескольких языков, см. 7.

Но что, если Вы работаете в многонациональной компании или организации? Ваше задание - сделать русскую версию английской интернет-страницы. Снова Вы используете машинный перевод. Но в этот раз Вы обнаружите, что текст содержит столько много ошибок, что будет быстрее, если Вы сами (без использования машинного перевода) напишите этот текст на русском языке.

4 Какова ситуация сегодня?

Сегодня есть решения для языковой технологии для северосаамского языка. В Тромсё мы планируем сделать подобные решения для языка кильдинских

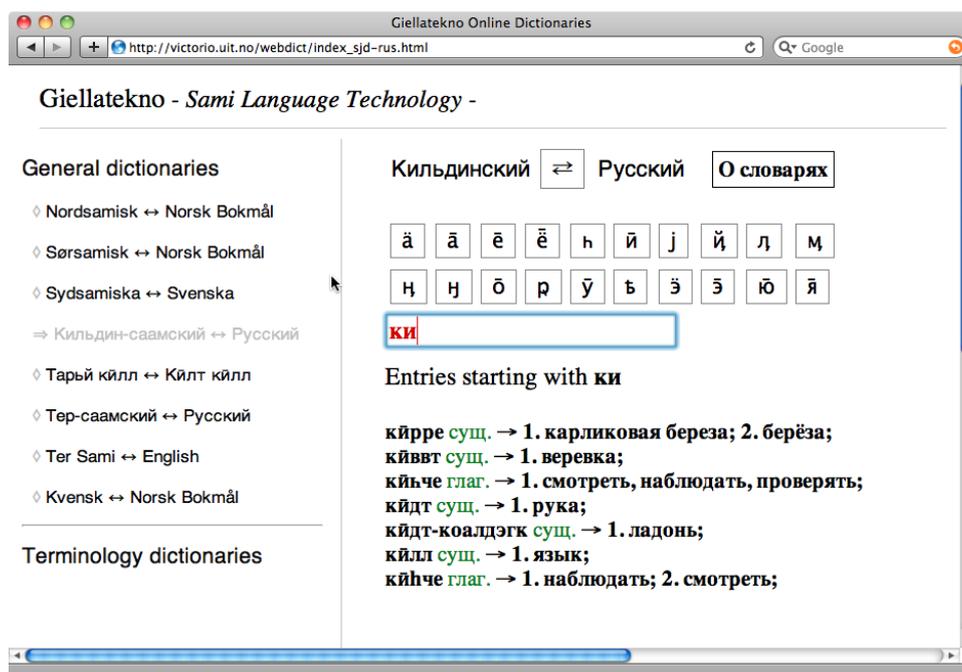


Рис. 7: Словар на интернете

саами, в сотрудничестве с кильдинскими языковыми активистами. Наш проект - это часть норвежской программы для Северных Территорий, но пока мы не знаем, будет ли финансирование для такого проекта.

Большая часть работы уже сделана без финансирования: прибавление символов кильдинского языка к системе Unicode, и первая версия электронного словаря для кильдинского языка.

5 Заключение

В ревитализации языка основной фокус должен быть на носителях языка. Те люди, которые знают язык, должны начинать использовать его, или использовать его больше. Люди, которые не знают язык, должны учить его.

Общество должно создать языковые гнёзда для детей. Эта, так называемая метода языковых гнёзд, эффективно используется в других языковых обществах.

Языковая технология может обеспечить языковое сообщество с необходимыми условиями для создания литературного языка и его поддержки.