

ВЕКТОРНО-КООРДИНАТНЫЙ МЕТОД ИЗМЕРЕНИЯ РАСХОЖДЕНИЙ МЕЖДУ ТЕРРИТОРИАЛЬНЫМИ РАЗНОВИДНОСТЯМИ ЯЗЫКА

Н. В. Беленов

*Самарский государственный социально-педагогический университет,
Самара, Россия, rectorat@sgspu.ru*

В работе предложена методика сравнения расхождений между территориальными разновидностями языка с помощью векторно-координатного метода. Рассмотрен пример соответствующего сравнения трёх близкородственных говоров эрзя-мордовского языка, бытующих в Похвистневском районе Самарской области, по кластеру географической терминологии.

Ключевые слова: компаративистика; математические методы; векторно-координатный метод; лексико-семантические различия; эрзя-мордовский язык.

Идея измерения в условных единицах расхождений между различными территориальными разновидностями языка (диалектами, группами говоров, говорами, подговорами), а также между различными языками не нова. История развития данной идеи насчитывает достаточное количество более или менее удачных примеров решения проблемы измерения межъязыковых и междиалектных различий.

Как правило, исследователи концентрируются на каком-то одном аспекте: лексическом, фонологическом, семантическом и т.д. [1; 3].

Упомянем, например, о методе измерения уровня расхождения между языками с помощью расстояния Левенштейна, который чаще используется в компьютерной лингвистике и теории информации, но может быть применим и в сравнительно-историческом языкознании [5].

Известной популярностью среди компаративистов пользуется графоаналитический метод, получивший наибольшую известность в свете исследований В.М. Иллич-Свитыча, стремившегося на его основе проиллюстрировать схему родства «ностратических» языков [2].

Предлагаемый нами векторно-координатный метод оценки расстояний между территориальными разновидностями языка отличается сравнительной простотой вычислений, при этом заключает в себе потенциал дальнейшего неограниченного расширения на все языковые аспекты сравнения исследуемых разновидностей.

В данной работе мы описываем двумерную модель для вычисления лексико-семантических расхождений между разновидностями языка (в нашем примере – между тремя близкородственными говорами эрзянско-

го языка). Материалы исследования были собраны и обработаны автором в течение полевых сезонов 2015 – 2024 гг.

При необходимости в перспективе можно ввести в модель оценку акцентологических, фонетических, морфологических и других различий, поскольку векторное пространство может быть n -мерно, а значения, которые затруднительно отобразить графически (после 3 измерения), будут вычисляться аналитически.

Идея метода состоит в следующем. Все действия выполняются в первой четверти декартовой системы координат. За начало координат выбирается один из сравниваемых говоров, все его лексико-семантические характеристики принимаются за точку (0; 0). Далее по оси абсцисс отсчитываются лексические отличия от него сравниваемого говора, а по оси ординат – семантические отличия.

Руководствуемся при этом следующими соображениями: лексема нет вообще (утрачена в принципе, заменена заимствованием из другого языка и т.д.) – шаг по оси абсцисс; сравниваемая лексема в обоих говорах присутствует, но её семантика отличается – шаг по оси ординат.

По окончании сравнения отмечаются координаты точки отклонения, которая фиксируется в результате суммы лексико-семантических отличий – это конец вектора, а начало координат, соответственно, начало вектора. Далее вычисляем длину вектора по известной формуле: квадратный корень из суммы квадратов координат, что и будет являться установленным расстоянием между говорами.

Проиллюстрируем применение модели на конкретном примере. Выбор материалов сравнения обусловлен научными интересами автора в области финно-угорского языкознания и диалектологии. Рассмотрим три близкородственных говора эрзя-мордовского языка, бытующих в Похвистневском районе Самарской области. Это говоры эрзянского населения сёл Большой Толкай, Малый Толкай и Большая Ёга. Носители данных говоров имеют общий этап заволжской истории (порядка 300 лет), при этом носители первых двух контактировали между собой и до переселения в Заволжье. Сами информанты неизменно отмечали, что их говоры близки между собой, однако признают, что большетолкайский и малотолкайский говоры находятся друг с другом в более тесном родстве, чем каждый из них с большеегинским.

Для сравнения говоров выберем кластер географической терминологии (выбор обусловлен теми же уже упомянутыми выше факторами). Разумеется, расширение выборки на другие кластеры лексики исследуемых говоров (можно использовать различные варианты списка М. Сво-

деша и его расширения с вариантами для различных языковых семей) приведёт к уточнению искомого расстояния. В целом, отбор лингвистического материала для применения математических методов в языкознании – отдельная тема, с нерешёнными дискуссионными вопросами, поэтому мы здесь подробно в неё погружаться не будем. Отметим лишь один факт: чем больше выборка, тем точнее результаты, соответственно, в любом подобном исследовании, при возможности, необходимо стремиться к сравнению максимально полных материалов по изучаемым разновидностям языка. Возвращаясь к выбранному кластеру географической лексики, отметим, что лексические единицы из этого кластера неизменно входят в различные варианты списков Сводеша для разных языков [4; 6], в свете чего его выбор для примера представляется оправданным.

К сравнению привлекались эрзянские эквиваленты, бытующие в соответствующих говорах, следующих лексем: река, озеро, лес, село, овраг, колодец, родник, кладбище, поле, роща, гора, переулоч, ручей, поляна, дорога, пруд, центральная часть села, лужа, брод, болото. Всего двадцать единиц.

За начало координат были выбраны данные малотолкайского говора эрзя-мордовского языка. Сравнение малотолкайского и большетолкайского говоров в данном кластере показало два лексических и два семантических отличия, что зафиксировало конец вектора в точке с координатами (2; 2). Расстояние между говорами в данном случае составило квадратный корень из 8 или, с округлением до тысячных: 2, 828.

При сравнении малотолкайского говора с большеегинским точка отклонения имела координаты (4; 2) – четыре лексических и два семантических отличия в кластере географической терминологии. При округлении до тысячных, это составляет 4, 472.

Отметим ещё раз: полученные в примере результаты приблизительны, в силу малой выборки и сосредоточения на одном конкретном лексическом кластере. Однако, даже они показали картину, близкую к реальности (как исторической, так и лингвopsихологической). Работа над совершенствованием предложенной методики оценки расхождения говоров будет продолжена автором.

Библиографический список

1. *Евдошенко А.П.* К вопросу о применении стереометрической модели в области фонологии // Исследования по структурной типологии. М.: АН СССР, 1963. С. 200 – 208

2. *Иллич-Свитыч В.М.* Опыт сравнения ностратических языков. М.: Наука, 1971. 370 с.
3. *Калужнин Л.А., Скороходько Э.Ф.* Некоторые замечания о лексической семантике. М.: АН СССР, 1963. С. 183 – 199
4. *Саенко М.Н.* Реконструкция праславянского списка Сводеша // Вопросы языкового родства. 2013. №10. С. 139 – 148
5. *Сергиенко М.А.* Методы анализа и структуризации базы нечётких правил. Дисс. к.т.н. Воронеж, 2010. 155 с.
6. *Трофимов А.А.* Опыт составления аннотированного списка Сводеша для ведийского языка // Вопросы языкового родства. 2020. №3. С. 67 – 88.