

RESEARCH

ЗВУКОВАЯ КАРТИНА ЯЗЫКА КАК СОВОКУПНОСТЬ ЦЕПОЧЕК ЗВУКОВ В РЕЧИ. ЧАСТЬ 1

Ю.А. Тамбовцев

Новосибирск, Россия

Summary: The article deals with the problem of the investigation of the sound picture of a language. A sound picture of a language is understood as the sum of the frequency of phonemes in its speech sound chain. It is possible to solve some linguistic problems only using certain statistical methods. Quantitative methods can contribute much to the investigation of the distances from Russian to some other Slavonic languages on the basis of their sound pictures. A huge sample of more than 20 million sounds, computed by the author, represents the sound picture of Russian.

Целью данной работы является изучение построения звуковой картины языка. По звуковой картине языка можно судить о месте этого языка среди других языков мира. Звуковая картина языка – это совокупность частоты встречаемости каждого элемента в звуковой цепочке, которая порождается в процессе речи. Звуковая картина каждого языка уникальна, так как она отличает один язык от другого. Процессы гомеостаза поддерживают постоянство частоты встречаемости фонем в звуковой цепочке языка. Любой естественный человеческий язык саморегулируется, что позволяет ему оставаться идентичным самому себе. Частоты звуков ученые объединяют в определенные классы и группы, чтобы понять фонологическую типологию языков с точки зрения работы артикуляторного аппарата человека [Тамбовцев 2006].

Звуковая картина языка связана с типологией упорядоченности звуковых цепочек [Тамбовцев 2003]. Типология упорядоченности звуковых цепей в языке в данном случае изучается при помощи методов фонологической статистики. Известно, что одной из основных задач фонологической статистики является нахождение, а затем и объяснение закономерностей частоты функционирования отдельных фонем, их групп и классов в потоке речи различных языков [Тамбовцев 1976, 1977, 1983, 2003]. В данной работе на материале русского языка будет более детально рассмотрено функционирование групп фонем в различных стилях, которые в совокупности составляют звуковую картину русского языка.

На наш взгляд, функционирование в языке фонемных групп может считаться универсальным признаком в связи с тем, что все языки мира, то

есть языки различного генетического происхождения и различного морфологического строя, в том или ином виде имеют основные типы (группы) фонем [Зиндер 1979; Maddieson 1981]. Исследования как по общей фонетике, так и по фонетике отдельных языков показывают, что наличие нескольких противопоставленных друг другу классов и групп фонем характерно для любого естественного человеческого языка, несмотря на то, что грамматические структуры, словарные запасы и фонологические системы у них различны [Тамбовцев 2001]. Разделение звуков речи на группы возникло уже в трудах ученых древности. Так, древнеиндийские, арабские (например, Авиценна за несколько сотен лет до н.э.) и древнегреческие ученые разделяли звуки речи на несколько категорий (групп) и довольно четко их определяли [Кодухов 1974: 61].

В данной работе определение принадлежности каждой фонемы к какой-то определенной группе имеет не только теоретическое, но и огромное прикладное значение. Неверное определение наполнения групп приведет к неверным выводам.

Впервые в типологические исследования вводится понятие звуковой картины языка и на ее основе изучается компактность языковых семей с точки зрения фоностатистических характеристик [Тамбовцев 1990]. Также впервые на основе звуковой картины языка мы предлагаем измерять расстояния между языками, исходя из типологии функционирования в их звуковых цепочках групп фонем. При этом основополагающим принципом будет следующий: чем больше общих признаков и чем более схожи их частоты, тем более сходны звуковые картины языков с типологической точки зрения.

Совокупность звуковых цепочек в языке как характеристика языка в целом

Любой исследуемый объект (в данном случае – звуковая картина языка) может быть охарактеризован некоторым конечным набором признаков, которыми в данном случае будут количественные характеристики функционирования в звуковой цепи некоторых фонемных групп, выделенных по работе активного органа, по характеру шумообразующей преграды и по работе голосовых связок. По этим количественным признакам можно характеризовать любой язык [Тамбовцев 1977]. Частота встречаемости 8 признаков согласных является той базой, на которой основано наше распознавание того или иного языка. Совокупность взятых 8 признаков образует пространство признаков, которое в данном случае охватывает все основные характеристики согласного звука с точки зрения его артикуляции в процессе речепроизводства. Размерность признакового пространства равна числу признаков – в нашем случае групп согласных. Следовательно, в данном случае это пространство является многомерным, а точнее восьмимерным.

Весь данный набор признаков характеризует в данном многомерном пространстве реально существующий язык как совокупность определенных черт типологии его звуковых цепей. Фактически, этот набор типоло-

гических черт образует вполне определенную точку в указанном многомерном пространстве. Языки сопоставляются друг с другом как точки этого пространства. Эти точки образуют в данном пространстве некоторые компактные совокупности (группы) точек. Мы их называем кластерами или таксонами. Кроме того, исходя из терминов и традиций лингвистики, такую компактную совокупность точек, образующую в пространстве относительно изолированную область, можно назвать на первом шаге языковой подгруппой (ветвью), на втором шаге – группой, на третьем – семьей, а на четвертом – общностью. Таким образом, для построения кластеров языков необходимо определить звуковую картину каждого языка. Можно сказать, что звуковая картина языка – это звуковой образ языка, и применить известные методы распознавания образов [Загоруйко 1972: 79, 21-24]. На наш взгляд, применение методов распознавания образов в языкознании может дать интересные результаты [Тамбовцев 1976]. Известно, что при неудачном выборе лингвистических признаков все языки или могут собраться в одной точке пространства, или же каждый язык составит отдельную группу. Ясно, что в первом случае мы задали признаки недостаточные, а во втором случае – избыточные для правильной классификации. Как мы увидим далее, у нас этого не происходит. Кроме того, мы взяли непересекающиеся признаки. Они являются непересекающимися потому, что первые четыре признака выделены по работе активного органа, следующие три признака выделены по характеру шумообразующей преграды, а последний признак характеризует работу голосовых связок. В языке нет ничего другого, что могло бы считаться еще одним непересекающимся признаком. Укажем еще раз, что наши признаки являются всеобъемлющими, в то же время они дополняют друг друга до полной системы признаков.

Далее мы покажем, как строить звуковую картину русского языка, которая нам будет нужна для сравнения с другими звуковыми картинами при построении языковых таксонов, например – семей [Тамбовцев 2005]. Проанализировав по указанным восьми признакам какой-либо язык (например, русский), мы далее будем сравнивать с ними звуковые картины тех языков, которые уже были отнесены к этой группе, т.е. славянские языки. Мы измеряем это расстояние. Затем берем какой-либо другой язык (например, украинский) и по очереди сравниваем со всеми языками славянской группы. Затем сравниваем, какой язык ближе к другим славянским языкам: русский или украинский. Тем самым мы устанавливаем ядерные и периферийные славянские языки.

Таким же образом мы поступаем со всеми языками, которые до сих пор не были бесспорно отнесены к той или иной группе или семье языков. Мы вычисляем точку в многомерном пространстве, которую занимает, например, японский язык, а затем устанавливаем расстояния от этой точки до той или иной совокупности точек, которые и составляют определенную семью языков. По величине этих расстояний можно будет предполагать, к какой языковой семье ближе всего японский язык по выбранным типологическим признакам.

Фонемные группы в качестве признаков описания звуковой картины языка

В данной работе мы базируемся на том, что фонетическая система – это наиболее консервативная составляющая любого языка, поскольку она зависит от артикуляционной базы, обычно сохраняющейся даже тогда, когда народ переходит на другой язык. Примером того, что фонетическая система остается неизменной, могут служить факты проживания населения в условиях двуязычия.

Процесс вычисления расстояний между языками мира включает, по нашему мнению, как минимум три основных этапа:

- 1) получение данных по частоте встречаемости определенных фонемных групп;
- 2) определение признаков, которые являются базовыми при построении межъязыковых расстояний;
- 3) построение решающего правила, то есть определения математических процедур, после применения которых мы получим искомые межъязыковые расстояния.

Наиболее трудоемкий этап – это этап полевого сбора материала и, затем, получения фонологической транскрипции текстов на различных языках. Следующий этап – это ввод фонемной транскрипции в ЭВМ. При этом выборка не должна быть маленькой [Тамбовцев 1985]. В принципе, описание любого языка – это бесконечный процесс. Мы ограничиваемся неполным описанием языка, так как не привлекаем его лексикологические, морфологические, синтаксические, семантические и другие признаки. Таким образом, приходится говорить только о фонетико-типологическом описании языка, которое, однако, является достаточным для распознавания языка и отделения одного языка от другого. Мы будем рассматривать расстояние как метрическую меру близости с точки зрения типологии функционирования фонемных групп, которые выделяются по работе активного органа речеобразования, по характеру шумообразующей преграды и по работе голосовых связок.

Таковыми базовыми признаками в нашем исследовании являются, прежде всего, признаки классов, а затем признаки групп. Признаками классов будет считаться частота встречаемости в речи гласных и согласных. Мы пока не будем рассматривать класс гласных. Для этого есть несколько причин, о которых далее мы говорим более подробно. Класс согласных в свою очередь разбивается на несколько групп.

По работе действующего органа выделяются следующие фонемные группы согласных:

- 1) губные; 2) переднеязычные; 3) среднеязычные, или палатальные;
- 4) заднеязычные, или веларные.

Ранее классификацию согласных по работе активного органа называли классификацией по месту артикуляции. На наш взгляд, правильнее говорить о работе активного органа речеобразования.

По характеру шумообразующей преграды или способу образования мы выделяем:

1) сонорные; 2) смычные шумные; 3) щелевые шумные.

По работе голосовых связок мы выделяем звонкие шумные. Фактически частота шумных глухих является дополняющей, поэтому можно брать только частоту шумных звонких. Обычно дополняющая характеристика ничего не добавляет при построении кластера или измерении расстояния.

Членение фонемной цепи на фонемные группы

Вначале фонемная цепь делится на согласные и гласные. Установлено, что в большинстве языков мира согласные имеют больший информационный вес и поэтому являются базой структуры слова. В связи с тем, что согласные несут основную семантическую нагрузку, гласные во многих языках в значительной мере избыточны (с точки зрения положений теории информации), так как слова могут быть опознаны по одним согласным. Специальные эксперименты показывают, что при сильном шуме и при искажениях для правильного восприятия слова решающее значение имеет правильная передача согласных, в то время как передача гласных играет второстепенную роль [Зиндер 1979: 111].

Гласные несут, в основном, служебную модифицирующую информацию: они служат неким фоном для разбавления согласных. Это подтверждается их частотностью. В большинстве языков в звуковой цепочке больше согласных, чем гласных [Тамбовцев 1979: 36-38; 2001].

Фоно-статистические характеристики помогают различать даже языки с идентичными номенклатурами фонем. Как совершенно справедливо замечает В.А.Никонов, два языка могут иметь даже одинаковый фонемный инвентарь и, тем не менее, различаться по частоте звуков. В то же время он подчеркивает, что между родственными языками сохраняется «статистическая близость» многих фонетических явлений [Никонов 1963: 259].

Северный диалект мансийского и казымский (северный) диалект хантыйского языков имеют почти идентичный состав фонемной номенклатуры. В то же время по частоте встречаемости фонем они все-таки различаются, хотя их частотные распределения фонем более близки друг другу, чем аналогичным распределениям фонем в других финно-угорских языках [Тамбовцев 1977, 1982].

Звуковая картина языка показывает свое отличие уже на уровне фонемных классов [Тамбовцев 1994]. Частотность двух основных классов (согласных и гласных) в двух разных языках говорит о различной частоте встречаемости. Эти различия показывают величины их консонантных коэффициентов. Это позволяет говорить о них как о различных языках. Если разница в величине этих коэффициентов небольшая, то можно говорить об этих языках как о типологически сходных. Часто это проявляется в близкородственных языках, например в русском и белорусском или в русском и украинском. Небольшая разница в звуковой картине языка говорит о большей близости каких-либо языков друг к другу, чем к другим языкам.

Академик Лев Владимирович Щерба выделял в русском языке следующие фонемы [Щерба 1963: 63]:

гласные: [i, u, l, e, o, a];

согласные: [p, p', b, b', v, v', f, f', m, m', t, t', d, d', ts, tʃ, ʃ, ʃ', ž, s, s', z, z', n, n', l, l', r, r', j, k, k', g, g', x, x'].

Классификация русских согласных по работе активного органа речеобразования (месту образования):

губные: [p, p', b, b', m, m', f, f', v, v'];

переднеязычные: [t, t', d, d', ts, n, n', s, s', z, z', tʃ, ʃ, ʃ', ž, l, l', r, r'];

среднеязычные: [j];

заднеязычные: [k, k', g, g', x, x'].

Классификация русских согласных по характеру шумообразующей преграды:

сонорные: [m, m', n, n', l, l', r, r', j];

шумные смычные: [b, b', p, p', t, t', d, d', k, k', g, g', tʃ, ts];

шумные щелевые: [v, v', f, f', s, s', z, z', ʃ, ʃ', ž, x, x'].

Классификация русских согласных по работе голосовых связок:

звонкие: [b, b', v, v', d, d', z, z', ž, g, g'].

Вероятно, А.М.Пешковский был первым из русских лингвистов, кто обратил внимание на частоту встречаемости звуков в русском языке как на то, что способно вызывать эстетическую оценку [Пешковский 1930: 134]. Он подсчитал частоту встречаемости фонем русского на материале художественной прозы на выборке объемом 10 000 фонем [Пешковский 1930: 167-191]. К сожалению, в столь маленьких выборках частоты групп согласных достаточно сильно колеблются, поэтому к данным этой выборки следует относиться с большой осторожностью. Сумма гласных по данным Пешковского – 42,54%. Величина отношения согласных к гласным (КК) – 1,35.

Только в 1964 г. появились новые данные по частоте встречаемости русских фонем. Они были получены В.Н.Елкиной и Л.С.Юдиной, которые обработали выборку объемом 94 000 слогов, что составляет примерно 235 000 фонем [Елкина и др. 1964: 58-78]. Они взяли три, по их мнению, наиболее важных и употребительных стиля: 1) художественный; 2) общественно-публицистический и 3) научно-технический. Эти три стиля были представлены в следующем количественном соотношении: художественный – 50%; общественно-публицистический – 25% и научно-технический – 25%. В качестве примера художественного стиля В.Н.Елкина и Л.С.Юдина взяли рассказы русских писателей: «Мещерский край» К.Паустовского, «Ионыч» А.П.Чехова, «Санта-Мария» В.Некрасова, «Почем фунт лиха» Г.Бакланова и «Три весны» Б.Ямпольского. Общественно-публицистический стиль был представлен докладом Н.С.Хрущева на 22 съезде КПСС. Научно-технический стиль представляли несколько статей из журнала «Электроника» № 45 за 1962 г. Сумма гласных – 42,13%. Величина КК – 1,37. Объем выборки составил 94 000 слогов.

Вероятно, потребность в надежной статистике русских фонем стала актуальной, так как в следующем 1965 г. А.В.Баклушин получил статистику на объеме 100 000 фонем [Баклушин 1965: 39-48]. К сожалению, он не указал стиль, на котором были произведены подсчеты. Сумма гласных – 42,29%. Величина КК – 1,36.

Достаточно большую выборку тоже объемом 100 000 фонем посчитали в 1968 г. Генри Кучера и Джордж Монроу. Их выборка состояла из художественной прозы XX века (60%), публицистики (20%), поэзии (10%) и научной литературы (10%). Нам кажется, что к их результатам нужно относиться с осторожностью, так как их компьютерная программа обрабатывала русский текст достаточно примитивным образом, не учитывая многих фонетических явлений (например, ассимиляцию, диссимиляцию и т.д.). Сумма гласных – 41,85%. Величина КК – 1,39.

В 1969 г. В.Н.Елкина, Л.С.Юдина и А.Г.Хайретдинова опубликовали статистику двухфонемных сочетаний (диад), которую мы пересчитали для наших целей, получив объем выборки в 1 269 419 фонем. Материалом послужили научно-технические тексты: А.Файнштейн «Основы теории информации», «Применение вычислительной техники для автоматизации производства»; Ж.Кампа де Ферье и др. «Функции математической физики»; Г.Грениевский «Кибернетика без математики», «Проблемы автоматического распознавания образов»; П.И.Ченцов «Основы теории информации»; Б.П.Асеев «Фазовые соотношения в радиотехнике»; А.П.Ершов и др. «Входной язык системы автоматического программирования»; О.Г.Чернов и др. «Элементарная теория статистических решений», «Вычислительные системы» (№ 10) и другие. Все математические формулы и символы опускались [Елкина и др. 1969: 48-72]. Сумма гласных – 496 991 (39,14%). Величина КК – 1,55. Выборка состояла только из технических текстов.

В 1977 г. Л.В.Бондарко, Л.Р.Зиндер и А.С.Штерн обработали достаточно большую выборку (100 000 фонем) на материале нескольких стилей: 1) художественная проза; 2) поэзия; 3) публицистика; 4) научная литература; 5) устная речь [Бондарко и др. 1977: 5]. Сумма гласных – 42,00%. Величина КК – 1,38.

Мы подсчитали частоту встречаемости русских фонем на материале устной диалогической речи объемом 31 000 фонем [Гамбовцев 2001]. Сумма гласных – 13 100 фонем (42,19%). Величина КК – 1,37.

Нами были также обработаны прозаические тексты из произведений современной фантастики, рассказы М.Шолохова, романы М.Шолохова, Б.Пастернака, М.Булгакова, Ю.Семенова и других современных писателей, а также проза писателей XIX века (М.Ю.Лермонтова, Л.Н.Толстого, М.Горького, Н.В.Гоголя, Ф.М.Достоевского). Выборка постепенно наращивалась. Начнем с наименьшей выборки объемом 69 315 фонем. Сумма гласных – 28 905 фонем (41,71%). Величина КК – 1,40.

Выборка большего объема показала примерно такие же результаты. Сумма гласных – 132 974 фонем (41,91%). Величина КК – 1,39.

Еще большая выборка по прозе XIX века и современной прозе объемом 1 826 649 фонем показала похожие результаты. Сумма гласных – 765 871 фонем (41,93%). Величина КК – 1,38.

Мы обработали очень большую выборку (9 045 000 фонем) по всем книгам А.Марининой, которые представили современный русский язык. Сумма гласных – 3 861 319 фонем (42,69%). Величина КК – 1,34.

Мы также подсчитали частоту встречаемости фонем русского языка в поэзии, прозе, письмах и других произведениях А.С.Пушкина, которые вошли в полное собрание его сочинений, на объеме 8 918 000 фонем. Сумма гласных – 3 744 357 фонем (41,99%). Величина КК – 1,38.

На наш взгляд, правильную звуковую картину русского языка может составить только совокупность стилей и совокупность выборок. Обобщенные результаты в наших исследованиях по звуковой картине русского языка составили данные указанных выше ученых (А.М.Пешковского, Л.В.Бондарко, Л.Р.Зиндера, А.С.Штерн, А.В.Баклушина, В.Н.Елкиной, Л.С.Юдиной, Генри Кучера, Джорджа Монроу, З.Ф.Оливериуса) и наши собственные данные по нескольким стилям. Сумма гласных – 42,18%. Величина КК – 1,37. Объем обобщенной выборки по русскому языку по всем стилям – 20 189 475 фонем. Таким образом, мы получили звуковую картину русского, которая может считаться самой надежной при сравнении с другими языками.

Табл. Частота встречаемости групп согласных в звуковой цепочке русского языка в процентах ко всем звукам в этой цепочке. Разные стили

Автор	Год	Губ- ные	Перед- неяз.	Сред- неяз.	Зад- неяз.	Сонорн.
Пешковский	1930	11,66	36,03	4,57	5,20	20,65
Елкина и др.	1964	12,87	35,49	3,81	5,70	22,94
Баклушин	1965	12,63	35,51	4,73	4,84	22,43
Кучера и др.	1968	12,37	35,64	4,14	6,00	22,69
Елкина и др.	1969	13,36	37,60	5,05	4,85	25,39
Бондарко	1977	12,29	36,30	3,66	5,75	22,61
Тамбовцев Устн. речь	2001	12,60	34,62	4,93	5,64	22,35
Тамбовцев ХудПроза-1	2001	13,29	35,71	3,77	5,52	22,62
Тамбовцев ХудПроза-2	2001	13,37	35,60	3,51	5,61	21,98
Тамбовцев ХудПроза-3	2001	13,24	35,75	3,75	5,33	22,72
Тамбовцев ХудПроза-4	2001	11,88	35,75	4,17	5,51	22,56
Тамбовцев ХудПроза-5	2001	12,46	34,66	4,82	6,07	23,37
Все выборки		12,63	35,18	4,27	5,74	23,07
Автор	Год	Смч.	Щел.	Звон.	Сум. согл.	Объем выборок
Пешковский	1930	21,69	15,12	11,79	57,46	10 000
Елкина и др.	1964	19,89	15,04	11,74	57,87	235 000
Баклушин	1965	18,49	16,79	12,51	57,71	100 000

Кучера и др.	1968	20,31	15,15	10,80	58,15	100 000
Елкина и др.	1969	20,75	14,72	10,68	60,86	1 269 419
Бондарко	1977	20,96	14,43	12,42	58,00	100 000
Тамбовцев Устн. речь	2001	21,50	13,96	11,84	57,81	31 050
Тамбовцев ХудПроза-1	2001	20,60	15,07	11,56	58,29	69 315
Тамбовцев ХудПроза-2	2001	20,81	15,30	12,39	58,09	317 311
Тамбовцев ХудПроза-3	2001	20,36	14,99	12,23	58,07	1 826 649
Тамбовцев ХудПроза-4	2001	20,77	13,98	11,09	57,31	9 045 000
Тамбовцев ХудПроза-5	2001	20,10	14,54	11,42	58,01	8 918 000
Все выборки		20,14	14,61	11,69	57,82	20 189 475

Выборка в 20 189 475 фонем по всем стилям русского языка дает нам уверенность, что получены надежные результаты по звуковой картине русского языка.

ЛИТЕРАТУРА

- Баклушин 1965 – *Баклушин А.В.* Статистико-комбинаторное выделение первого морфологического типа в русском языке // Статистико-комбинаторное моделирование языков. М.-Л.: Наука, 1965.
- Бондарко и др. 1977 – *Бондарко Л.В., Зиндер Л.Р., Штерн А.С.* Некоторые статистические характеристики русской речи // Слух и речь в норме и патологии. Л.: ЛВВОГ, 1977.
- Елкина и др. 1964 – *Елкина В.Н., Юдина Л.С.* Статистика слогов в русской речи // Вычислительные системы, вып. 10. Новосибирск, 1964.
- Елкина и др. 1969 – *Елкина В.Н., Юдина Л.С., Хайретдинова А.Г.* Статистика двух- и трехфонемных сочетаний русской речи // Вычислительные системы, вып. 37. Новосибирск: ИМ СОАН, 1969.
- Загоруйко 1972 – *Загоруйко Н.Г.* Методы распознавания и их применение. М.: Сов. радио, 1972.
- Зиндер 1979 – *Зиндер Л.Р.* Общая фонетика. М.: Высшая школа, 1979.
- Кодухов 1974 – *Кодухов В.И.* Общее языкознание. М.: Высшая школа, 1974.
- Никонов 1963 – *Никонов В.А.* Фоно статистика как средство определения языкового родства // Конференция по структурной лингвистике, посвященная базисным проблемам фонологии, 20-23 мая 1963 г. Тезисы докладов. М.: Изд-во АН СССР, 1963.
- Пешковский 1930 – *Пешковский А.М.* Принципы и приемы стилистического анализа и оценки художественной прозы // Вопросы методики родного языка, лингвистики и стилистики. М.-Л.: Гос. изд., 1930.

- Тамбовцев 1976 – Тамбовцев Ю.А. Распознавание фонем человеком. Эмпирическое предсказание и распознавание образов // Вычислительные системы, вып. 67. Новосибирск: Институт математики СО АН СССР, 1976.
- Тамбовцев 1977 – Тамбовцев Ю.А. Некоторые характеристики распределения фонем мансийского языка // Советское финно-угроведение, том XIII, № 3, 1977.
- Тамбовцев 1979 – Тамбовцев Ю.А. Распределение гласных фонем в звуковых последовательностях мансийского языка (по материалам поэзии) // Известия Сибирского отделения АН СССР. Сер. Обществ. наук, вып. 3, 1979, № 11.
- Тамбовцев 1982 – Тамбовцев Ю.А. Эмпирическое распределение частотности фонем в казымском диалекте хантыйского языка // Ученые записки Тартуского ГУ, № 628, 1982.
- Тамбовцев 1983 – Тамбовцев Ю.А. Эмпирическое распределение частотности фонем в орочском языке // Ученые записки Тартуского ГУ, № 658, 1983.
- Тамбовцев 1985 – Тамбовцев Ю.А. Устойчивость консонантного коэффициента в зависимости от величины выборки в некоторых финно-угорских языках // Исследование языков народов СССР. Новосибирск: НГУ, 1985.
- Тамбовцев 1990 – Тамбовцев Ю.А. Компактность финно-угорской языковой семьи по данным консонантного коэффициента // Linguistica Uralica, vol. 26, № 1, 1990.
- Тамбовцев 1994 – Тамбовцев Ю.А. Динамика функционирования фонем в звуковых цепочках языков различного строя. Новосибирск: НГУ, 1994.
- Тамбовцев 2001 – Тамбовцев Ю.А. Компендиум основных статистических характеристик функционирования согласных фонем в звуковой цепочке английского, немецкого, французского и других индоевропейских языков. Новосибирск: Новосибирский классический институт, 2001.
- Тамбовцев 2003 – Тамбовцев Ю.А. Типология функционирования фонем в звуковой цепочке индоевропейских, палеоазиатских, урало-алтайских и других языков мира: компактность подгрупп, групп, семей и других языковых таксонов. Новосибирск: Сибирский Независимый институт, 2003.
- Щерба 1963 – Щерба Л.В. Фонетика французского языка. М.: Высшая школа, 1963.
- Kuchera et. al. 1968 – Kuchera H., Monroe G.K. A comparative phonology of Russian, Czech and German. New York: Elsevier, 1968.
- Maddieson 1981 – Maddieson I. UPSID: UCLA Phonological Segment Inventory Database // UCLA Working Papers in Phonetics, № 53, November, 1981.
- Tambovtsev 2005 – Tambovtsev Yuri. Language Taxons and the Naturalness of their Classification // California Linguistic Notes, vol. XXX, № 2, Fall 2005.

- Tambovtsev 2006 – *Tambovtsev Yuri*. Degrees of Dispersion of Various Language Taxa as the Index of Their Typological Similarity // *Favete Linguis. Studies in Honour of Viktor Krupa*. Bratislava: Institute of Oriental Studies of Slovak Academy of Sciences, 2006.
- Tambovtsev 2007 – *Tambovtsev Yuri*. How can typological distances between Latin and some Indo-European language taxa improve its classification? // *The Prague Bulletin of Mathematical Linguistics*, Number 88, December, 2007.
- Tambovtsev 2008 – *Tambovtsev Yuri*. Typological proximity of Esperanto to the classical language taxa of the Indo-European family // *California Linguistic Notes*, Vol. XXXIII, № 1, Winter, 2008.