

НПММВЯ-178, 23 марта 2013 г.

П.М. Аркадьев

(Ин-т славяноведения РАН/РГГУ/МГГУ)

peterarkadiev@yandex.ru

**Квантитативная
ареальная типология
префиксальной перфективации**

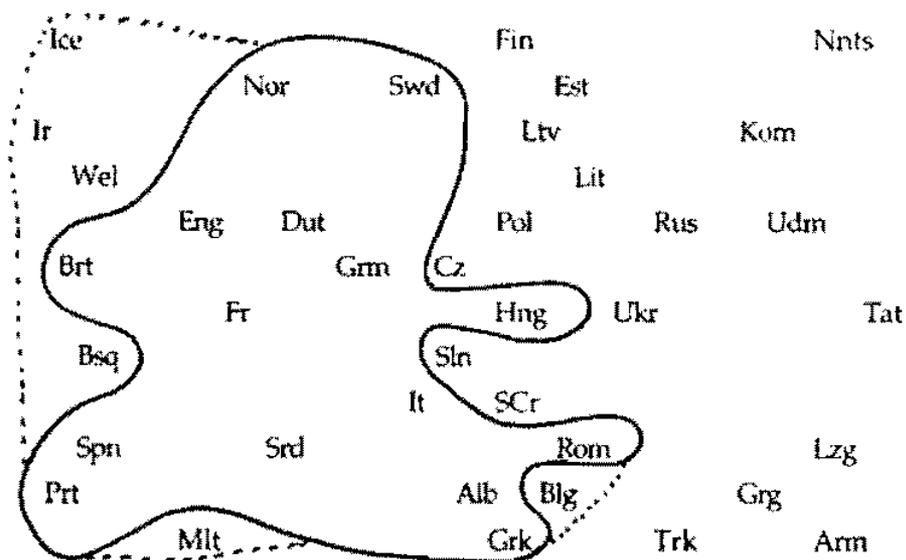
Что такое ареальная типология?

Areal linguistics is traditionally concerned with similarities between geographically contiguous languages, in particular when they cannot be ascribed to a common protolanguage, and the identification of so-called Sprachbünde ... or ... linguistic areas. (Dahl 2001: 1456)

Areal typology, then, is the study of patterns in the areal distribution of typologically relevant features of languages. It ... looks both at the patterns themselves and the processes that give rise to them. (ibid.)

Методы ареальной типологии

1. **Изоглосса** — обозначенное на карте множество идиомов (языков, диалектов), демонстрирующих одно и то же значение того или иного признака.



— definite and indefinite article present

- - - - only definite article present

Map 107.1: Definite and indefinite article

(Haspelmath 2001: 1494)

Методы ареальной типологии

Основной недостаток изоглосс:

типологические признаки не обязательно кластеризуются, т.е. разные изоглоссы выделяют разные множества языков, которые оказывается трудно сравнивать.

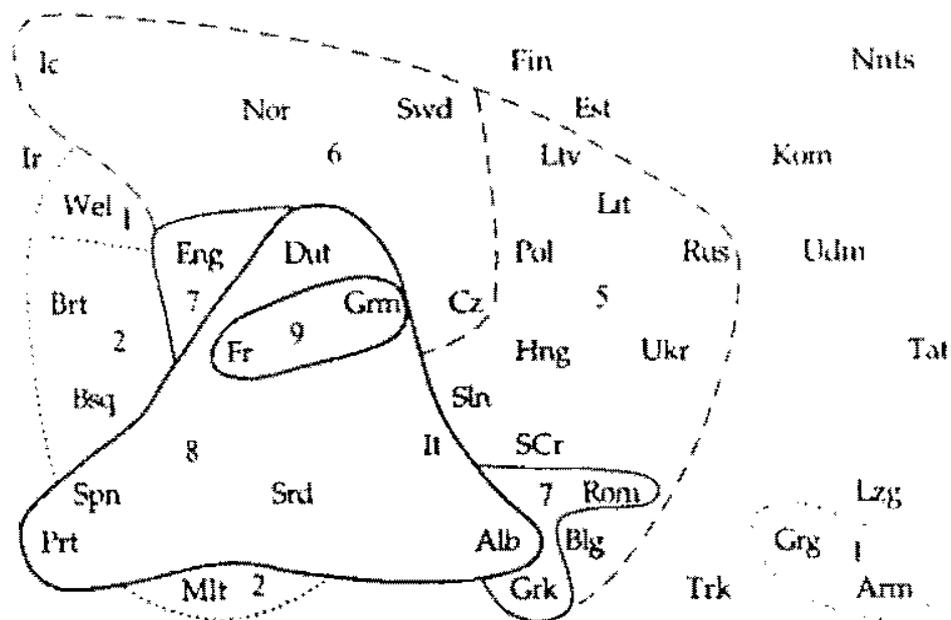
«Больной вопрос» теории «языковых союзов» (ср. Dahl 2001, Campbell 2006):

сколько общих признаков необходимо и/или достаточно для признания множества языков «языковым союзом»?

Одни изоглоссы более важны, чем другие?

Методы ареальной типологии

Изоплеты (isopleth) — lines showing the geographical distribution of languages which share the same number (not necessarily the same subset) of some set of features. (Dahl 2001: 1458)



Map 107.13: A cluster map combining nine features

(Haspelmath 2001: 1505)

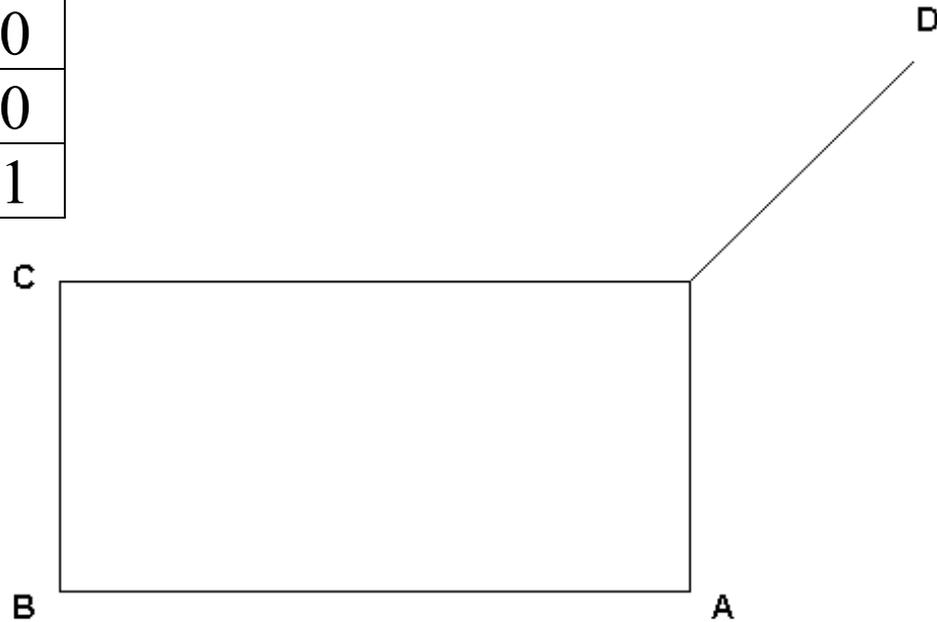
Методы ареальной типологии

Филогенетические графы (Neighbor Net в программе SplitsTree) — более сложное и математически нетривиальное средство представления сходств и различий между какими-либо объектами, основанное на множестве признаков.

Изначально разработано в эволюционной генетике (Bryant, Moulton 2004, Huson, Bryant 2006), в лингвистике применялось в сравнительно-исторических и типологических исследованиях (см. напр. McMahon, McMahon 2005; Bryant et al. 2005; Nichols, Warnow 2008; Соловьев 2010).

Методы ареальной типологии

| | α | β | γ | δ |
|---|----------|---------|----------|----------|
| A | 0 | 1 | 0 | 1 |
| B | 0 | 1 | 1 | 0 |
| C | 1 | 1 | 1 | 0 |
| D | 1 | 0 | 0 | 1 |



Методы ареальной типологии

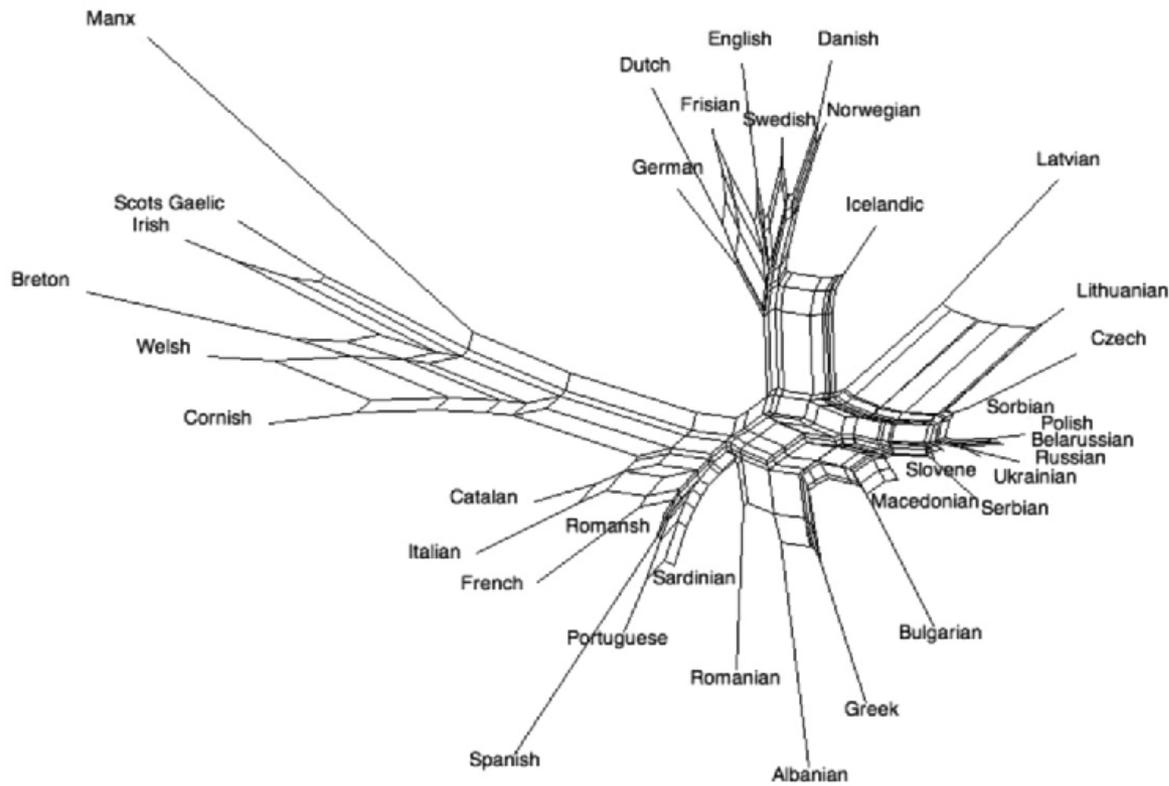


Figure 2. NeighborNet clustering of 36 European languages based on WALS morphosyntactic features.

(Donohue 2012: 102)

Префиксальная перфективация

Перфективация — операция, превращающая глагол в терминативный, т.е. обозначающий завершённую, ограниченную во времени ситуацию.

«Прототипический» случай перфективации связан с выражением противопоставления предельного действия в развитии (*писал письмо*) и достижения этим действием предела (*написал письмо*).

Перфективация также может выражать различные модификации глагольной семантики, как-то начинательность (*петь ~ запеть*), семельфактивность (*толкать ~ толкнуть*), временную ограниченность (*спать ~ поспать*) и др.

Префиксальная перфективация

Во многих языках Восточной Европы перфективация выражается с помощью словообразовательного средства — присоединения к глагольной основе **преверба**, т.е. морфемы, по крайней мере в существенной части своих употреблений выступающей как префикс (ср. Breu 1992).

Префиксальная перфективация

русский (славянские): *резать* ~ *разрезать*

литовский (балтийские): *skaityti* ‘читать’ ~ *perskaityti* ‘прочитать’

идиш (германские): *washn* ‘мыть’ ~ *oyswashn* ‘вымыть’

венгерский (уральские): *olvasni* ‘читать’ ~ *felolvasni* ‘прочитать’

осетинский (индоиранские): *фыста* ‘писал’ ~ *ныффыста* ‘написал’

грузинский (картвельские): *çers* ‘пишет’ ~ *daçers* ‘напишет’

Задача

- 1) Выявить характер и пределы варьирования глагольных систем с префиксальной перфективацией.
- 2) Установить, в какой степени сходства и различия между такими глагольными системами обусловлены генетическим родством языков, общетипологическими тенденциями и языковыми контактами.

Пути решения

Многофакторный типологический анализ (ср. Tomelleri 2009, 2010):

- 1) Исследование глагольных систем с префиксальной перфективацией на основе единого множества параметров.
- 2) Анализ кластеризации значений этих параметров (возможных корреляций между параметрами).
- 3) Анализ кластеризации глагольных систем на основе значений этих параметров.

Исследуемые языки

все (основные) славянские

балтийские (литовский, латышский)

германские (идиш + немецкий)

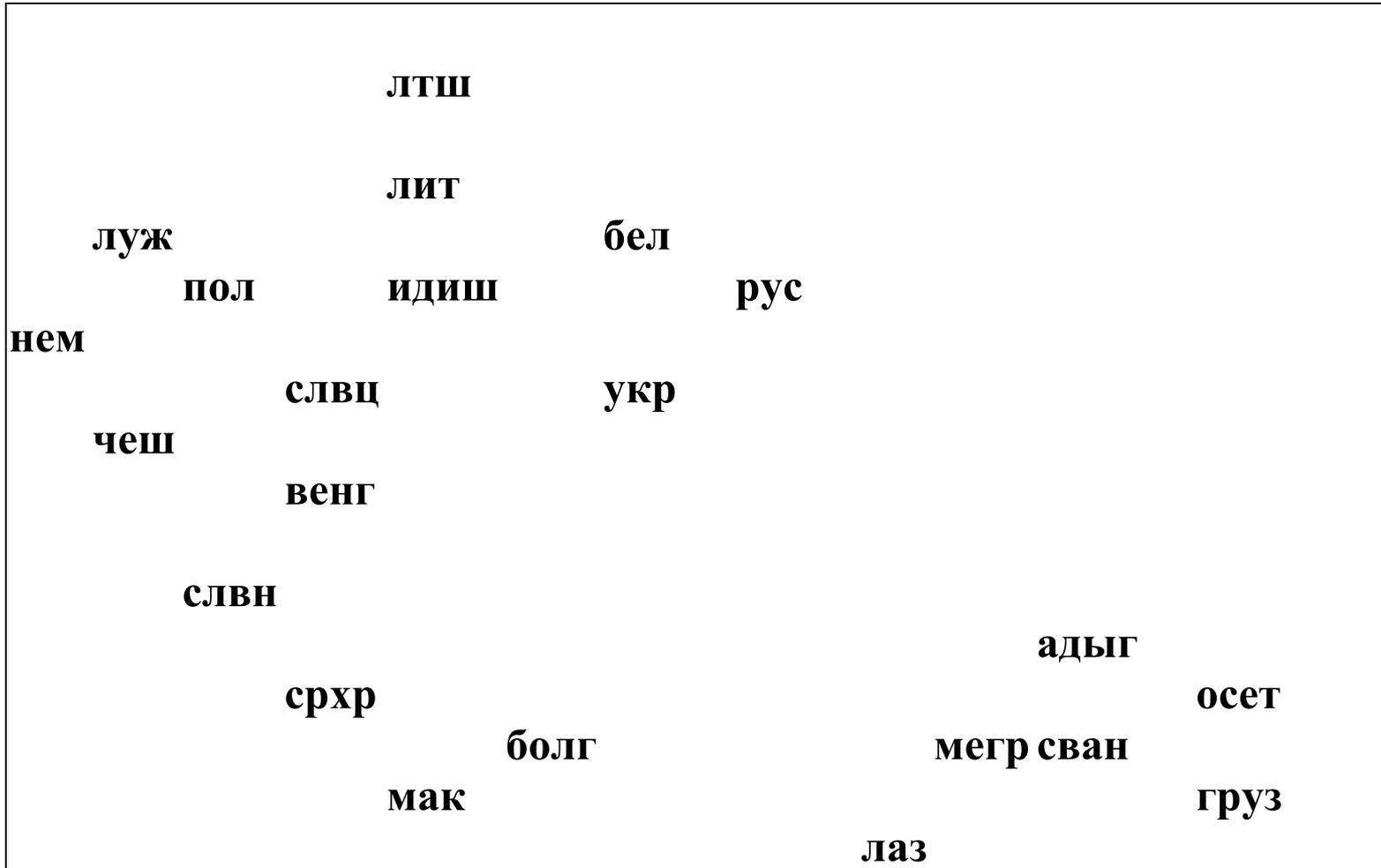
венгерский

осетинский

кارتвельские (грузинский, сванский, мегрельский, лазский)

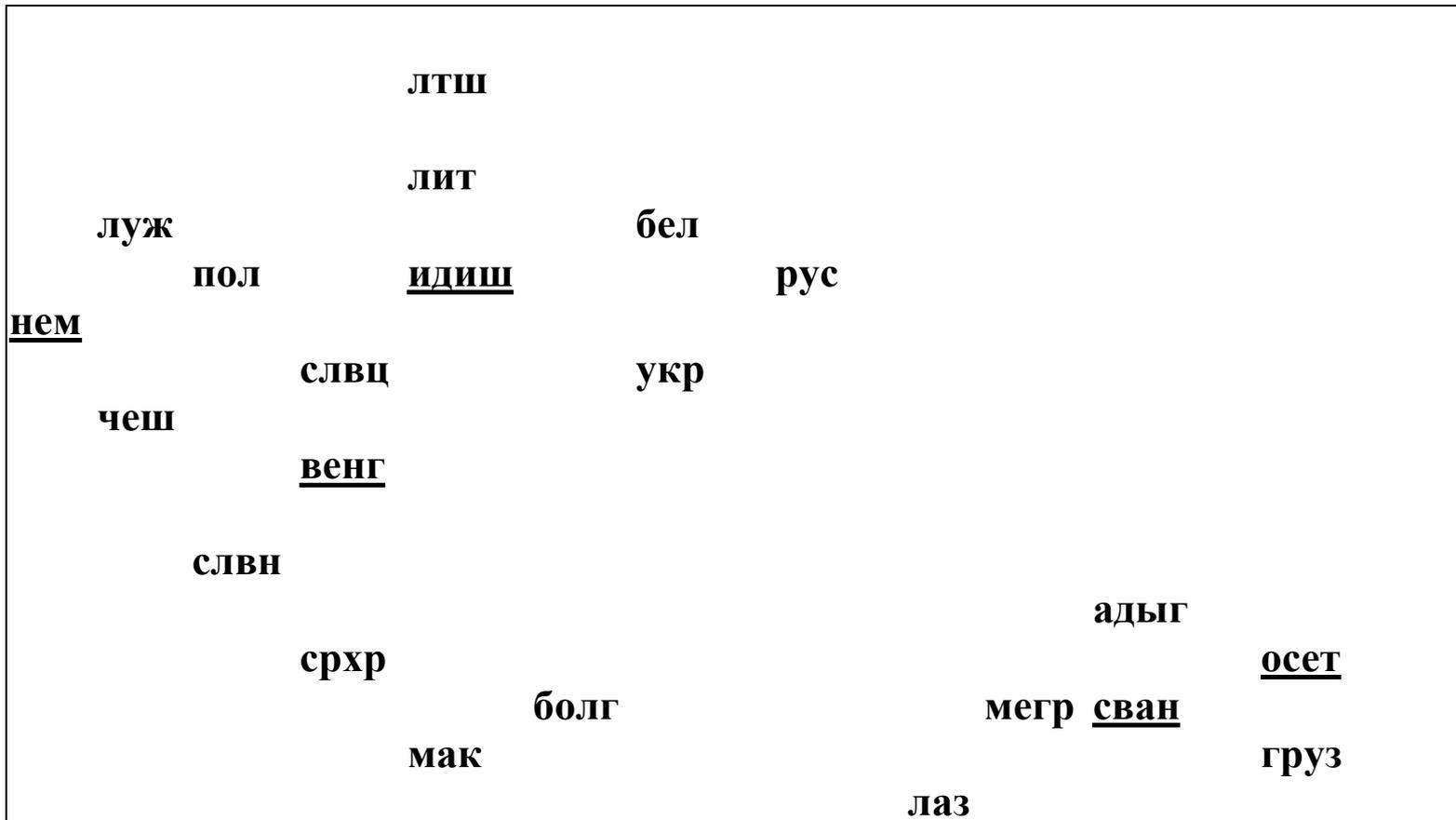
+ адыгейский

Исследуемые языки



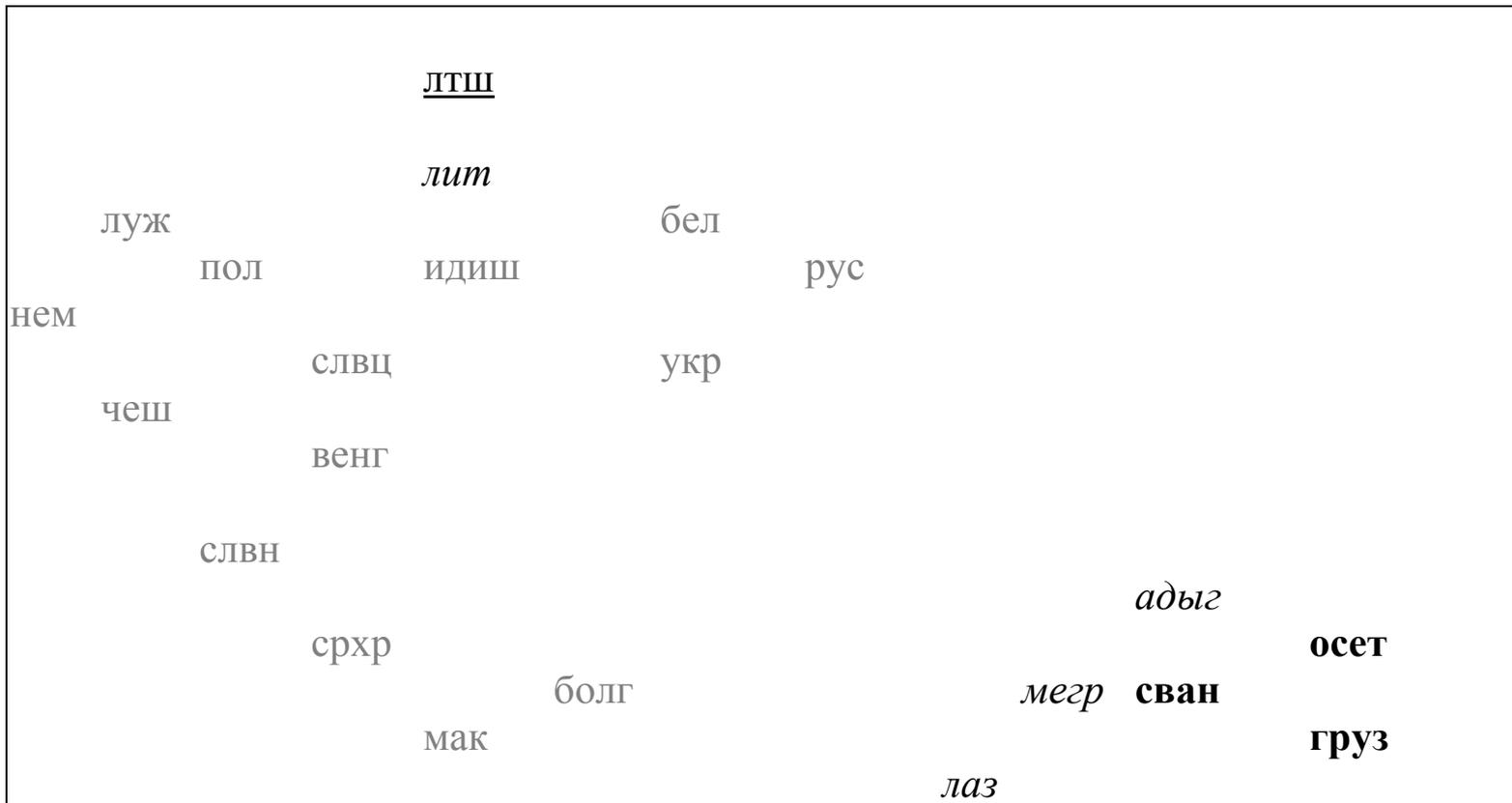
Типологические параметры

1.1. Отделимость превербов: 0 отсутствует; 1 допускается.



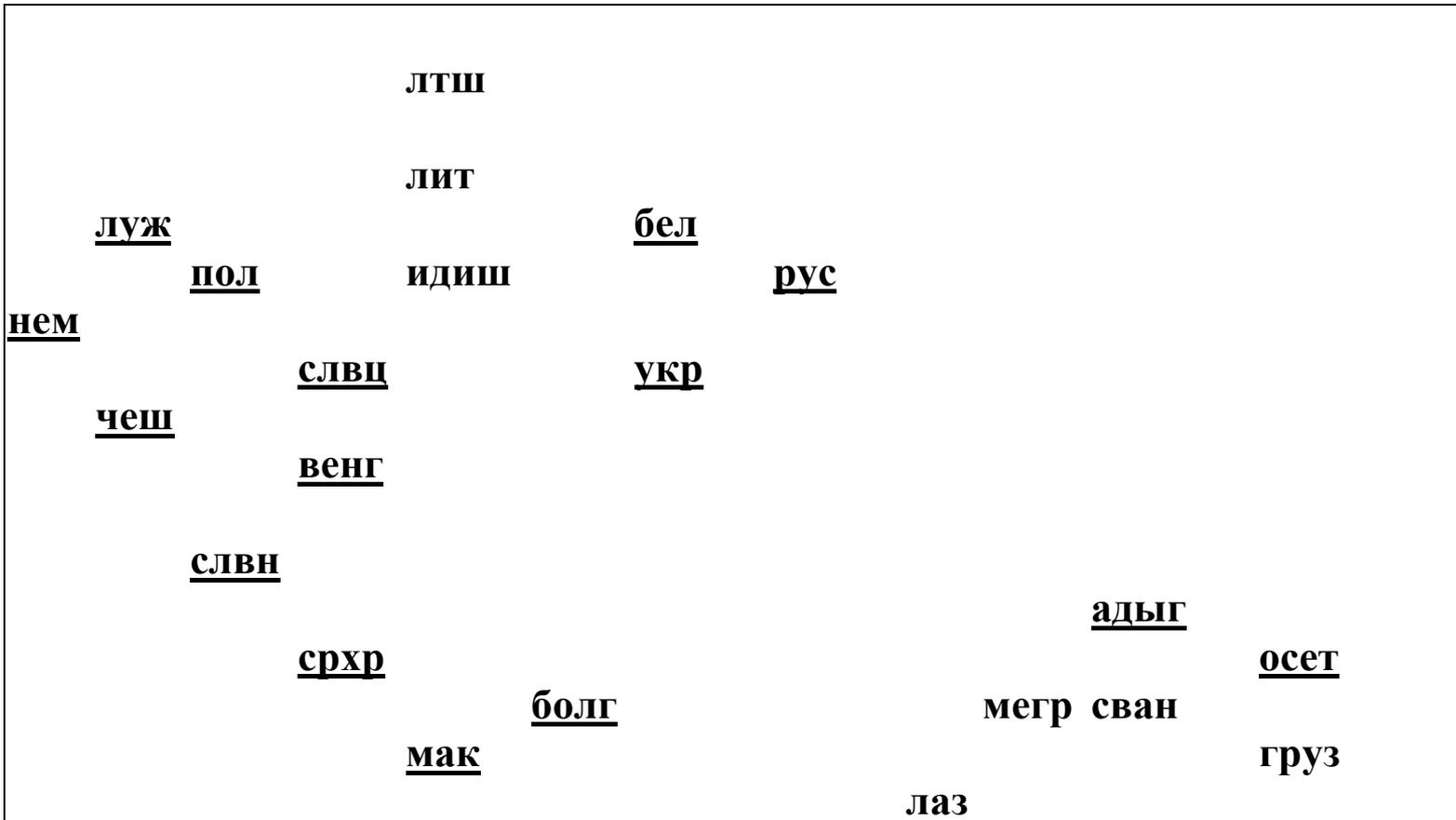
Типологические параметры

1.3. Позиция превербов в глагольной словоформе (при условии положительного значения признака 1.2): 1 **начальная**; 2 *срединная*; 3 предосновная.



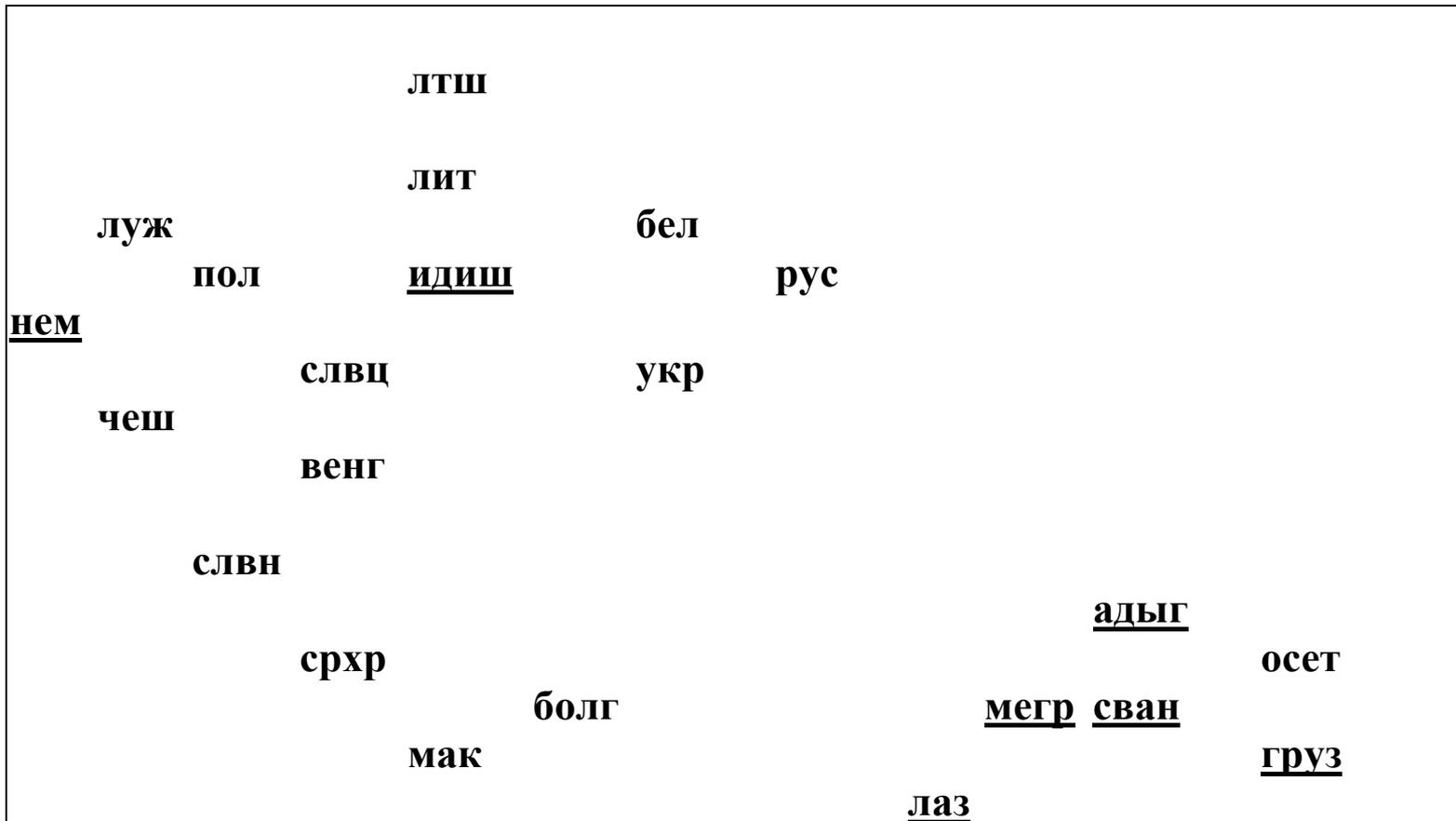
Типологические параметры

1.4. Итерация провербов: 0 отсутствует; 1 допускается.



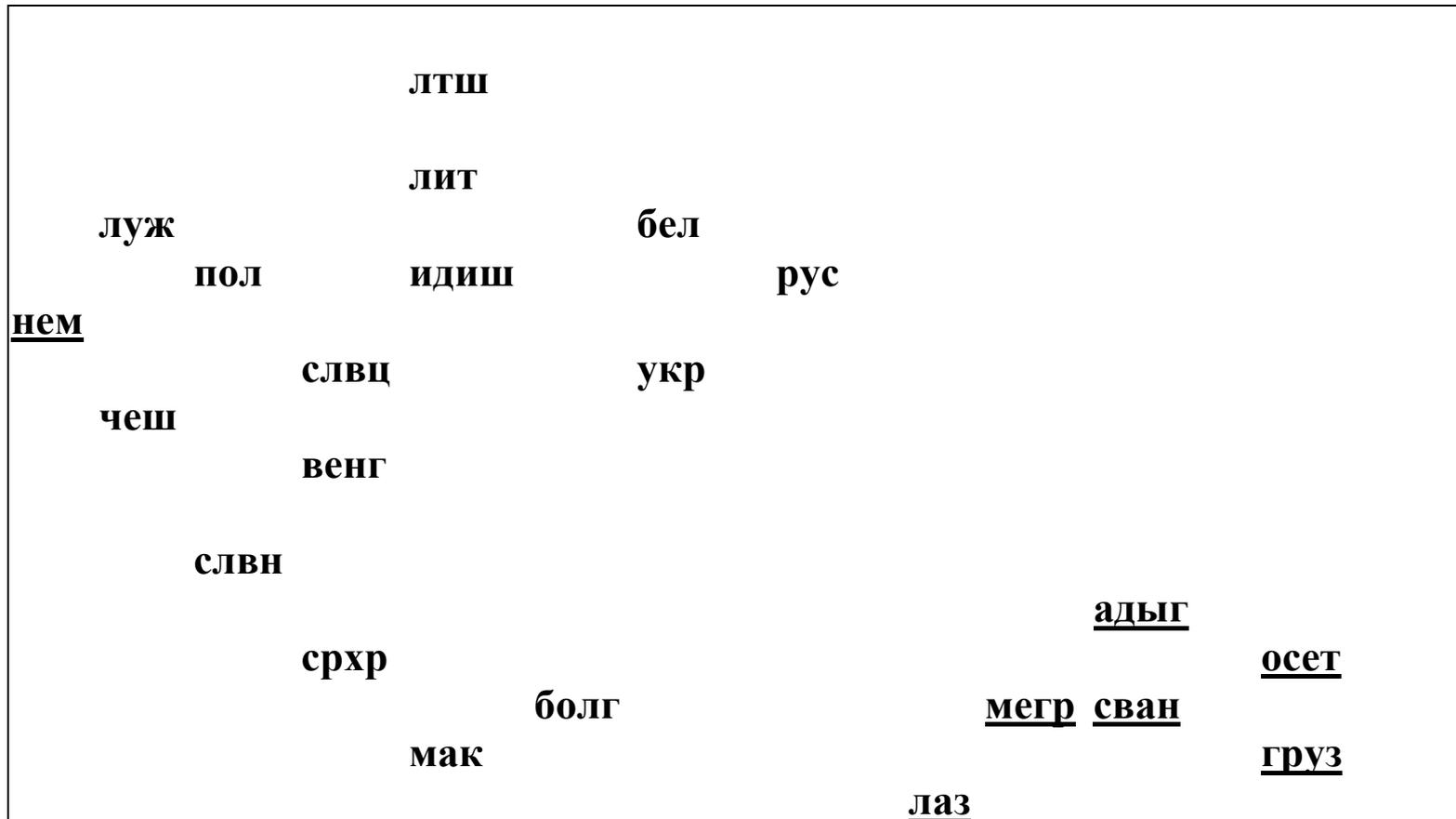
Типологические параметры

1.5. Морфологические подклассы превербов: 1 да; 0 нет.



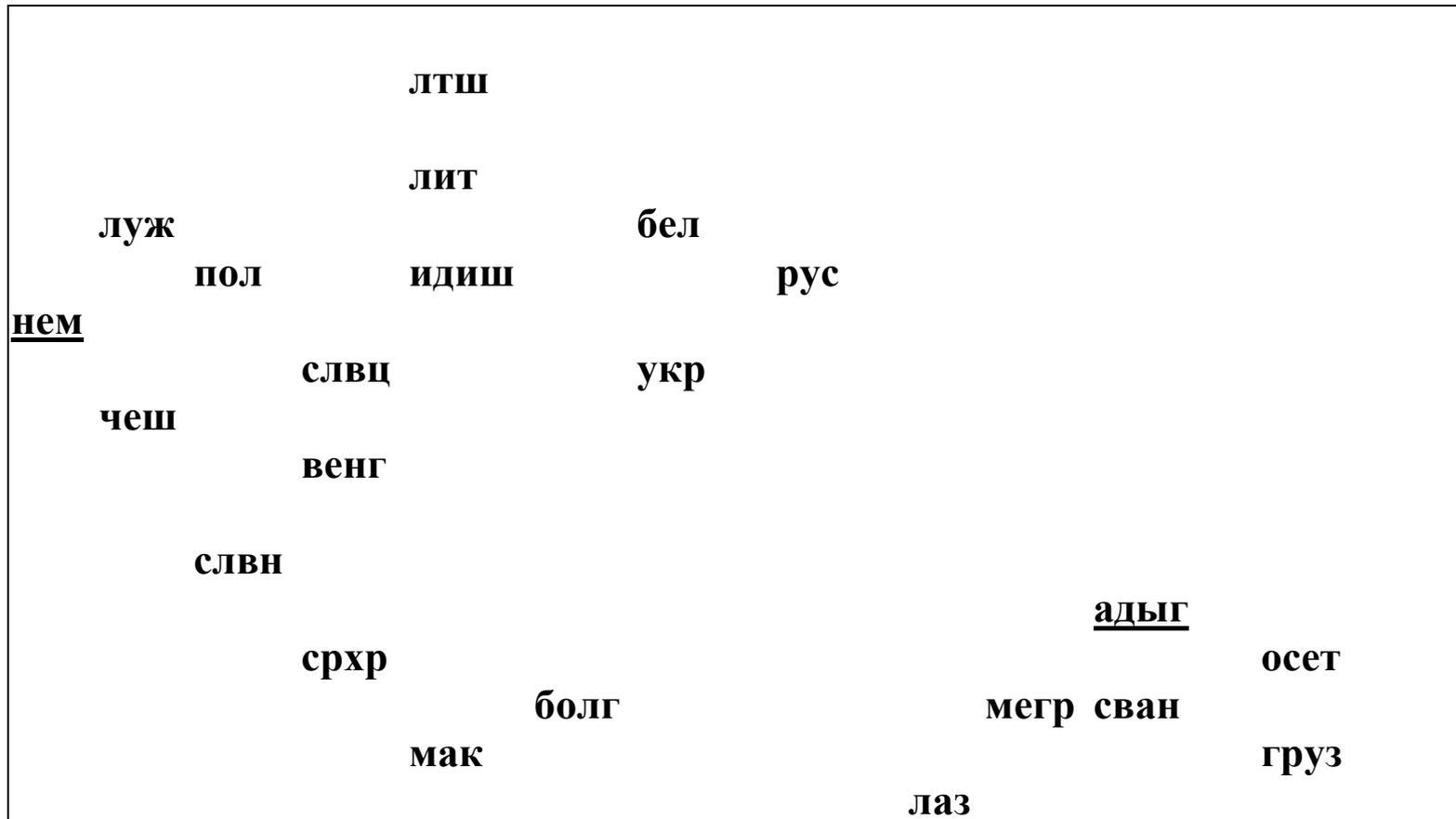
Типологические параметры

2.1. Выражение дейксиса: 0 систематическое отсутствует или маргинально; 1 систематическое.



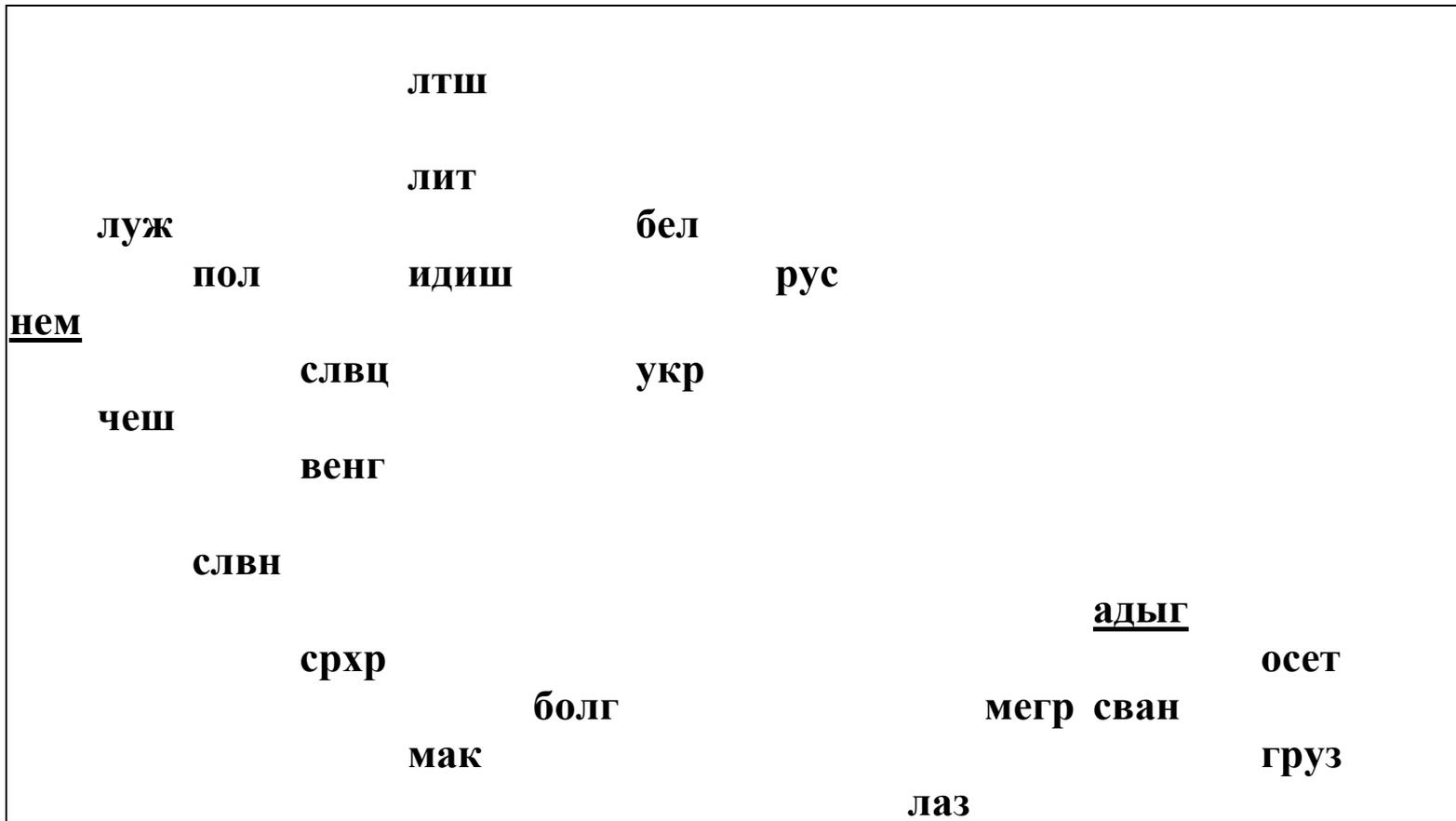
Типологические параметры

2.2. Перфективирующая функция превербов: 0 отсутствует или маргинальна; 1 продуктивна.



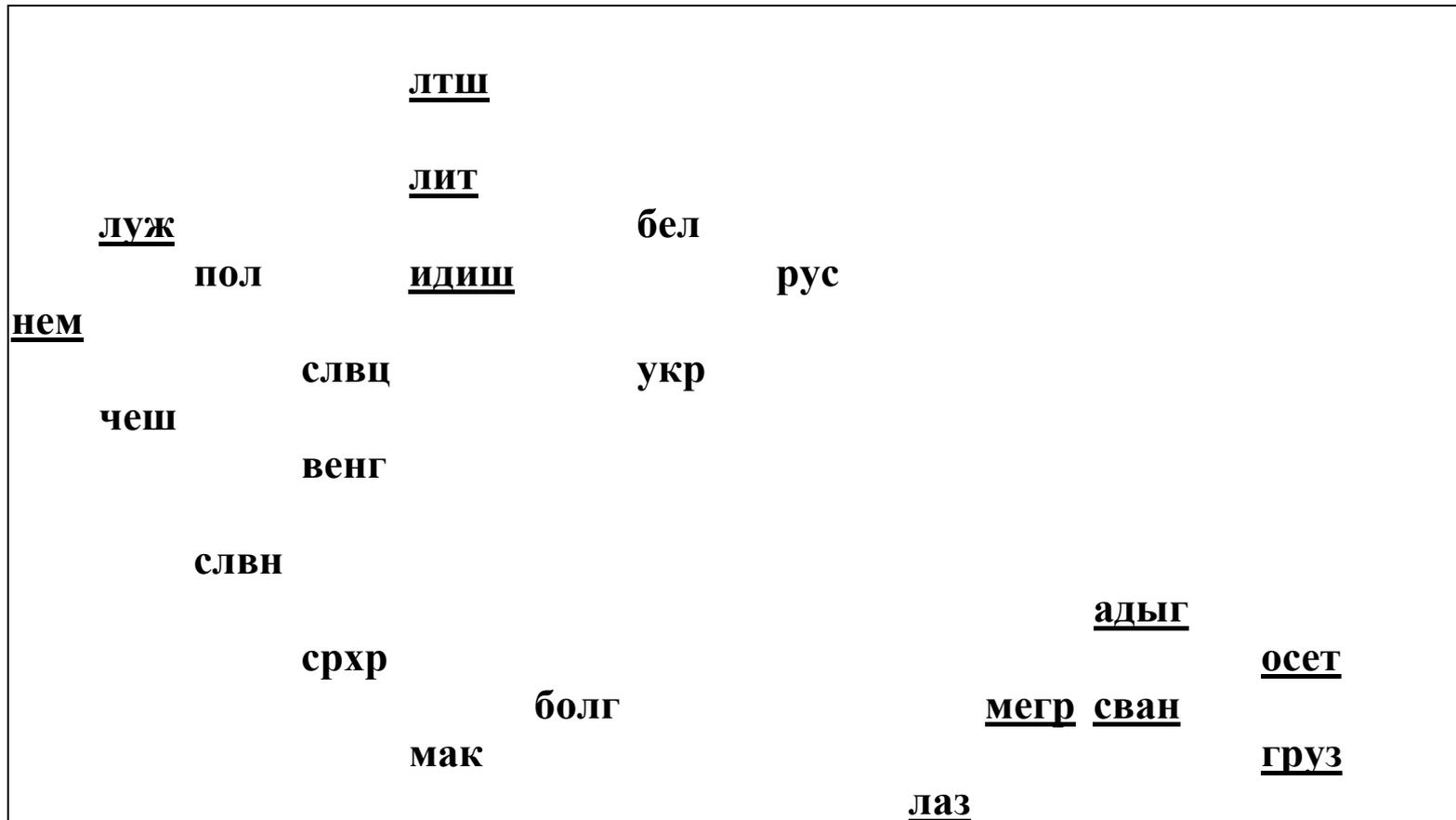
Типологические параметры

2.3. «Чистая перфективация»: 0 отсутствует или маргинальна; 1 продуктивна.



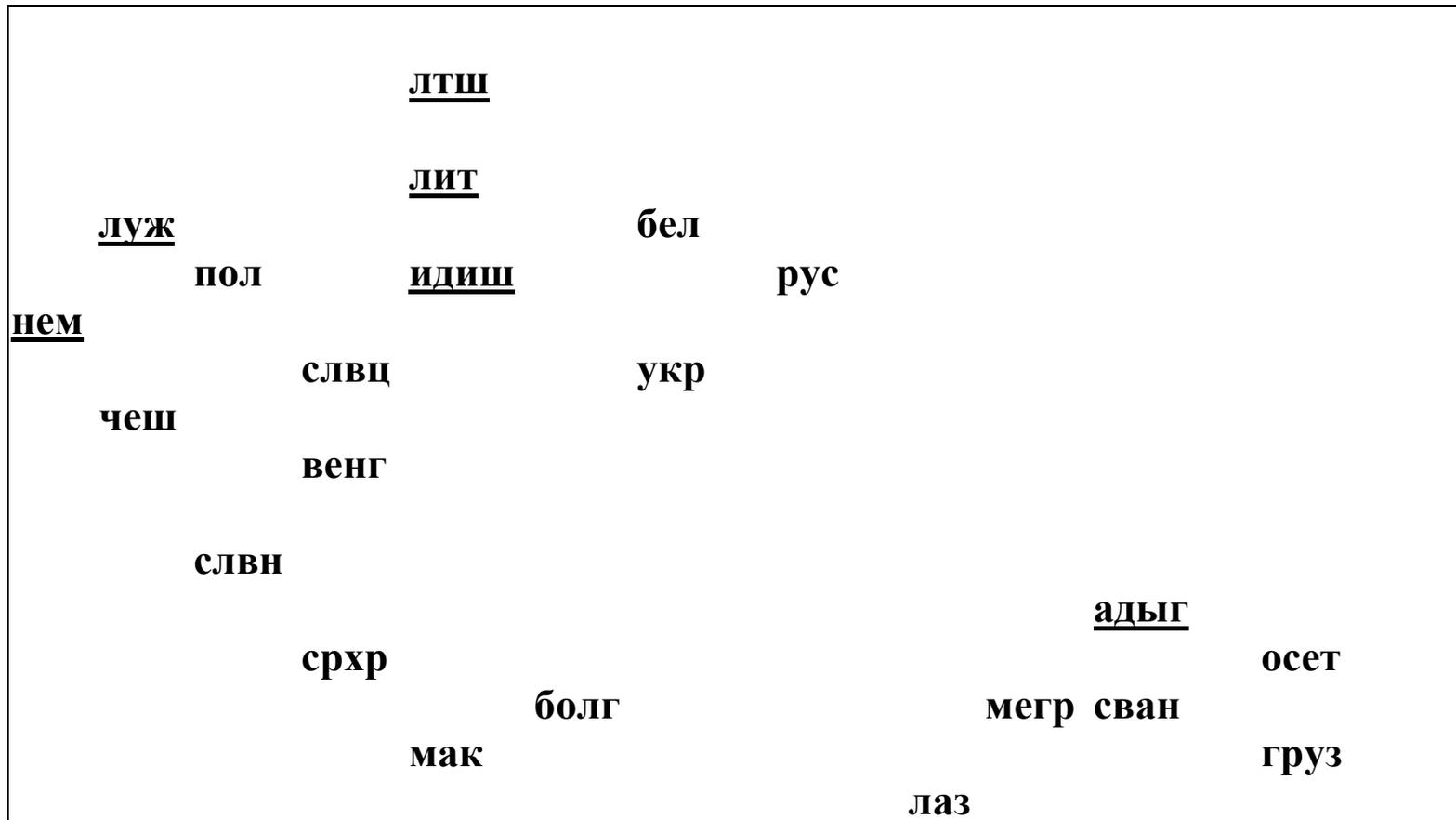
Типологические параметры

2.5. Дуративное употребление превербных глаголов перемещения: 0 запрещено; 1 допускается.



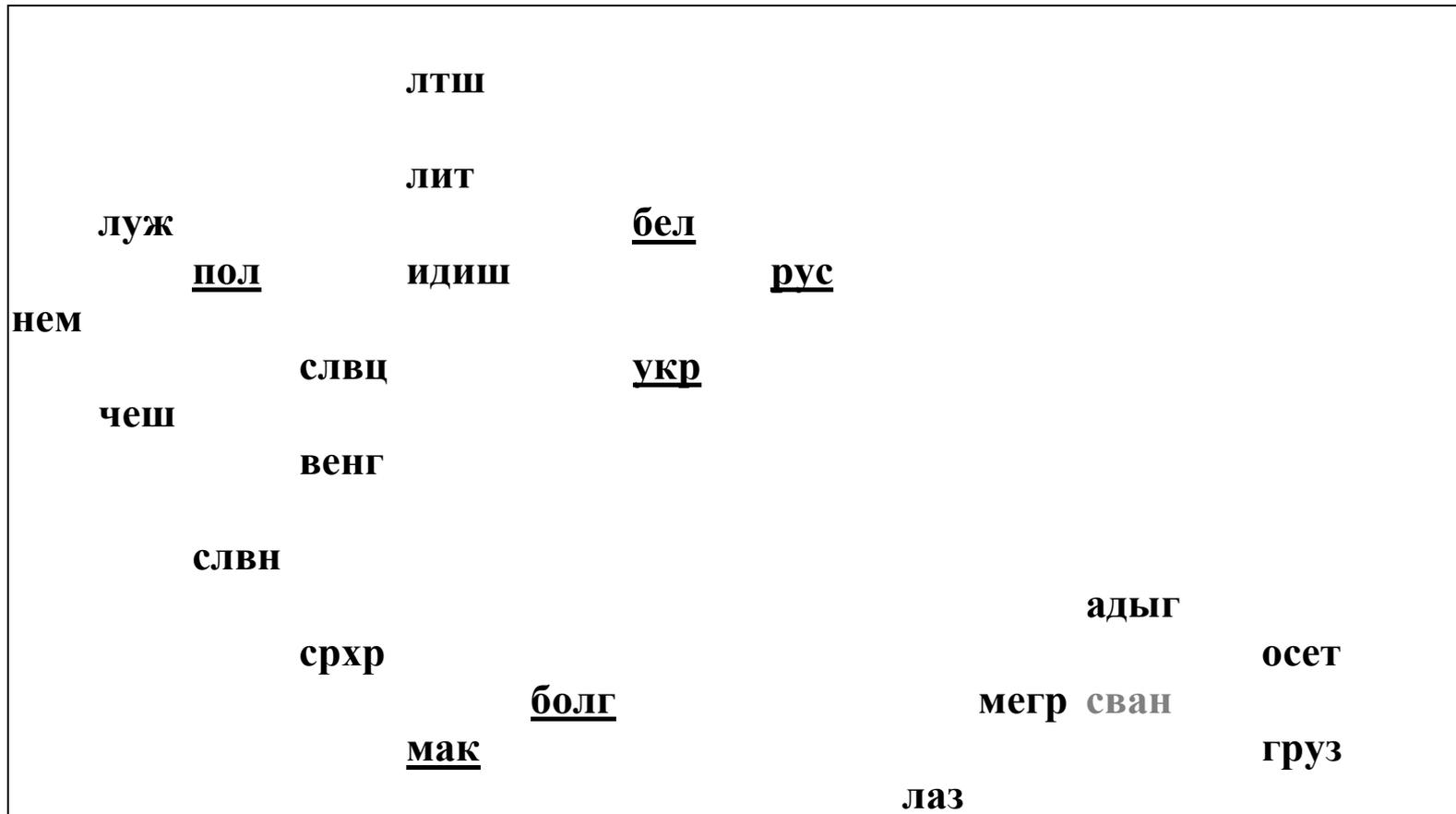
Типологические параметры

2.6. Дуративное употребление превербных глаголов других семантических классов: 0 запрещено; 1 допускается.



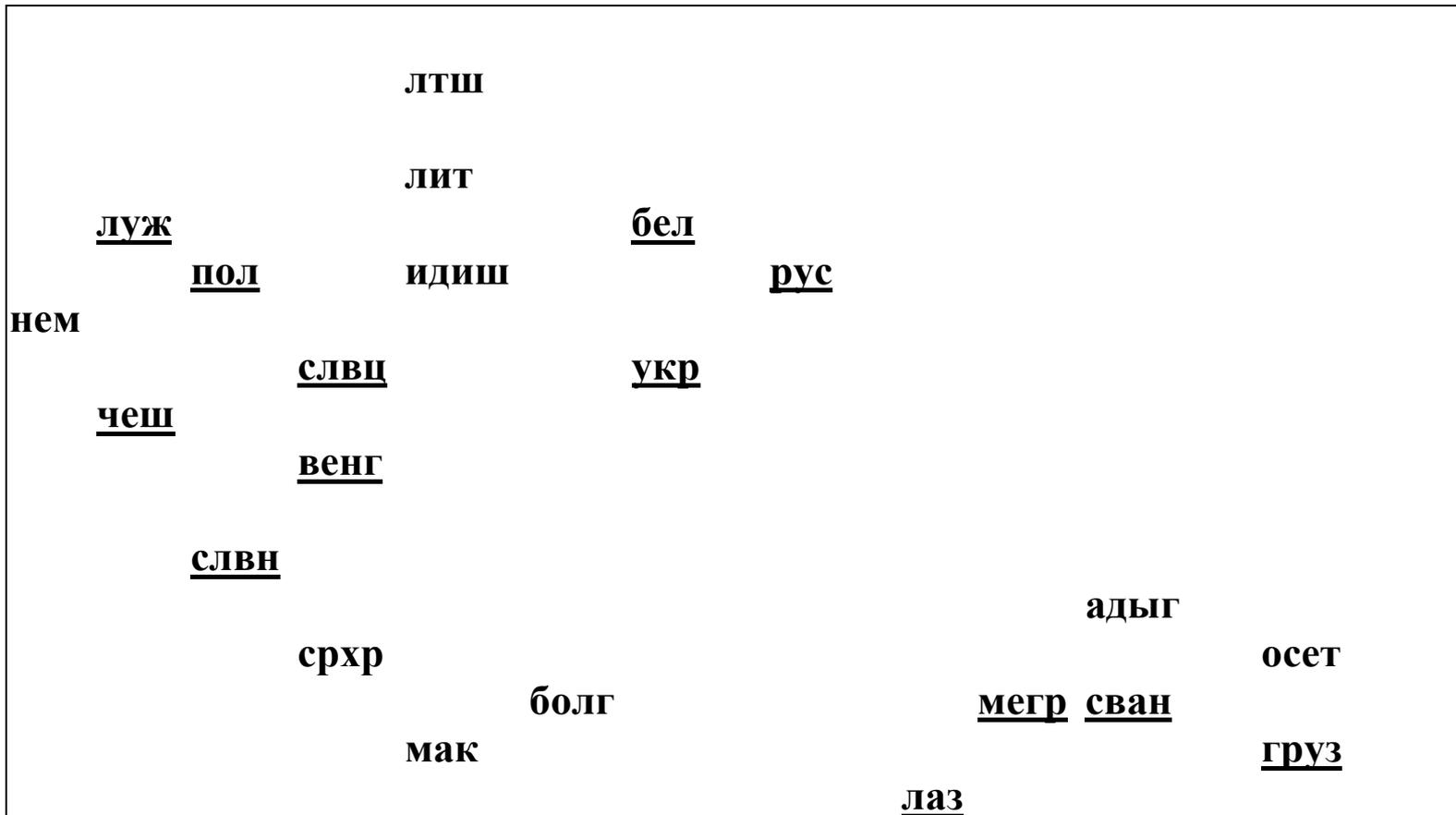
Типологические параметры

3.1. Использование презенса преевербных глаголов в значении *praesens historicum*: 1 да; 0 нет.



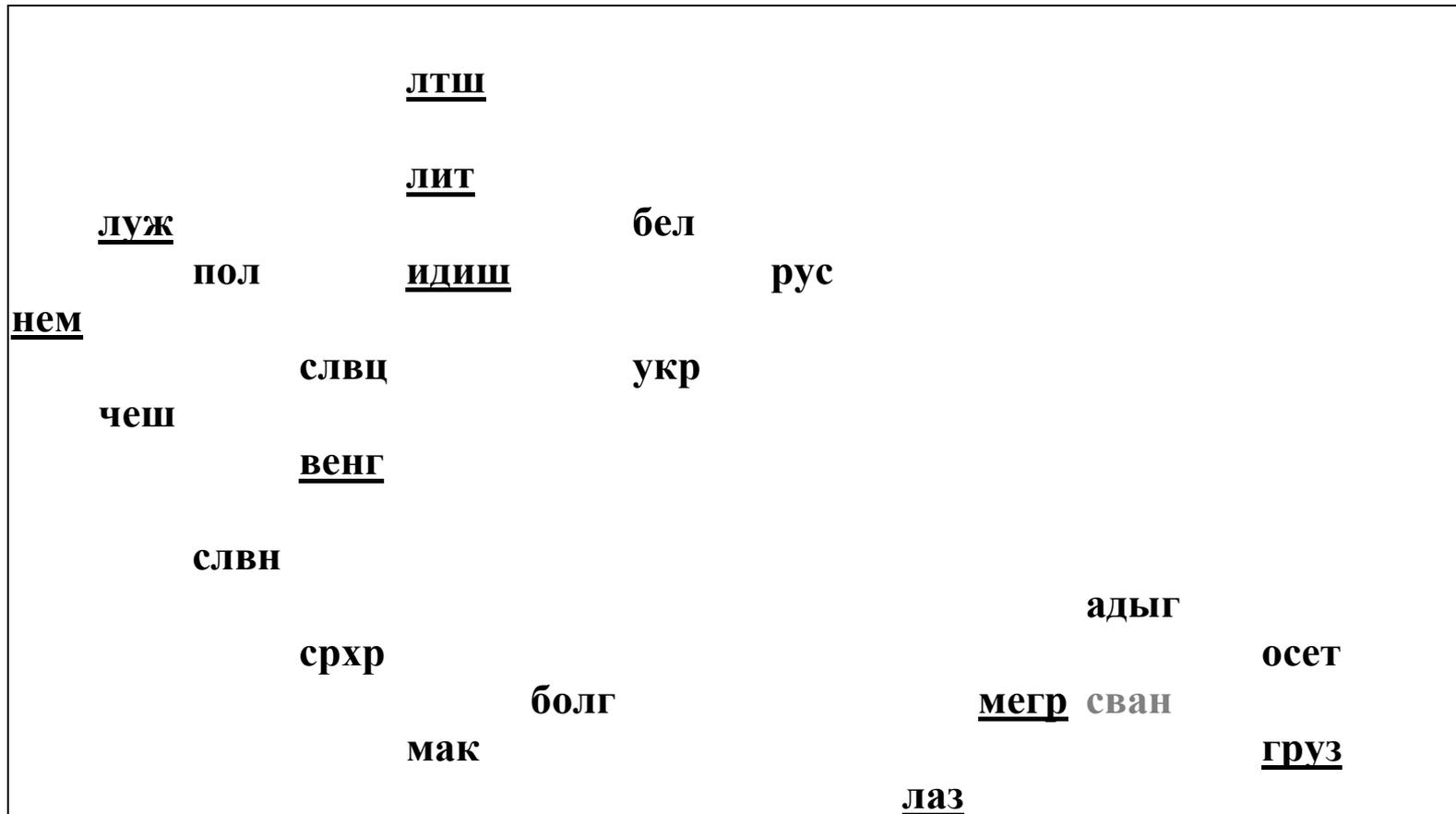
Типологические параметры

3.2. Продуктивное футуральное употребление презенса пре-
вербных глаголов: 1 да; 0 нет.



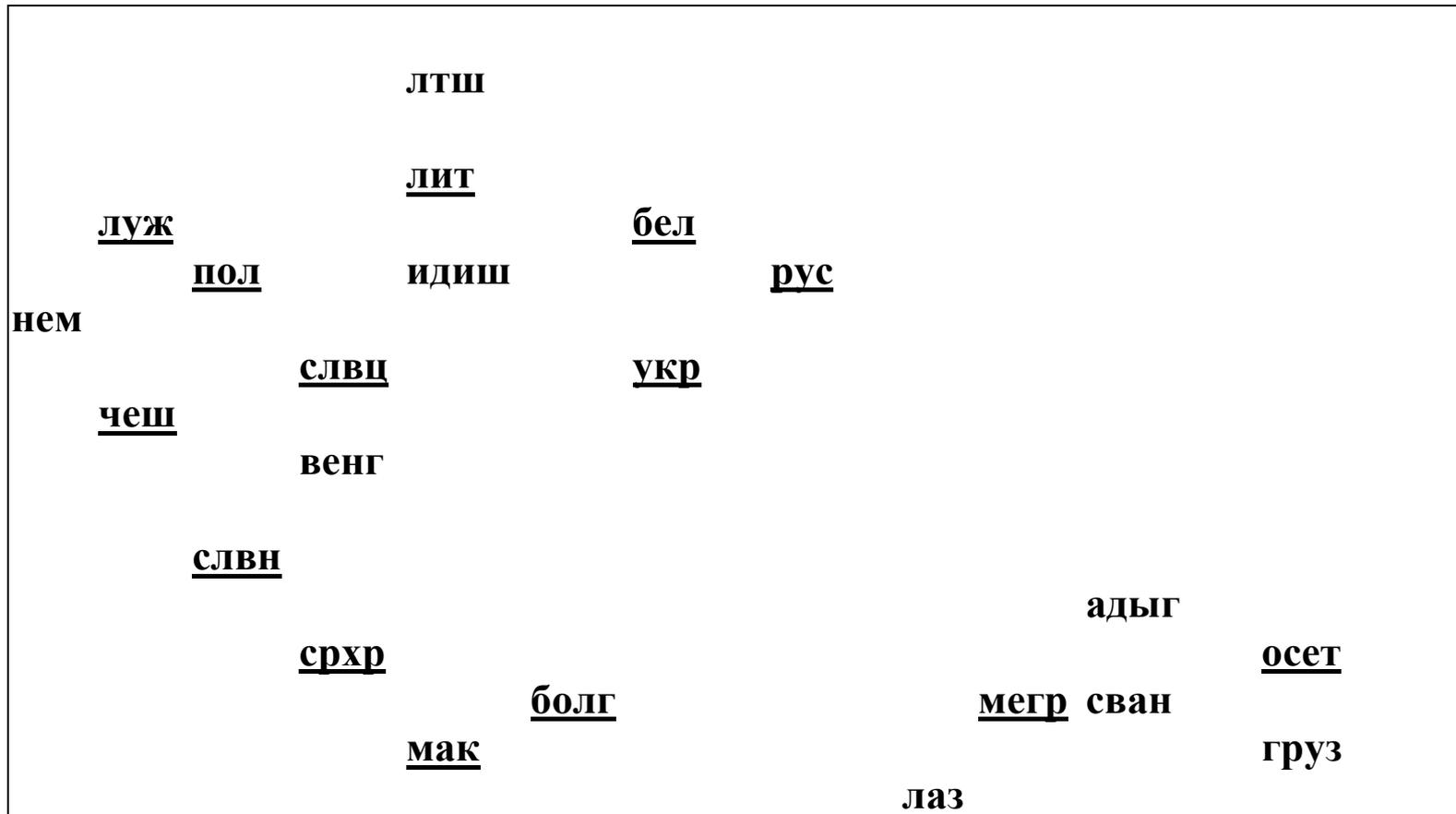
Типологические параметры

3.3. Использование превербных (или вообще терминативных) глаголов с фазовыми предикатами: 1 да; 0 нет.



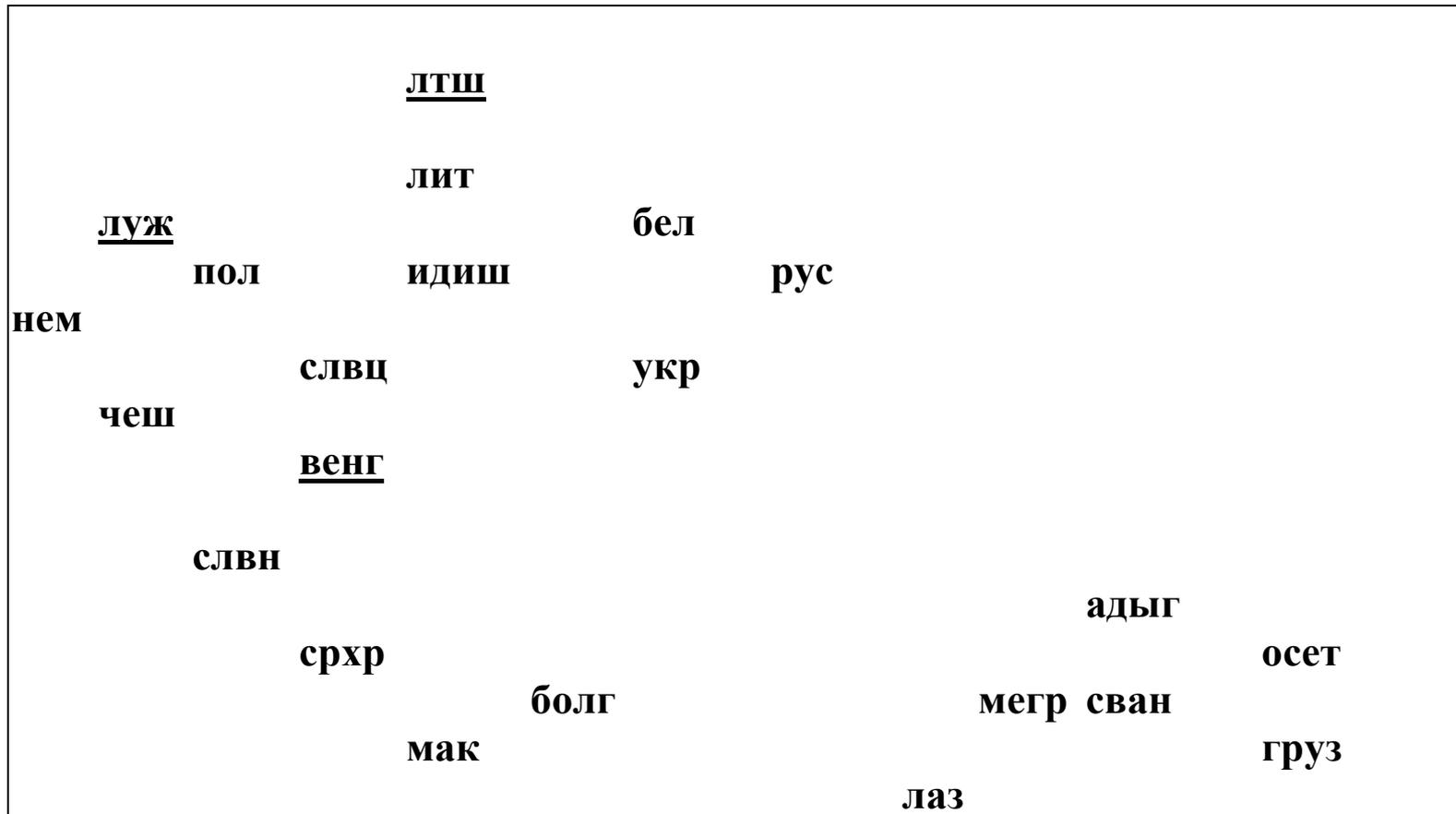
Типологические параметры

3.4. Морфологические средства вторичной имперфективации: 0 нет; 1 да.



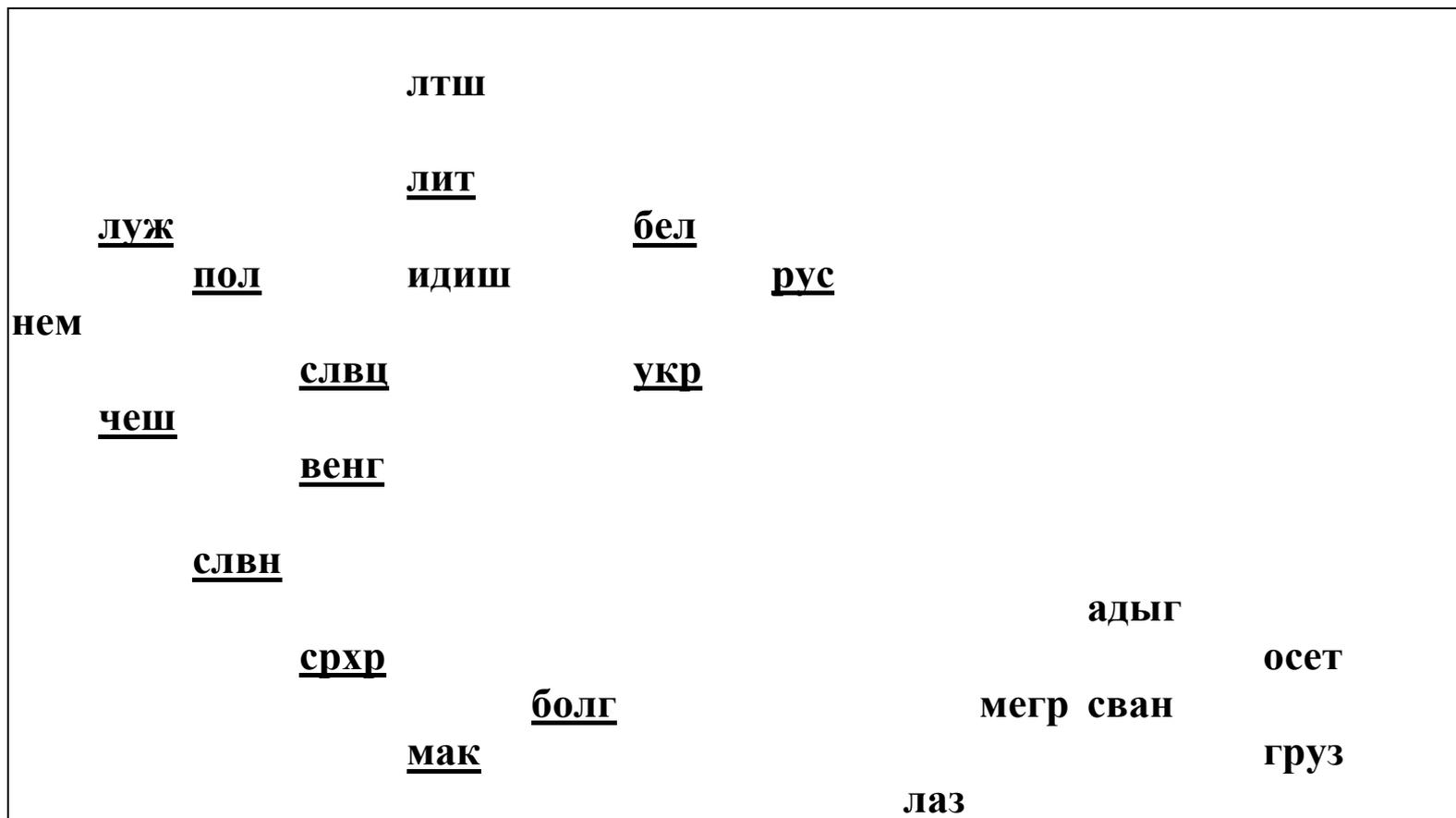
Типологические параметры

3.5. Синтаксические средства вторичной имперфективации: 0 нет; 1 да.



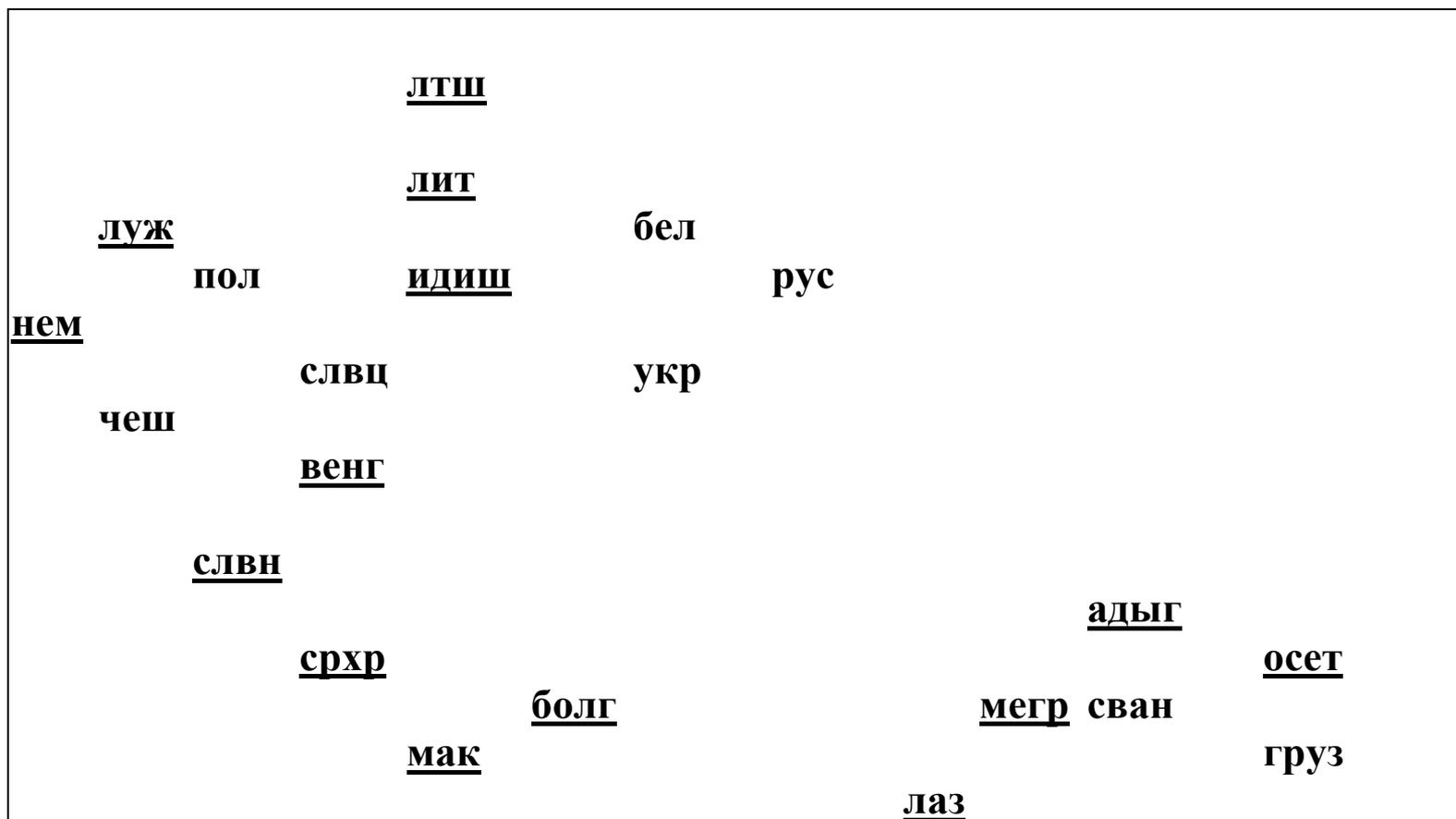
Типологические параметры

3.6. Средства перфективации помимо превербов: 1 да; 0 нет.



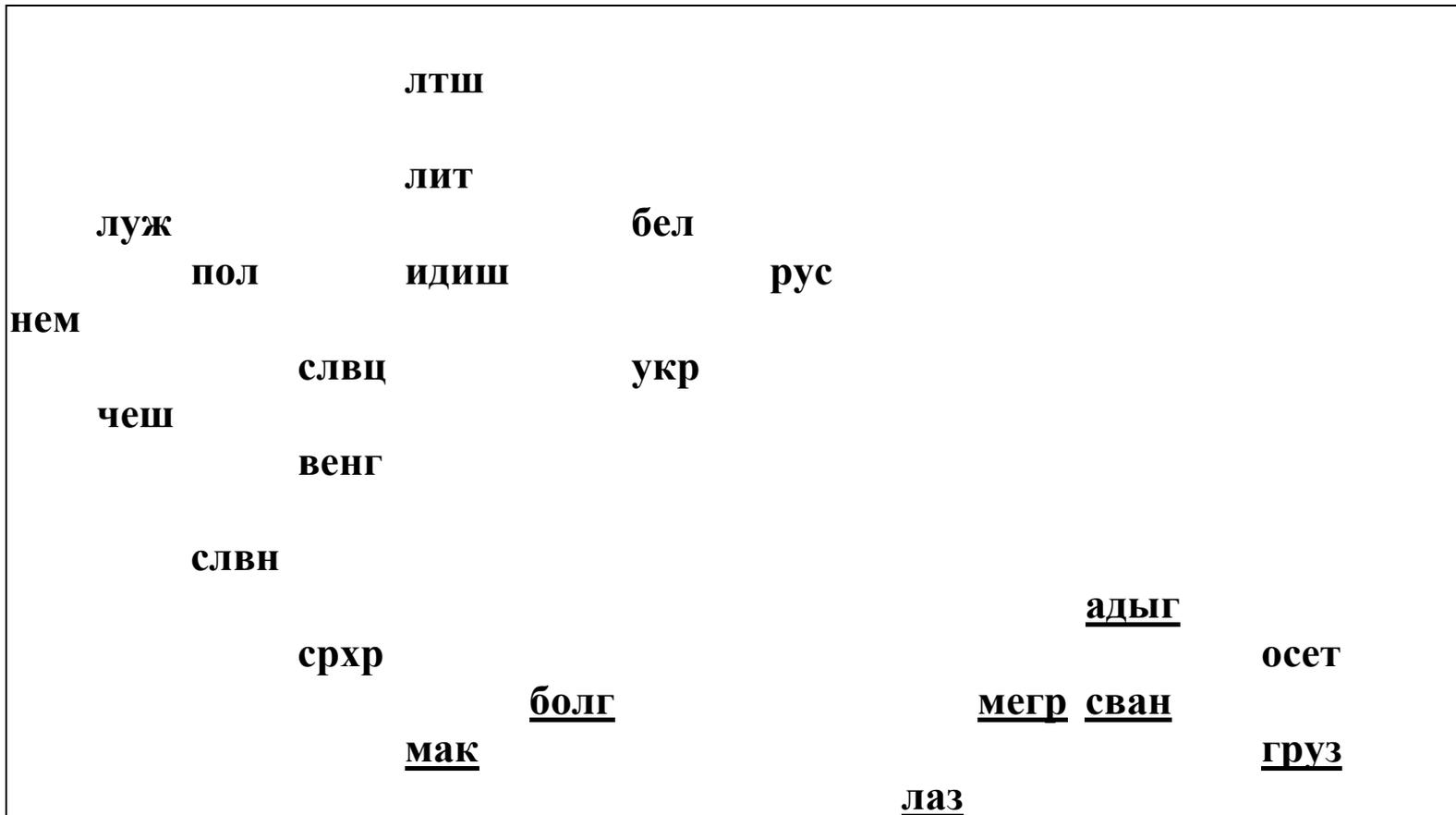
Типологические параметры

3.7. Аспектуально-нейтральный футурум: 1 да; 0 нет.



Типологические параметры

3.8. Аорист и имперфект наряду с префиксальным перфективом: 0 нет или маргинальны; 1 да.



Сводная таблица

| признаки | рус | пол | чеш | слвн | срхр | болг | мак | луж | лит | лтш | нем | идиш | венг | осет | груз | сван | мегр | лаз | адыг |
|-----------------------|-----|-----|-----|------|------|------|-----|-----|-----|-----|-----|------|------|------|------|------|------|-----|------|
| 1.1 отделимость | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| 1.2 др. префиксы | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1.3 позиция | | | | | | | | | 2 | 3 | | | | 1 | 1 | 1 | 2 | 2 | 2 |
| 1.4 итерация | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| 1.5 морф.подтипы | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 2.1 дейксис отд. | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2.2 перфективация | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 2.3. чистая перф. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 2.4. делимитатив | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2.5. непред. движ. | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2.6. непред. другие | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 3.1. перф. презенс | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 |
| 3.2. перф.през. = фут | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 3.3. фаз. глаголы | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | | 1 | 1 | 1 |
| 3.4. 2импф морф | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 3.5. 2импф синт | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3.6. непрев. перф. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3.7. нейтр. футур. | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| 3.8. аор/имперф | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |

Изоплеты?

Какой набор значений признаков принять за «эталон»?

Например, все «1».

Но таким набором не характеризуется ни один язык!

верхнелужицкий: 13

осетинский, мегрельский: 12

литовский, грузинский, лазский: 11

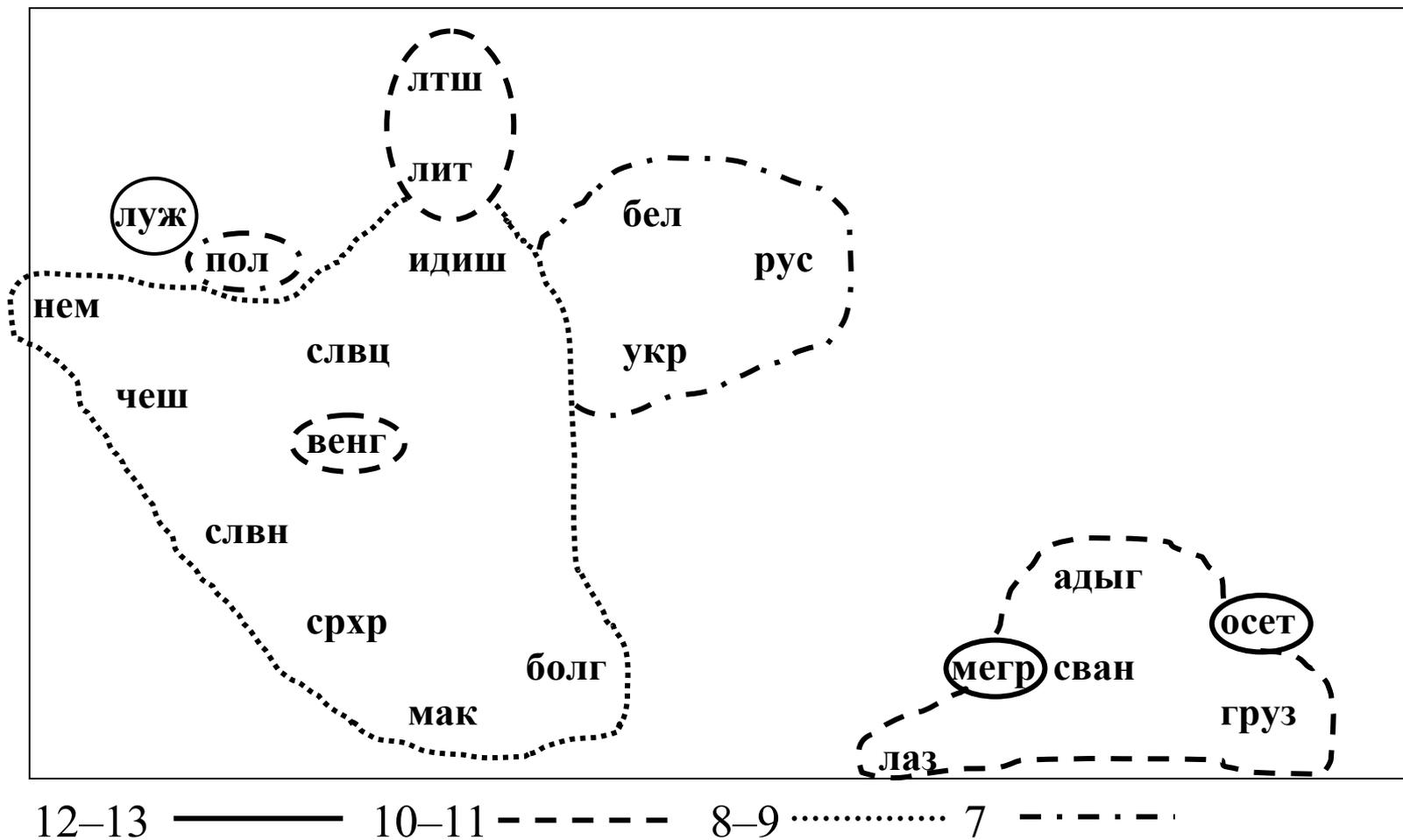
латышский, венгерский, сванский, адыгейский: 10

словенский, немецкий, идиш: 9

чешский, словацкий, сербохорватский, болгарский, македонский: 8

русский, украинский, белорусский, польский: 7

Карта изоплет по критерию «1»



Изоплеты?

Более того, легко заметить, что многие из языков с одинаковым числом «единиц» на самом деле имеют между собою мало общего:

литовский и грузинский (11): 9 общих признаков из 19 (vs. грузинский и лазский: 17 из 19)

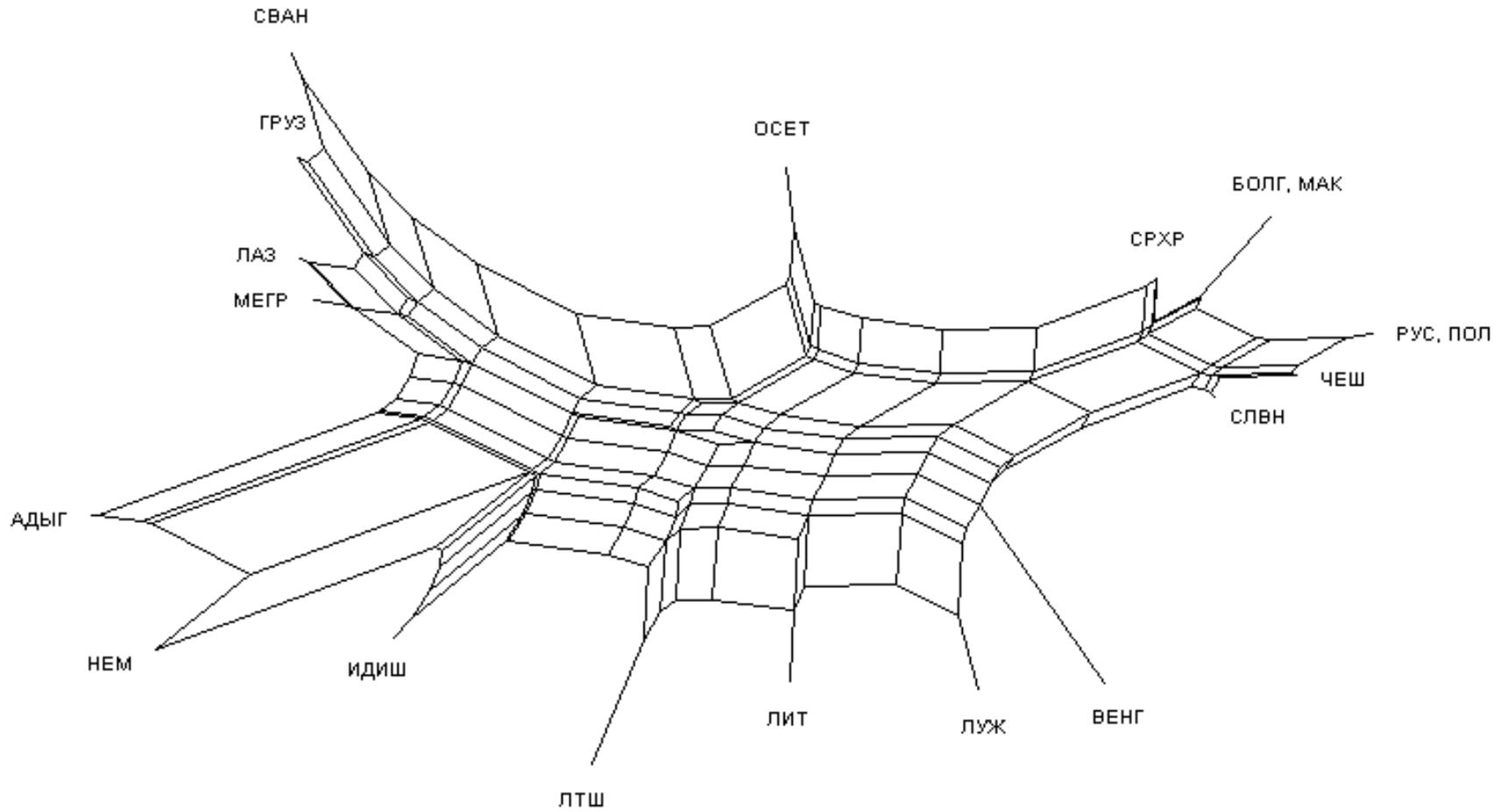
венгерский и адыгейский (10): 6 общих признаков из 19

Тем не менее, сравнение с таким «эталоном» показывает:

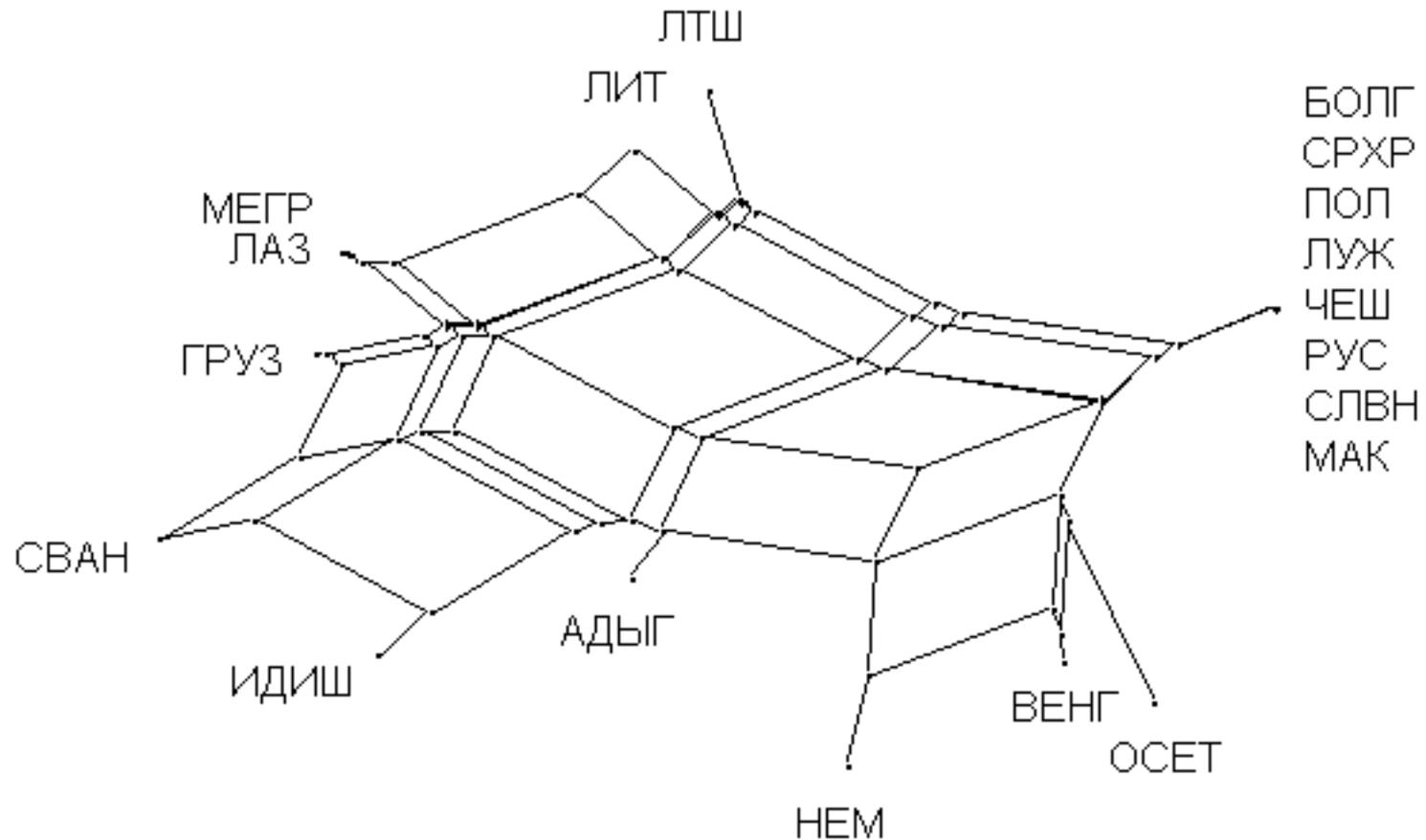
– основные славянские языки имеют больше отрицательных значений признаков, чем неславянские, в особенности кавказские;

– верхнелужицкий существенно выделяется на фоне славянских.

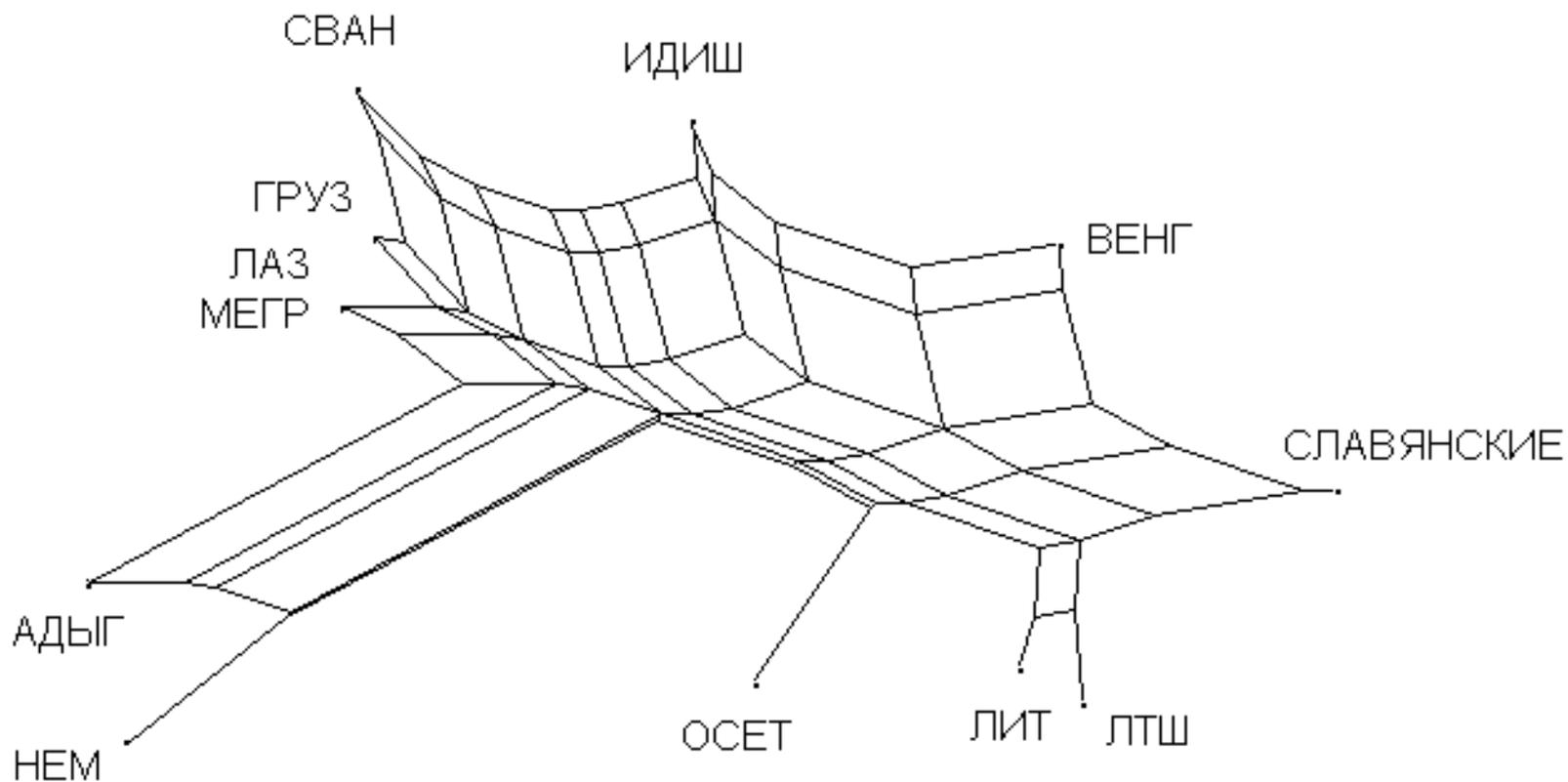
Neighbor Net



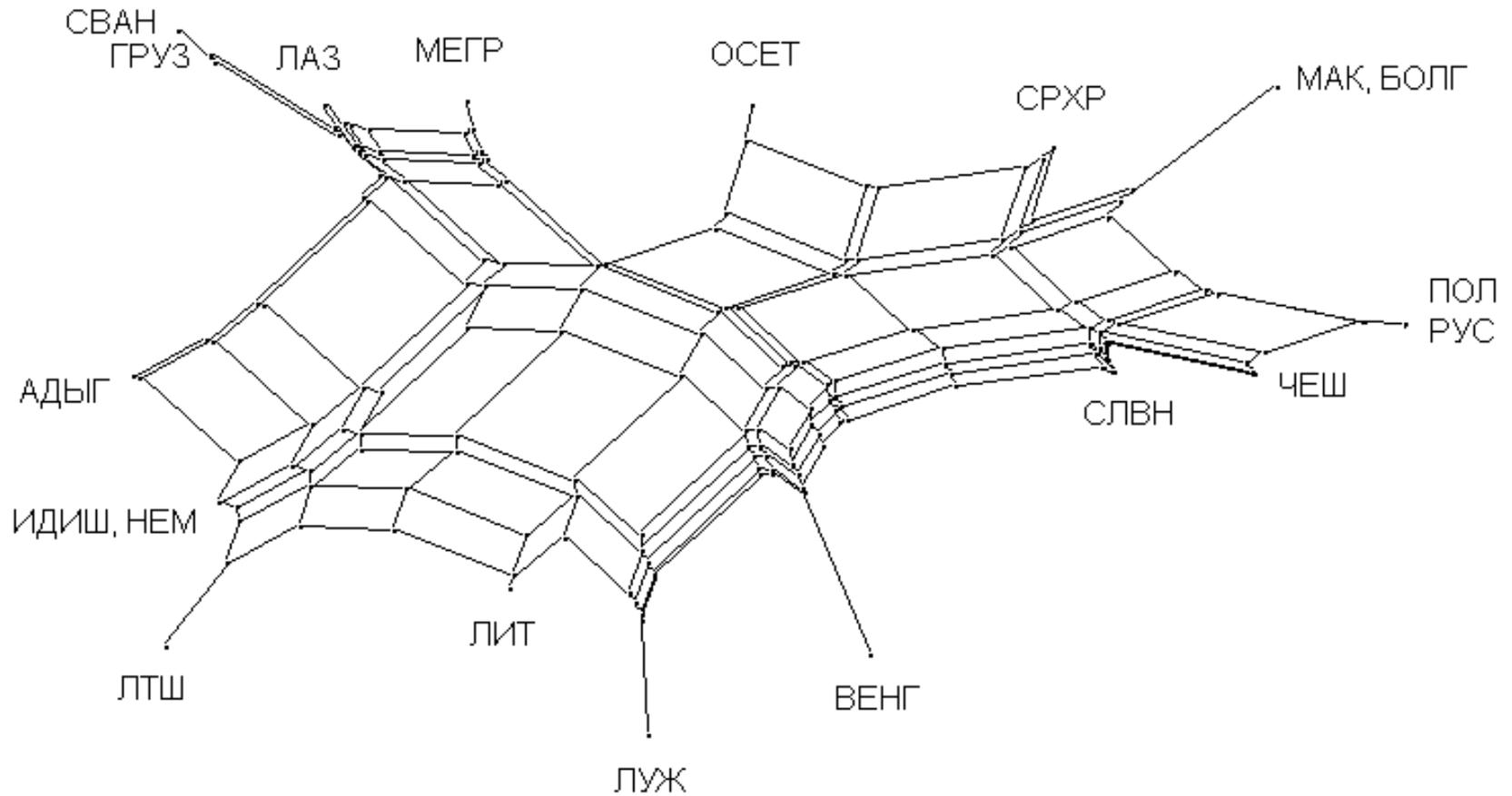
Neighbor Net: только морфология (1.1–1.5)



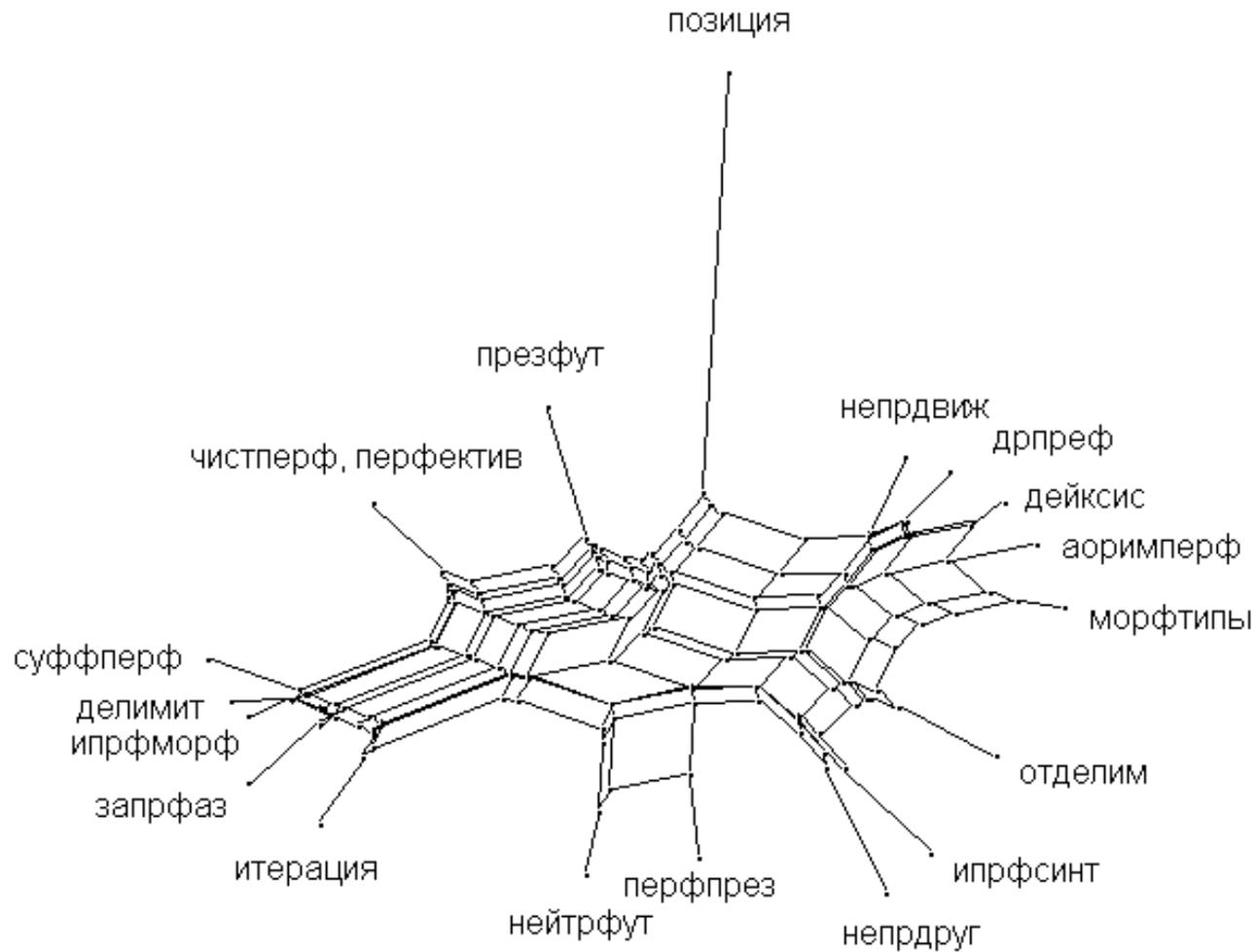
Neighbor Net: только превербы (1.1–2.4)



Neighbor Net: только функции (2.5–3.8)



Neighbor Net: кластеризация признаков



Мера степени сходства между языками

$$M(L1,L2) = 100\% \cdot \frac{|\{\alpha : \alpha(L1) = \alpha(L2)\}|}{|\{\alpha : \alpha(L1) \neq \emptyset \ \& \ \alpha(L2) \neq \emptyset\}|}$$

Степень сходства пары языков определяется как процентное отношение числа признаков, значения которых для данной пары одинаковы, к числу признаков, значения которых определены для каждого из членов пары.

Например, для русского определены 18 признаков из 19, для сванского — 17, общее число определённых для данной пары признаков — 16, число совпадений — 6, значит $M(\text{рус,сван}) = 100\% \cdot 6/16 = 37,5\%$

Для простоты, в дальнейшем M округляется до целого.

Мера степени сходства между языками

Средняя степень сходства для $L1$ $M^*(L1)$ — среднее арифметическое степеней сходства для всех пар $(L1, Ln)$, кроме $(L1, L1)$ (NB подсчитывается по полной таблице с учётом всех языков).

Аналогично может быть подсчитана средняя степень сходства для любого множества языков.

Также принимается в расчёт стандартное отклонение M .

Мера степени сходства между языками

| | рус | пол | чеш | слвн | срхр | болг | мак | луж | лит | лтш | нем | идиш | венг | осет | груз | сван | мегр | лаз | адыг | среднее |
|------|-----|-----|-----|------|------|------|-----|-----|-----|-----|-----|------|------|------|------|------|------|-----|---------|---------|
| рус | 100 | 94 | 88 | 82 | 88 | 88 | 65 | 53 | 35 | 18 | 29 | 59 | 53 | 41 | 40 | 41 | 35 | 18 | 63 (28) | |
| пол | 100 | 94 | 88 | 82 | 88 | 88 | 65 | 53 | 35 | 18 | 29 | 59 | 53 | 41 | 40 | 41 | 35 | 18 | 63 (28) | |
| чеш | 94 | 94 | 94 | 88 | 82 | 82 | 71 | 59 | 41 | 24 | 35 | 65 | 59 | 47 | 40 | 47 | 41 | 24 | 65 (25) | |
| слвн | 88 | 88 | 94 | 94 | 88 | 88 | 76 | 65 | 47 | 29 | 41 | 71 | 65 | 41 | 33 | 53 | 47 | 29 | 67 (23) | |
| срхр | 82 | 82 | 88 | 94 | 94 | 94 | 71 | 71 | 53 | 35 | 47 | 65 | 71 | 35 | 27 | 47 | 41 | 35 | 66 (22) | |
| болг | 88 | 88 | 82 | 88 | 94 | 100 | 65 | 65 | 47 | 29 | 41 | 59 | 65 | 29 | 27 | 41 | 35 | 29 | 63 (25) | |
| мак | 88 | 88 | 82 | 88 | 94 | 100 | 65 | 65 | 47 | 29 | 41 | 59 | 65 | 29 | 27 | 41 | 35 | 29 | 63 (25) | |
| луж | 65 | 65 | 71 | 76 | 71 | 65 | 65 | 76 | 71 | 41 | 53 | 71 | 53 | 41 | 27 | 53 | 47 | 41 | 60 (13) | |
| лит | 53 | 53 | 59 | 65 | 71 | 65 | 65 | 76 | 78 | 41 | 65 | 47 | 61 | 50 | 38 | 67 | 61 | 56 | 59 (10) | |
| лтш | 35 | 35 | 41 | 47 | 53 | 47 | 47 | 71 | 78 | 47 | 71 | 53 | 56 | 56 | 44 | 56 | 61 | 56 | 51 (12) | |
| нем | 18 | 18 | 24 | 29 | 35 | 29 | 29 | 41 | 41 | 47 | 76 | 47 | 53 | 53 | 53 | 53 | 58 | 88 | 41 (19) | |
| идиш | 29 | 29 | 35 | 41 | 47 | 41 | 41 | 53 | 65 | 71 | 76 | 59 | 53 | 65 | 67 | 65 | 71 | 65 | 51 (16) | |
| венг | 59 | 59 | 65 | 71 | 65 | 59 | 59 | 71 | 47 | 53 | 47 | 59 | 47 | 47 | 47 | 47 | 53 | 35 | 56 (9) | |
| осет | 53 | 53 | 59 | 65 | 71 | 65 | 65 | 53 | 61 | 56 | 53 | 53 | 47 | 56 | 63 | 61 | 56 | 50 | 57 (6) | |
| груз | 41 | 41 | 47 | 41 | 35 | 29 | 29 | 41 | 50 | 56 | 53 | 65 | 47 | 56 | 94 | 83 | 89 | 61 | 52 (18) | |
| сван | 40 | 40 | 40 | 33 | 27 | 27 | 27 | 26 | 38 | 44 | 53 | 67 | 47 | 63 | 94 | 75 | 81 | 50 | 47 (18) | |
| мегр | 41 | 41 | 47 | 53 | 47 | 41 | 41 | 53 | 67 | 56 | 53 | 65 | 47 | 61 | 83 | 75 | 94 | 67 | 55 (15) | |
| лаз | 35 | 35 | 41 | 47 | 41 | 35 | 35 | 47 | 61 | 61 | 58 | 71 | 53 | 56 | 89 | 81 | 94 | 72 | 54 (18) | |
| адыг | 18 | 18 | 24 | 29 | 35 | 29 | 29 | 41 | 56 | 56 | 88 | 65 | 35 | 50 | 61 | 50 | 67 | 72 | 42 (20) | |

Средняя степень сходства для всех пар: 58 % (с.о. 22)

Мера степени сходства между языками

М* для славянских языков (без лужицкого) — 89 % (с.о. 7)

М* для картвельских языков — 87 % (с.о. 7)

М* для остальных языков (кроме адыгейского) — 60 % (с.о. 11)

При этом исключение верхнелужицкого и/или осетинского почти не меняет результат.

М* для кавказских языков — 70 % (с.о. 16)

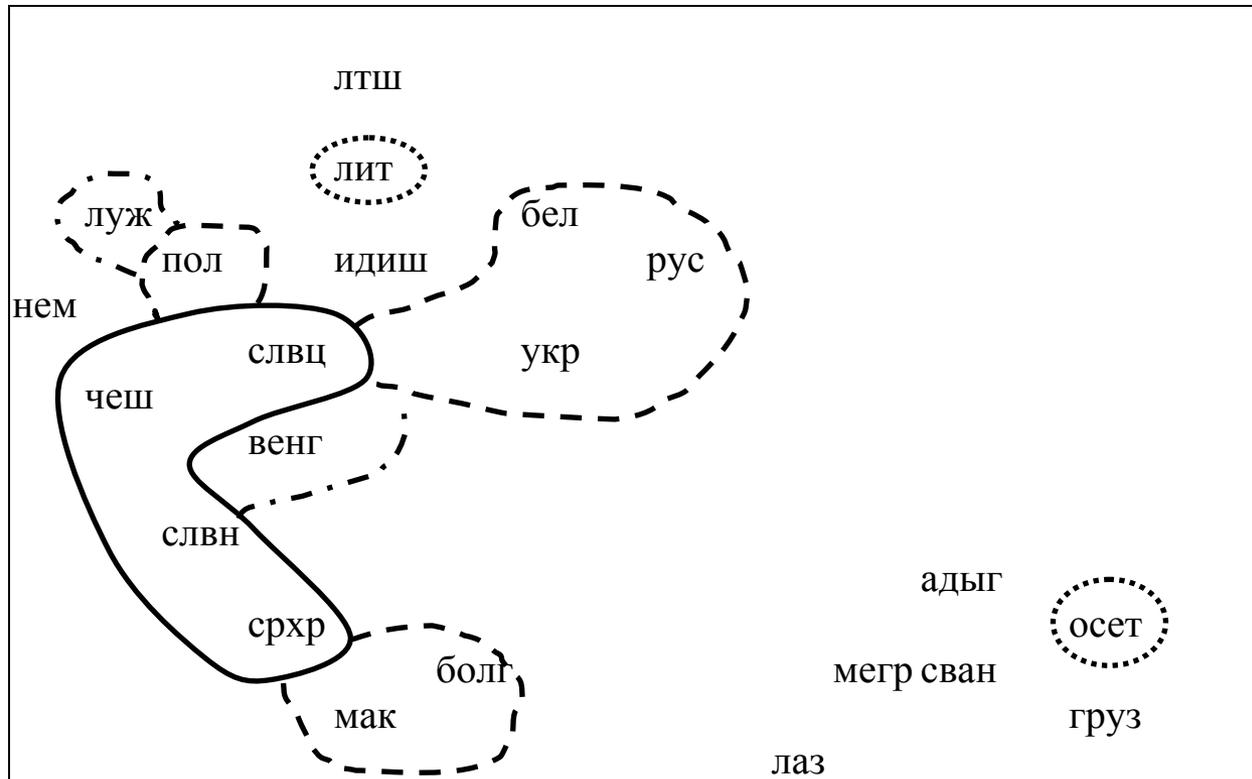
без адыгейского — 74 % (с.о. 16)

без осетинского — 78 % (с.о. 13)

Эффект сходства аутсайдеров — немецкий и адыгейский.

Снова изоплеты

В качестве эталона разумно взять язык, максимально похожий на остальные. Словенский ($M^* = 67$).

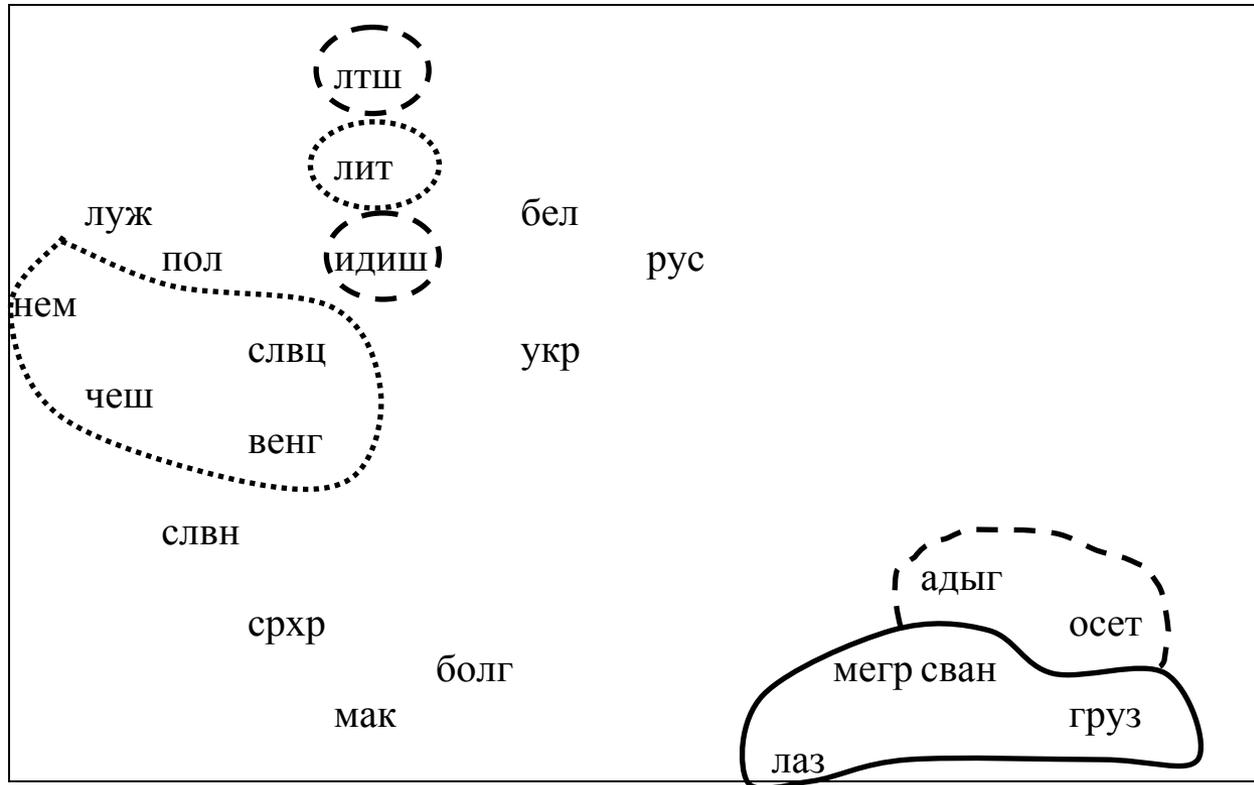


— > 15 общих признаков
- - - 14-15 общих признаков

- · - · 12-13 общих признаков
····· 10-11 общих признаков

Снова изоплеты

Для сравнения: что получится, если в качестве «эталона» взять грузинский ($M^* = 52$).



— > 14 общих признаков
- - - 10–11 общих признаков

..... 8–9 общих признаков

Некоторые выводы

Высокая межъязыковая вариативность, нет сильных корреляций между признаками.

Два основных кластера, основанных на генетическом родстве, — славянский и картвельский, характеризующиеся различными наборами значений признаков.

Остальные языки образуют промежуточную область, не будучи существенно сходны ни друг с другом, ни с основными кластерами, однако тяготея скорее к славянскому прототипу, нежели к картвельскому (даже осетинский!).

Некоторые выводы

Контактные явления могут «пересиливать» родство, ср. верхнелужицкий vs. остальные славянские, однако их роль не следует преувеличивать: она скорее локальна (контакт затрагивает только отдельные признаки, по большей части функциональные), нежели глобальна.