

УДК 811.512.1:004

Мультимедийные информационные технологии и компьютерная лингвистика в системе языковых процессов мировой цивилизации**Каримов Бахтиёр Рахманович**

доктор философских наук, профессор кафедры общественных наук Ташкентского государственного технического университета. Республика Узбекистан 100095, г. Ташкент, ул. Университетская, 2. E-mail: karimov.bahtiyor@yahoo.com;

Муталов Шоахмад Шоназарович

научный сотрудник Интернационального культурного центра Республики Узбекистан. Республика Узбекистан 100100, Ташкент, ул. Бабура, 4 E-mail: shahahmad.mutal@gmail.com.

Аннотация. Посредством использования методов математической и компьютерной лингвистики предлагается создать усредненные языки на основе усреднения лексических фондов, языковых правил и норм генеалогически родственных языков. Предлагается идея создания среднемирового языка посредством усреднения усредненных и изолированных языков. Рассмотрены социальные, культурные, лингвистические, информационные, коммуникативные проблемы усредненных языков и пути использования мультимедийных технологий в этой сфере.

Ключевые слова: усредненные языки; языковая идентичность; этнолингвизм; математическая и компьютерная лингвистика; мультимедийные и информационные технологии.

Введение. При обмене информацией обменивающиеся ею стороны должны понимать язык, на котором закодирована информация. Язык социальной коммуникации, то есть устный язык (звуковая кодировка информации), письменный язык (текстовая кодировка информации), а также частично и знаковый язык, выработанный для сообщества немых, лежат и в современную эпоху в основе информационного обмена между людьми. Без знания и понимания языка социальной коммуникации в социуме, в котором задействованы мультимедийные технологии и информационные технологии коммуникации эти технологии теряют свою эффективность, так как без языка теряется семиотическая и семантическая целостность системы кодов социальной коммуникации. Поэтому в большинстве исследований мультимедийных технологий и информационных технологий исходят из презумпции знания языка обмена информацией. Однако реальная ситуация в системе мировой цивилизации в корне отличается от этой идеализированной модели, которая таким образом оказывается применимой лишь к социальной группе знающей определенный подразумеваемый в презумпции общий язык. Для приближения к адекватному познанию реальности необходимо выйти за рамки этой идеализированной модели и создать более широкую модель, охватывающую её как частный случай.

Языковая ситуация в мире изучается во многих исследованиях, которые показывают сложность, противоречивость, конфликтность, драматизм, трагичность отношений языковых групп в мировой цивилизации. Хотя мировая цивилизация признаёт, что каждый язык является общечеловеческой ценностью, исследования показывают, что каждые две недели погибает один из живых языков, превращаясь в мертвый язык. Языковая идентичность тесно связана с национальной и этнической идентичностью, поэтому языковые процессы взаимосвязаны с национально-этническими процессами.

Современный ход языкового развития мировой цивилизации разворачивается на основе процесса создания национальных государств, в которых язык так называемой государствообразующей нации признается государственным языком, а языки других наций на этой территории, в

большинстве случаев, приобретают второстепенное положение. Так как большинству наций не удастся создать такое свое национальное государство, то язык этих наций оказывается во второстепенном статусе. Такие нации составляют большинство наций мира, а их языки – большинство языков мира.

1. Постановка проблемы. В статье предложен иной путь решения этой фундаментальной языковой проблемы социальной коммуникации в современной мировой информационной цивилизации, основанный на рациональной научной регламентации информационного обмена и языковых процессов основанный на методах математической и компьютерной лингвистики.

Несмотря на некоторые особенности процесса формирования наций, большинство существующих наций развились из родственных племен и общенациональный язык, литературный язык нации развивался на основе племенных диалектов в результате их взаимодействия. С точки зрения генезиса кажется естественным, что зародыши наций находятся среди родственных племен, населявших соседствующие территории и имевших более тесные коммуникации. Кроме того, чем ближе были их языки, тем легче была коммуникация.

Согласно ойкуменической теории нации (Karimov 2003) такие этнические общности как род, племя, народность и нация следует рассматривать как исторические стадии развития общего родового феномена этнация. Этнация формируется и развивается вместе со своей ойкуменой (Karimov 2003), то есть с системой этний с которыми она взаимодействует, поддерживает коммуникацию.

Человечество в будущем будет единым целым. Стадии развития, когда родственные нации образуют общности более высокого уровня, представляются неизбежными. Но следует иметь в виду, что могут существовать неверные пути консолидации наций, такие как имевшие место в прошлом попытки сформировать единые народы. Проводя политику консолидации наций, следует принимать во внимание, что этническое самосознание является общечеловеческой ценностью.

Этнический лингвистический паннизм, сокращенно этнолингвопаннизм (Karimov 2003) является применением философской концепции потенциального единства к определенной этнической группе. Близость определенной группы народов и сознание того, что люди этой группы могут понимать друг друга в определенной мере, даже если каждый из них знает только свой язык, является основой любого специфического вида этнолингвопаннизма. Подобие процесса коммуникации между людьми, говорящими на диалектах одного языка, и представителями вышеупомянутых народов является гносеологическим основанием возникновения соответствующего паннизма. Если есть возможность, то этнолингвопаннизм использует и такие характеристики, как близость культуры, религии, литературы, происхождения, исторических судеб и так далее этносов, близость чьих языков служила основой при обосновании соответствующего этнолингвопаннизма.

В настоящее время единство человечества необходимо для его выживания, целесообразно сознательно искать пути для формирования и сохранения такого единства. Объединение близких наций в федерации, конфедерации или унитарные государства согласно их волеизъявлению является одним из таких путей.

В многонациональных государствах возникает языковая проблема. В таких государствах должен быть язык, служащий в качестве общего языка общения, называемый языком-посредником. В мире существует много языков-посредников. Модно рассматривать английский язык как мировой язык-посредник, и в мире существуют много регионов с их собственными языками-посредниками. Например, суахили является таким языком в Восточной Африке, хинди – в Индостане, испанский и португальский в – Латинской Америке и т.д. Но национальный язык в качестве языка-посредника имеет, по меньшей мере, два неприемлемых свойства.

Во-первых, с точки зрения справедливости в духовной жизни, т.к. носители других языков будут чувствовать себя ущемленными в своих правах применения их родных языков. Во-вторых, с экономической точки зрения, все народы, кроме носителей языка-посредника, будут отставать на годы из-за необходимости изучать язык-посредник наряду со своими родными языками для получения мировой информации. Таким образом, можно прийти к заключению, что ни один национальный язык не пригоден для роли языка-посредника.

Региональным языком-посредником должен быть язык, в равной мере близкий к языкам народов региона, где предполагается, что он будет языком общения. Язык всемирного общения должен отвечать таким же требованиям. По нашему мнению, региональные языки межнационального общения должны быть созданы на базе групп родственных языков согласно специальному методу, обеспечивающему его нейтральность, а также и оптимальную степень понятности как можно большему числу людей среди носителей этих языков.

Ускорение темпов социальных процессов, формирование системы глобальных проблем человечества приводит к необходимости перехода к ноосфере, в частности, к моделированию и управлению информационными и языковыми процессами. Метод создания языка-посредника должен соответствовать объективному механизму процесса формирования койне в ходе общения носителей родственных языков, имитационно моделировать его. Чем у большего числа: 1) родственных языков и 2) индивидов носителей этих языков встречается определенная форма при реализации некоторого языкового феномена, тем больше вероятность включения этой нормы в систему норм койне.

Механизм, или метод, создания языка-посредника важен, так как в осознании своей этнической принадлежности личностями, считающими своим родным языком родственные языки, важную (хотя и не решающую) роль играет степень близости этих языков, осознание личностью своих языков как самостоятельных языков или как диалектов одного языка. В системах диалектов различных языков мера близости диалектов, приводящая к осознанию носителями диалектов принадлежности к одному языку, существенно различается.

Для лексики необходимо выбрать определенной длины верхнюю часть частотного словаря одного из языков. Этот язык служит базовым. Значения полисемантических слов, которые помечены в словаре основного языка цифрами, считаются отдельными лексемами. Выбранные слова в списке нумеруются. Для каждой лексемы упорядоченного списка должны быть найдены эквиваленты в языках или диалектах, на базе которых строится усредненный язык.

2. Концепция решения проблемы. Метод создания языка-посредника должен соответствовать объективному механизму процесса

формирования койне в ходе общения носителей родственных языков, имитационно моделировать его. Чем у большего числа: 1) родственных языков и 2) индивидов носителей этих языков встречается определенная форма при реализации некоторого языкового феномена, тем больше вероятность включения этой нормы в систему норм койне.

3. Реализация концепции

Сутью предложенного метода (Karimov, Mutalov 1982; 1992; 1993; 2008) является выявление общего фонда для группы диалектов или языков посредством специальной математической процедуры и кодификация данного общего фонда в качестве стандарта, т. е. как единицы соответствующего уровня создаваемого языка. Процедура следующая: единицам уровней языка сопоставляются векторы. Как правило, единицы уровней имеют варианты: различные значения слов, позиционные варианты фонем и морфем, различные способы выражения синтаксических отношений. Векторное пространство, сопоставленное единицам уровня, построено таким образом, что главное значение единицы сопоставлено со значением 1, другие варианты соответствуют числам, меньше 1, согласно формуле

$$X_{n,\alpha}^i = 1 - \frac{i-1}{2s}, \quad (1)$$

где "n" - номер единицы в списке единиц; "α" - определенный язык; "i" - номер варианта единицы уровня с номером "n"; "s" - число вариантов уровня с номером "n". Эти числа образуют набор компонентов вектора. Мы назвали этот метод методом усреднения, а языки, построенные таким способом, - усредненными.

Следующим шагом является упорядочение корней в соответствии со значениями данной функции.

Затем таблицы простых списков слов нужно преобразовать в, где однокоренные слова-эквиваленты расположены на одной строке.

Для однокоренных слов вводится функция по формуле

$$F_n^j(\bar{x}) = \sum_{\alpha=1}^A \left(1 + \frac{K_\alpha}{K} \right) X_{n,\alpha}^j, \quad (2)$$

где A - полное число усредняемых языков, K_α - численность носителей языка α, $\bar{K}$ - среднее арифметическое носителей языка одной группы родственных языков, которое вычисляется как частное от деления общего числа носителей всех языков группы на число языков в группе, j означает порядковый номер однокоренных слов в упорядоченный в таблицах.

Роль коэффициента $1 + \frac{K_\alpha}{K}$ - это учет численности носителей языка в группе, для которой строится усредненный язык.

Корни, помещенные на первом месте, рекомендуются для включения в словарь усредненного языка в качестве основных лексических единиц. Другие единицы могут формировать запас слов-синонимов для обогащения словаря усредненного языка при необходимости.

Вся работа может быть выполнена при помощи компьютеров. Преимущества предлагаемого решения мировой языковой проблемы заключаются в следующем:

1) так как усреднение проводится на базе этнических языков (диалектов), носители данных языков (или диалектов), на основе которых строится усредненный язык, будут понимать его в определенной мере без предварительного изучения;

2) не принадлежа ни одной этнической группе, усредненный язык не дает преимущества ни одной из них, так что он не будет способствовать национальной розни на основе языковой политики;

3) усредненный язык устраняет некоторый произвол в выборе одного из местных языков в качестве официального государственного языка, так же как межэтнические конфликты;

4) усредненный язык позволяет не вводить язык бывших колонизаторов в качестве единственного официального государственного языка, ослабляет зависимость от бывшей метрополии в сфере культуры и образования.

5) многие народы, говорящие на одном языке разделены государственными границами. Таким образом, усредненный язык, построенный предложенным методом, мог бы играть роль макропосредника.

Проблемы этнической консолидации различных диалектных групп в основном связаны со степенью их близости. Степень близости, когда носители диалекта считают себя говорящими на одном языке, сильно различается, т. е. в некоторых случаях носители не чувствуют себя относящимися к одному и тому же языку, хотя существует объективная близость диалектов, не осознаваемая носителями. Для объективной, научно-обоснованной оценки этой близости языков и диалектов целесообразно использовать метод количественной оценки их близости (Karimov, Mutalov 1985).

Здесь можно обратить внимание на еще одно преимущество усредненных языков – они могут оказаться спасением для исчезающих языков, многие из которых являются родственными языками малочисленных народов и диалектами.

В связи с созданием усредненных нейтральных языков для родственных языков возникает вопрос о перспективах развития коммуникации между носителями неродственных языков в мировом масштабе, глобальной коммуникации всех народов.

Заключение. Мы полагаем, что всемирный вспомогательный язык может быть создан тем же путем, что и усредненные языки для групп родственных языков, и он будет следующим шагом в нейтрализации языков. Создание всемирного вспомогательного языка важно для решения проблемы информационных барьеров, так же как и необходимости преодоления информационного колониализма, и необходимости достижения равенства народов в условиях, когда существует клуб так называемых мировых языков, и стремление установить новый информационный порядок. Эта актуальная глобальная проблема может быть решена группой специалистов под эгидой ООН. При создании искусственного всемирного языка рекомендуется принять во внимание факты всех языков, и в связи с этим использование среднемирового языка создаваемого методом усреднения имеет то преимущество, что обеспечивается нейтральность языка и равноправие языков народов мира (Karimov, Mutalov 1982; 1992; 1993; 2008).

Развитие мультимедийных технологий обеспечивает их применение во все более широких областях человеческой деятельности. В частности, в обмене информацией и образовании (дистанционное образование, международное, непрерывное образование). Эти сферы необыкновенно тесно связаны с употреблением языка, язык, в свою очередь, связан с самосознанием и

национальной самоидентификацией. Для обеспечения равноправия наций и локальных цивилизаций в языковом аспекте целесообразно развивать усредненные языки, создать методом усреднения среднемировой язык и внедрять их в сферы обмена информацией и образования.

Әдебиеттер тізімі/ Список литературы

1. Каримов Б.Р., Муталов Ш.Ш. К вопросу о языковой политике в многоязычных развивающихся странах // Тезисы докладов республиканской научно-теоретической конференции молодых ученых и специалистов по общественным наукам (Ташкент, 2-3 марта 1982 г.). – Ташкент, 1982. – С.94-95.
2. Каримов Б.Р., Муталов Ш.Ш. О количественной оценке синхронической близости родственных диалектов и языков// Тюркское языкознание: материалы международной конференции. - Ташкент, 1985.
3. Каримов Б.Р., Муталов Ш.Ш. Усредненные языки: попытка решения мировой языковой проблемы. – Ташкент: Фан, 2008.
4. Каримов Б.Р., Муталов Ш.Ш. Ўрта турк тили. – Ташкент, 1992.
5. Karimov B.R., Mutalov Sh.Sh. Averaged languages: an attempt to solve the world language problem. Tashkent: Fan, 1993 (2008).
6. Karimov B.R. The oikumenic concept of the nation and development of languages. Ойкуменическая концепция нации и развитие языков. – Qarshi, 2003.

Reference

- Karimov 2003 – Karimov, BR 2003, *The oikumenic concept of the nation and development of languages. Oikumenicheskaya koncepciya natsii i razvitie yazykov*, Qarshi. (in Rus / inEng).
- Karimov, Mutalov 1982 – Karimov, BR, Mutalov, ShSh 1982, K voprosu o yazykovoj politike v mnogoyazychnyh razvivayushchihya stranah, *Tezisy dokladov respublikanskoj nauchno-teoreticheskoy konferencii molodyh uchenykh i specialistov po obshchestvennym naukam*. Tashkent, 2-3 marta 1982 g., Tashkent, P.94-95. (in Rus).
- Karimov, Mutalov 1985 – Karimov, BR, Mutalov, ShSh 1985, O kolichestvennoj ocenke sinhronicheskoy blizosti rodstvennykh dialektov i yazykov, *Tyurkskoe yazykoznanie. Materialy mezhdunarodnoj konferencii*, Tashkent. (in Rus).
- Karimov, Mutalov 1992 – Karimov, BR, Mutalov, ShSh 1992, *Ўрта турк тили*, Toshkent. (in Uzb).
- Karimov, Mutalov 1993 – Karimov, BR, Mutalov, ShSh 1993 (2008), *Averaged languages: an attempt to solve the world language problem*, Fan, Tashkent. (in Eng).
- Karimov, Mutalov 2008 – Karimov, BR, Mutalov, ShSh 2008, *Usrednennye yazyki: popytka resheniya mirovoj yazykovoj problemy*, Fan, Tashkent. (in Rus).

Мультимедиялық ақпараттық технологиялар және дүниежүзілік өркениеттің тілдік үрдістер жүйесіндегі компьютерлік лингвистика

Каримов Бахтиёр Рахманович

философия ғылымдарының докторы, Ташкент мемлекеттік техникалық университетінің қоғамдық ғылымдар кафедрасының профессоры. Өзбекстан Республикасы 100095, Ташкент қ., Университет көшесі, 2. E-mail: karimov.bahtiyor@yahoo.com.

Муталов Шоахмат Шоназарович

Өзбекстан Республикасы Интернационалдық мәдени орталықтың ғылыми қызметкері. Өзбекстан Республикасы 100100, Ташкент., Бабур көшесі, 4 E-mail: shahahmad.mutal@gmail.com.

Түйін. Математикалық және компьютерлік лингвистика әдістерін қолдану арқылы генеалогиялық туыс тілдердің тіл ережелері және нормаларын, лексикалық қорларды орташыландыру арқылы орташыландырылған тілдер құру ұсынылады. Орташаландырылған және алшақтатылған тілдерді орташыландыру арқылы орташа әлемдік тіл құру идеясы ұсынылады. Орташаландырылған тілдердің әлеуметтік, мәдени, лингвистикалық, ақпараттық-коммуникативтік мәселелері және осы салада мультимедиялық технологияларды қолдану жолдары қарастырылды.

Түйін сөздер: орташаландырылған тілдер; тілдік ұқсастық; этнолингвопанизм; математикалық және компьютерлік лингвистика; мультимедиялық және ақпараттық технологиялар.

Multimedia information technology and computational linguistics in the system of linguistic processes of world civilization

Karimov Bakhtiyor Rakhmanovich

Doctor of Philosophy, Professor of the Department of Social Sciences, Tashkent State Technical University. 100095 Republic of Uzbekistan, Tashkent, 2 Universitetskaya str. E-mail: karimov.bahtiyor@yahoo.com

Mutalov Sh.Sh.

researcher of the International Cultural Center of the Republic of Uzbekistan. Republic of Uzbekistan 100100, Tashkent, 4, Babura str. E-mail: shahahmad.mutal@gmail.com

Abstract. By using mathematical methods and computational linguistics, it is proposed to create averaged languages by averaging lexical fund, language rules and norms of genealogically related languages. The idea of creating the world averaged language by averaging averaged and isolated languages is proposed. Social, cultural, linguistic, information, communicative problems of averaged languages and ways of using multimedia technologies in this field are considered.

Keywords: averaged languages; linguistic identity; etno-linguapанизm; mathematical and computational linguistics; media and information technology.