

Перестановочный тест частично подтверждает существование алтайской языковой макросемьи

Публикуем обзор статьи (Kassian et al., 2021) [Permutation test applied to lexical reconstructions partially supports the Altaic linguistic macrofamily](#) от одного из ее соавторов. В статье методами компаративистики показана неслучайность сходств языковых семей, входящих в алтайскую макросемью.

Илья Егоров, н.с. ШАГИ РАНХиГС (Москва)

Алтайская проблема

Алтайская гипотеза предполагает родство, т.е. происхождение от общего языка-предка, пяти языковых семей Евразии: тюркской, монгольской, тунгусо-маньчжурской, японской (кроме японских диалектов включает рюкюские языки) и корейского языка. Эта гипотеза существует в двух версиях. “Узкая” версия предполагает только родство тюркских, монгольских и тунгусо-маньчжурских языков. В московской школе компаративистики их принято называть *ядерными алтайскими языками*. “Широкая” версия включает также корейский и японский. Все вместе это называется *алтайской макросемьей*. Иногда термин *алтайская (макро)семья* используют только для клады, включающей тюркские, монгольские и тунгусо-маньчжурские языки, а говоря о единстве с корейским и японским, употребляют термин *трансевразийская макросемья*.

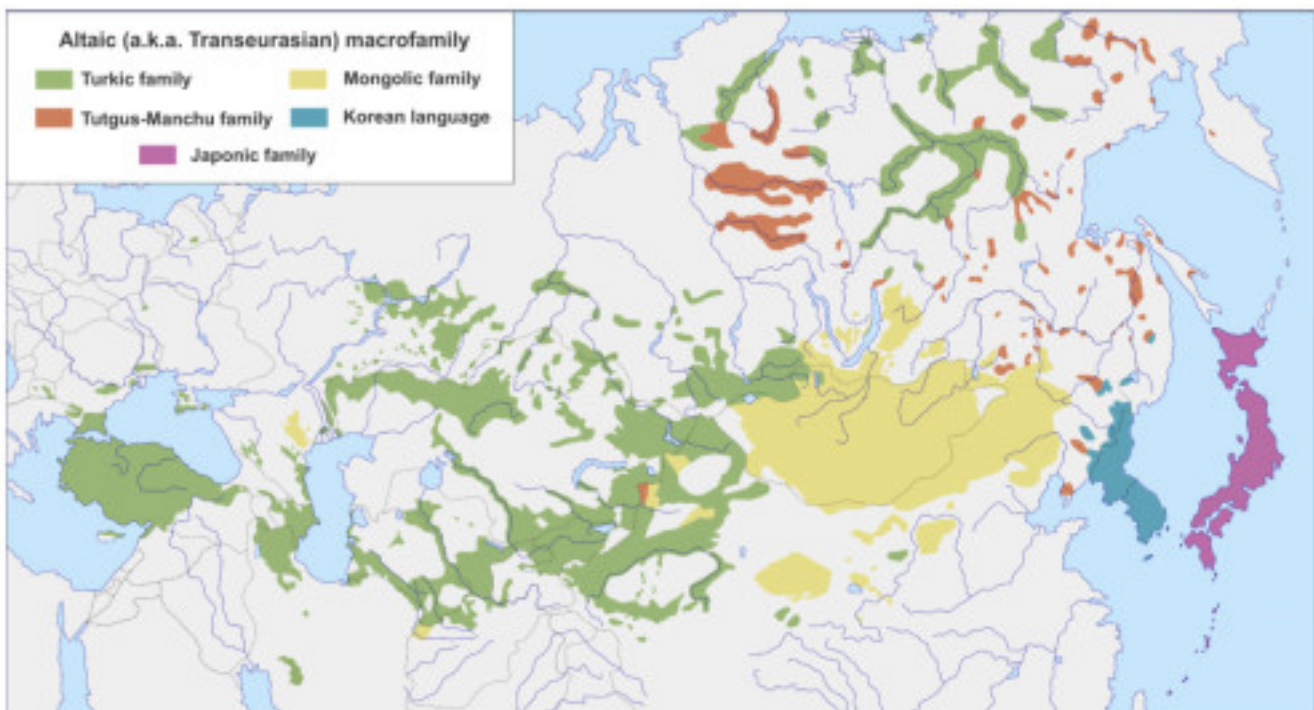


Рисунок 1. Современная территория распространения языковых семей, которые включаются в алтайскую (трансевразийскую) макросемью.

Алтайская гипотеза в ее узком варианте с добавлением корейского была сформулирована Густавом Рамstedтом еще в начале XX века (см. [Рамстедт 1957](#)). Однако первая последовательная фонетическая реконструкция праалтайского языка появилась только на рубеже тысячелетий с выходом “Этимологического словаря алтайских языков” ([Starostin, Dybo, Mudrak 2003](#)).

Гипотеза Рамстедта вызвала активную критику со стороны специалистов по отдельным семьям, с выходом алтайского словаря поднялась новая волна критики. История споров об алтайской проблеме изложена в обобщающей книге Вацлава Блажека (2019), которая также является хорошим введением в алтаистику.

Неслучайность сходств в фонетике, грамматике и лексике тюркских, монгольских и тунгусо-маньчжурских языков можно считать общепризнанной. Но трактовка этих сходств различна. Противники алтайской гипотезы утверждают, что они вызваны тесными и очень продолжительными контактами. Странники теории алтайского родства связывают часть сходств с общим происхождением этих языков, обычно признавая, что другие сходства вызваны языковыми контактами. Контактное объяснение возможно и для корейского с японским. Неслучайность сходств между двумя гипотетическими кладами ядерной алтайской и японско-корейской нуждается в дополнительном обосновании.

Доказательство родства vs предворительное сравнение

Стандартным доказательством языкового родства считается нахождение регулярных фонетических соответствий в базисной лексике. Базисная лексика — часть словаря, не зависящая от типа культуры и демонстрирующая особенную устойчивость.

На практике поиску регулярных фонетических соответствий обычно предшествует этап, когда лингвист “на глазок” оценивает сходство лексики. Формализовать этот этап работы можно путем *сравнения слов по консонантным классам*. Консонантные фонемы, т.е., проще говоря, согласные, подвержены изменению с ходом времени гораздо меньше, чем гласные. При этом изменения в глухости / звонкости, мягкости / твердости, придыхательности или в таких деталях, как приподнятость или опущенность кончика языка (сравните английский *t* и русский *т*), встречаются все-таки довольно часто. Если как бы закрыть глаза на эти менее существенные признаки, то согласные можно объединить в определенные классы, например П-образные, Т-образные, Н-образные, М-образные, Р-образные, шипяще-свистящие и т.д.

P-class (labials)	p b β β f v ...
T-class (dentals)	t d ɸ θ ð ɟ ʈ ...
S-class (sibilant fricatives)	s z š ž ...
ʃ-class (sibilant affricates)	ç ʒ č ʒ ...
Y-class (palatal glides)	y ...
W-class (labial glides)	w w ...
M-class (labial nasals)	m ŋ ...
N-class (non-labial nasals)	n ŋ ɲ ŋ ...
Q-class (lateral affricates)	ʎ ...
R-class	r r ...
L-class	l ʎ ʎ ...
K-class (velars and uvulars)	k g x q ɣ ʁ ...
Zero-class or H-class	h ʕ h ʕ ? h ʕ ? and any vowels

Таблица 1. Консонантные классы.

Каждое привлекаемое к сравнению слово редуцируется до консонантных классов. Поскольку одна из самых частотных моделей корня в языках Евразии — это согласный-гласный-согласный, удобно сравнивать структуры из двух согласных. Например, прамонгольское слово **nogoha* ‘зеленый’ превращается в *NK*, также как и пратунгусское **nog* ‘зеленый’, а прамонгольское **čila* ‘камень’ и пратунгусское **žolo* ‘камень’ становятся *ʒL*.

А что, собственно, сравниваем?

Как уже было сказано, для доказательства родства нужно исследовать в первую очередь базисную лексику. Один из самых

популярных списков базисной лексики — 100-словный список Морриса Сводеша. Московской школой компаративистики в 2010 году была разработана версия этого списка со строгим уточнением каждого значения (Kassian et al. 2010). Было, например, установлено, что в значении ‘нога’ нужно брать эквивалент ‘foot’, а не ‘leg’; в значении ‘плавать’ брать глагол описывающий действия человека, а не лодки или бревна.

Для исследований в области дальнего языкового родства ключевым является принцип ступенчатости реконструкции. Это значит, что к сравнению привлекаются не слова современных языков, а реконструированные формы праязыков, чье существование не вызывает сомнений. Так, говоря об алтайском родстве, мы должны работать с пратюркскими, прамонгольскими, пратунгусскими, пракорейскими и праяпонскими формами. На практике успешную реконструкцию удается сделать только для четырех семей. Данные древнекорейских памятников и корейских диалектов слишком скудные, чтобы можно было сделать надежную реконструкцию списка Сводеша. Поэтому вместо пракорейского используется список, собранные по памятникам позднего среднекорейского периода (XV-XVI вв.).

Фонетическая реконструкция — довольно строгая процедура, но все же она допускает некоторую свободу в интерпретации. Разные исследователи могут восстанавливать немного разные фонемы праязыка на основании одних и тех же регулярных соответствий или признавать разные системы соответствия. Иногда эти расхождения могут быть серьезными и касаться принадлежности согласного к одному или другому консонантному классу. Так в случае с японским Сергей Анатольевич Старостин реконструировал *d*- в тех местах, где другие лингвисты реконструируют *u*-. Это обстоятельство заставляет использовать две версии праяпонской реконструкции.

Если фонетическая реконструкция уже давно является строгим методом, то реконструкция значения долгое время делалась на глазок. Ноу-хау московской школы компаративистики — строгий метод *ономасиологической реконструкции*.

Ономасиологическая реконструкция предполагает, что исследователь подбирает оптимальный корень (или основу), которым могло выражаться данное значение в праязыке. Иными словами, это реконструкция от значения к форме. Строгим этот метод делают 5 критериев, по которым осуществляется такой отбор:

- Представленность корня в разных подгруппах языковой семьи должна быть такая, чтоб она предполагала минимальное количество семантических переходов в истории языков этой семьи.
- Даже если корень нашелся в языках разных подгрупп, они не должны быть географически смежными. Это нужно чтобы исключить заимствования и ареальные влияния.
- Если корень имеет внешние параллели за пределами семьи или группы, он более предпочтителен, чем тот, что представлен лишь в данной группе или семье.
- Предпочтение отдается непроезным корням, а основы с аффиксами и корнесложения считаются более слабыми кандидатами.
- Сценарий изменений значения корня в отдельных языках должен предполагать типологически более вероятные развития с точки зрения типологии семантических переходов.

Таким образом было реконструировано четыре списка (пратюркский, прамонгольский, пратунгусо-маньчжурский, праяпонский), при этом японский в двух версиях (с *d*-реконструкцией С.А. Старостина и с консервативной *u*-реконструкцией), и один список собран по среднекорейским памятникам.

Как доказать, что сходства неслучайны?

Для доказательства неслучайности сходств используется перестановочный тест. Принцип перестановочного теста прост. Одинаковые структуры из двух согласных (биконсонантные структуры) в каждой паре языков помечаются как совпадения. Дальше один из списков перемешивается в случайном порядке, и снова оценивается количество совпадений. Если сделать таких перестановок достаточно много, то можно понять среднюю вероятность случайного совпадения биконсонантных структур между списками. На основе этого показателя можно оценить, какова вероятность того, что совпадения, имеющиеся в исходных списках случайные. Чем меньше такая вероятность, тем более статистически значимы эти совпадения. Так выглядит классический перестановочный тест, он не взвешенный, это значит, что все потенциальные совпадения биконсонантных структур имеют одну цену (один вес).

Но в настоящей работе использован взвешенный перестановочный тест (еще одно ноу-хау московской школы), который устроен следующим образом. 110 сводешевских концептов ранжированы по своей типологической устойчивости, например, слово ‘глаз’ в целом стабильнее в языках мира, чем слово ‘грудь’, слово ‘новый’ стабильнее, чем слово ‘круглый’. Чем выше индекс стабильности у концепта, тем выше цена за биконсонантное совпадение в этом концепте. Дешевизна совпадений в

низко стабильных концептах оправдана тем, что чем менее стабилен концепт, тем выше вероятность лексической замены, которая, в свою очередь, может быть случайно созвучна слову из другого языка.

Что показало автоматическое сравнение?

Алгоритм обнаружил 66 пар биконсонантных совпадений в списках Сводеша пяти языковых семей. Сравним этот результат со взглядом авторов “Этимологического словаря алтайских языков”, в котором этимологически родственные слова определяются по регулярным фонетическим соответствиям, а не по консонантным классам. Получается, что среди выявленных алгоритмом соответствий 11 (17%) ложноположительные, т.е. эксперты их не признают родственными. А еще 74 пары, которые лингвисты считают родственными, алгоритм не выявил. В этих случаях фонетические изменения привели к смене консонантного класса. Относительно небольшое количество ложноположительных результатов можно считать достоинством метода.

Перестановочный тест показал высокую статистическую значимость совпадений между ядерными алтайскими семьями (тюркской, монгольской и тунгусо-маньчжурской). Сравнение с японским дало более слабый результат: статистическую значимость имеют только японско-тюркские и японско-тунгусские совпадения.

А вот совпадения лексики любого из четырех языков с корейским статистической значимости не показали. Это может быть связано с историей корейской фонетики. Большинство неначальных согласных в этом языке исчезли, а в начале некоторых слов добавилось *s*. Эти процессы очень сильно изменили структуру корня. Неудивительно, что в такой ситуации метод консонантных классов не смог найти достаточно совпадений в других алтайских семьях.

	Mongolic	Tungusic	Korean	Japonic (S. Starostin)
Turkic	7.8×10^{-5}	8.8×10^{-4}	0.577	1.8×10^{-4}
Mongolic	—	$<10^{-6}$	0.085	0.094
Tungusic	—	—	0.234	0.004
Korean	—	—	—	0.011
	Mongolic	Tungusic	Korean	Japonic ('conservative')
Turkic	7.8×10^{-5}	8.8×10^{-4}	0.577	5.4×10^{-5}
Mongolic	—	$<10^{-6}$	0.085	0.103
Tungusic	—	—	0.234	0.006
Korean	—	—	—	0.026

Таблица 2. Вероятность фонетических совпадений между пятью семьями, составляющими гипотетическую алтайскую макро-семью, по данным взвешенного перестановочного теста. Сверху результаты с использованием праапонской реконструкции С.А. Старостина, снизу с консервативной реконструкцией. Цветом обозначены статистически значимые значения: зеленым — $\alpha = 0.001$; желтым — $\alpha = 0.01$; красным — $\alpha = 0.05$.

Доказано ли родство?

Может ли перестановочный тест доказать родство языков? Строго говоря, нет. Для доказательства родства недостаточно просто показать, что какие-то слова похожи. Должна быть продемонстрирована регулярность фонетических соответствий. Можно даже сказать, что систематические различия интересуют компаративистов намного больше, чем внешнее сходство слов. Статистически значимые результаты перестановочного теста, скорее, нужно трактовать как эвристическое указание на связь языков. Но эта связь может быть как генетической, так и обусловленной контактами.

Вместе с тем, отсутствие статистически значимых результатов перестановочного теста для биконсонантных структур не опровергает родство языков. Хотя мы и ожидаем, что большинство фонетических изменений происходит в пределах одного консонантного класса, ничто не запрещает фонетические процессы, приводящие к смене класса. Например, переход $\check{c} > s$ или $z > r$ довольно часто встречается в языках мира, не говоря уже просто об отпадении конечных согласных.

Высокая статистическая значимость совпадений между ядерными алтайскими семьями (тюркской, монгольской и тунгусо-маньчжурской) позволяет трактовать их и как результат контактов, и как следы генетического родства. Мы знаем, что контакты между языками этих семей имели место в древности и продолжают по сей день. Другое дело статистически значимые совпадения в парах японский-тунгусский и японский-тюркский. Ввиду географической близости можно было бы спекулировать о древних японско-тунгусских контактах (хотя, конечно, намного больше это похоже на древнее родство). Но вот контактный сценарий с заимствованной лексикой между пратюркским и праапонским требовал бы некоторых сильных допущений по причине громадной географической дистанции между ареалами этих языковых семей. Сценарий языкового родства в тюркско-японском случае представляется куда более экономным.

Оригинальная статья

Kassian, A., Starostin, G., Egorov, I., Logunova, E., & Dybo, A. (2021). [Permutation test applied to lexical reconstructions partially supports the Altaic linguistic macrofamily](#). *Evolutionary Human Sciences*, 3, E32. doi:10.1017/ehs.2021.28

Литература

Рамstedт, Г.И. (1957). [Введение в алтайское языкознание](#). Москва.

Blažek, V. (2019). [Altaic languages: History of research, survey, classification and a sketch of comparative grammar](#). Masaryk University Press.

Kassian, A. S., Starostin, G., Dybo, A., & Chernov, V. (2010). [The Swadesh wordlist](#). An attempt at semantic specification. *Journal of Language Relationship*, 4, 46–89.

Starostin, S. A., Dybo, A. V., & Mudrak, O. A. (2003). [Etymological dictionary of the Altaic languages](#) (Vols 1–3). Brill.