

Филогенетические исследования в лингвистике: Tree Model vs. Wave model

Деревья, неукорененные деревья, NeighbourNets

Г. Мороз

11 марта 2016

Как описать отношения между родственными идиомами?

Tree model:

[Schleicher 1852]

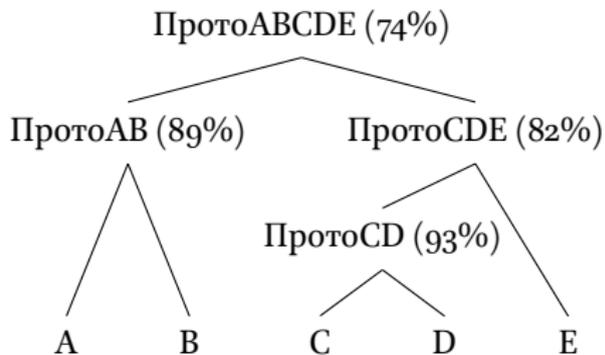
- на основе количества регулярных соответствий некоторых языковых единиц (лексических, звуковых или других) строится гипотетическая цепочка изменений идиомов, объясняющая разнообразие некоторого синхронного состояния
- предполагается, что смоделированная цепочка изменений отражает филогенетическое ветвление идиомов или групп идиомов

Древесная модель

tree vs. wave

NeighborNet

адыгские
идиомы

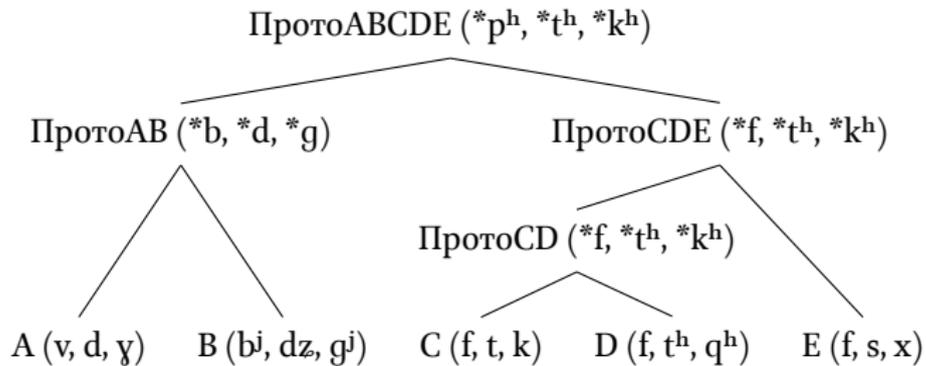


Древесная модель

tree vs. wave

NeighborNet

адыгские
идиомы

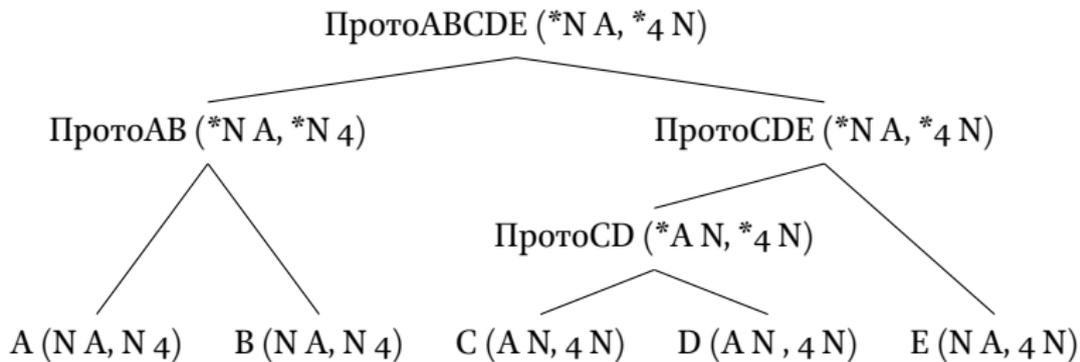


Древесная модель

tree vs. wave

NeighborNet

адыгские
идиомы



N --- имя, A --- прилагательное, 4 --- числительное

Древесная модель

Проблемы:

- регулярность соответствий --- святое
 - спорадические переходы создают "шум"
 - параллельное независимое развитие "шумит"
- как определить направление изменений?
см. [Fallon 2013] (абруптивные), [Cser 2003] (лениция и усиление) и другие...
кроме того некоторые асимметрии звуковых изменений объясняют фонетические модели Джона Охалы

Древесная модель

Проблемы:

- регулярность соответствий --- святое
 - спорадические переходы создают "шум"
 - параллельное независимое развитие "шумит"
- как определить направление изменений?
см. [Fallon 2013] (абруптивные), [Cser 2003] (лениция и усиление) и другие...
кроме того некоторые асимметрии звуковых изменений объясняют фонетические модели Джона Охалы
- контакты, субстраты, суперстраты...
- интенсивные контакты или диффузия [Dixon 1997]
- запрет на двойную аффилиацию
- историчность...

Модель прерванного равновесия

Данная модель предложена в работе [Dixon 1997] и языковое изменение она описывает, как чередование:

- периодов *равновесия* (equilibrium)
языковые признаки подвергаются значительной конвергенции

Модель прерванного равновесия

Данная модель предложена в работе [Dixon 1997] и языковое изменение она описывает, как чередование:

- периодов *равновесия* (equilibrium)
языковые признаки подвергаются значительной конвергенции
- и периодов *прерывания равновесия* (punctated equilibrium)

Волновая модель

Wave model:

[Schmidt 1872]



tree vs. wave

NeighborNet

адыгские
идиомы

Филогенетические сети

Под понятием филогенетические сети объединяют много разных методов, некоторую классификацию которых можно посмотреть в работе [Huson, Bryant 2006].

- NeighborNet [Bryant, Moulton 2002]

Искусственный пример 1

tree vs. wave

NeighborNet

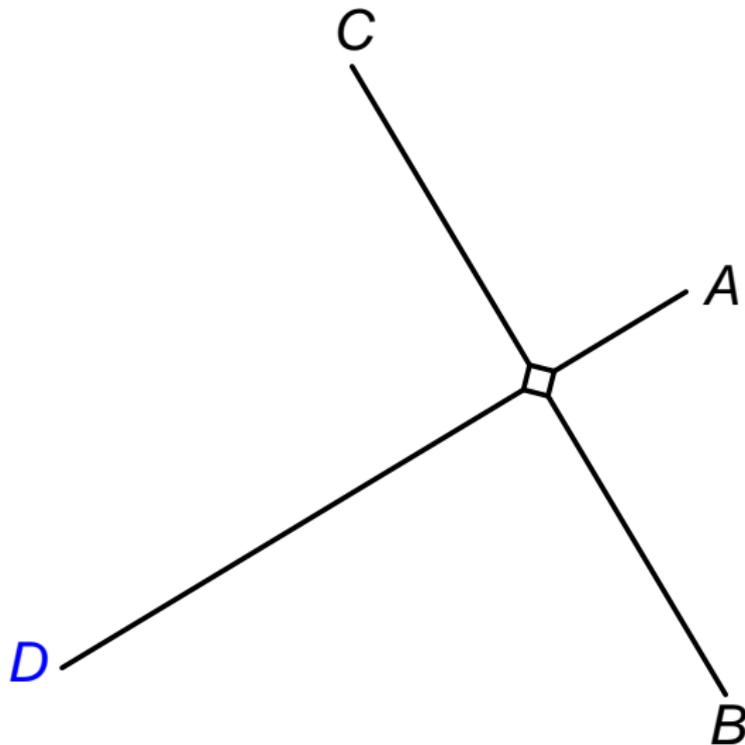
адаптивные
идиомы

	*p	*b	*t	*d	*k
A	p	b	t	d	k
B	p	b	ts	d	k
C	p	b	t	t	k
D	p	v	t	d	x

	*p	*b	*t	*d	*k
A	0	0	0	0	0
B	0	0	1	0	0
C	0	0	0	1	0
D	0	1	0	0	1

	A	B	C
B	1.000000		
C	1.000000	1.414214	
D	1.414214	1.732051	1.732051

Искусственный пример 1



Искусственный пример 1

tree vs. wave

NeighborNet

адыгские
идиомы

```
library(phangorn)
dat <- data.frame(
  A = c(0, 0, 0, 0, 0),
  B = c(0, 0, 1, 0, 0),
  C = c(0, 0, 0, 1, 0),
  D = c(0, 1, 0, 0, 1))
dat.dist <- dist(t(dat))
nnet <- neighborNet(dat.dist)
par("mar- rep(1, 4))
plot(nnet, "2D"
      # расстояние
      # NeighborNet
      # исправление границ
      # график
      # размер
      cex = 2)
```

Искусственный пример 1

tree vs. wave

NeighborNet

адыгские
идиомы

```
library(phangorn)
dat <- data.frame(
  A = c(0, 0, 0, 0, 0),
  B = c(0, 0, 1, 0, 0),
  C = c(0, 0, 0, 1, 0),
  D = c(0, 1, 0, 0, 1))
dat.dist <- dist(t(dat))
nnet <- neighborNet(dat.dist)
par("mar= rep(1, 4))
plot(nnet, "3D")
```

расстояние
NeighborNet
исправление границ
график

Результат

Искусственный пример 2

tree vs. wave

NeighborNet

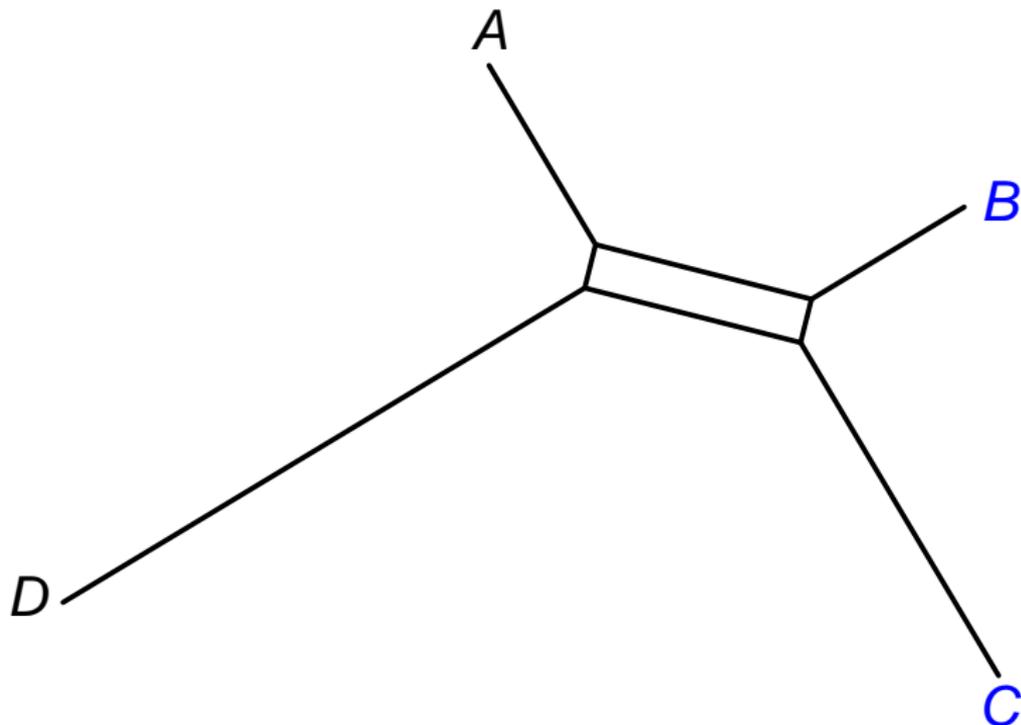
адагские
идиомы

	*p	*b	*t	*d	*k
A	p	b	t	d	k
B	p	b	ts	d	k
C	p	v	ts	t	k
D	p	v	t	d	x

	*p	*b	*t	*d	*k
A	0	0	0	0	0
B	0	0	1	0	0
C	0	0	1	1	0
D	0	1	0	0	1

	A	B	C
B	1.000000		
C	1.414214	1.000000	
D	1.414214	1.732051	2.000000

Искусственный пример 2



Искусственный пример 2

tree vs. wave

NeighborNet

адыгские
идиомы

```
library(phangorn)
dat <- data.frame(
  A = c(0, 0, 0, 0, 0),
  B = c(0, 0, 1, 0, 0),
  C = c(0, 0, 1, 1, 0),
  D = c(0, 1, 0, 0, 1))
dat.dist <- dist(t(dat))
nnet <- neighborNet(dat.dist)
par("mar- rep(1, 4))
plot(nnet, "2D"
      # расстояние
      # NeighborNet
      # исправление границ
      # график
      # размер
      cex = 2)
```

Искусственный пример 2

tree vs. wave

NeighborNet

адыгские
идиомы

```
library(phangorn)
dat <- data.frame(
  A = c(0, 0, 0, 0, 0),
  B = c(0, 0, 1, 0, 0),
  C = c(0, 0, 1, 1, 0),
  D = c(0, 1, 0, 0, 1))
dat.dist <- dist(t(dat))
nnet <- neighborNet(dat.dist)
par("mar= rep(1, 4))
plot(nnet, "3D")
```

расстояние
NeighborNet
исправление границ
график

Результат

Адыгские идиомы на территории РФ

tree vs. wave

NeighborNet

адыгские
идиомы

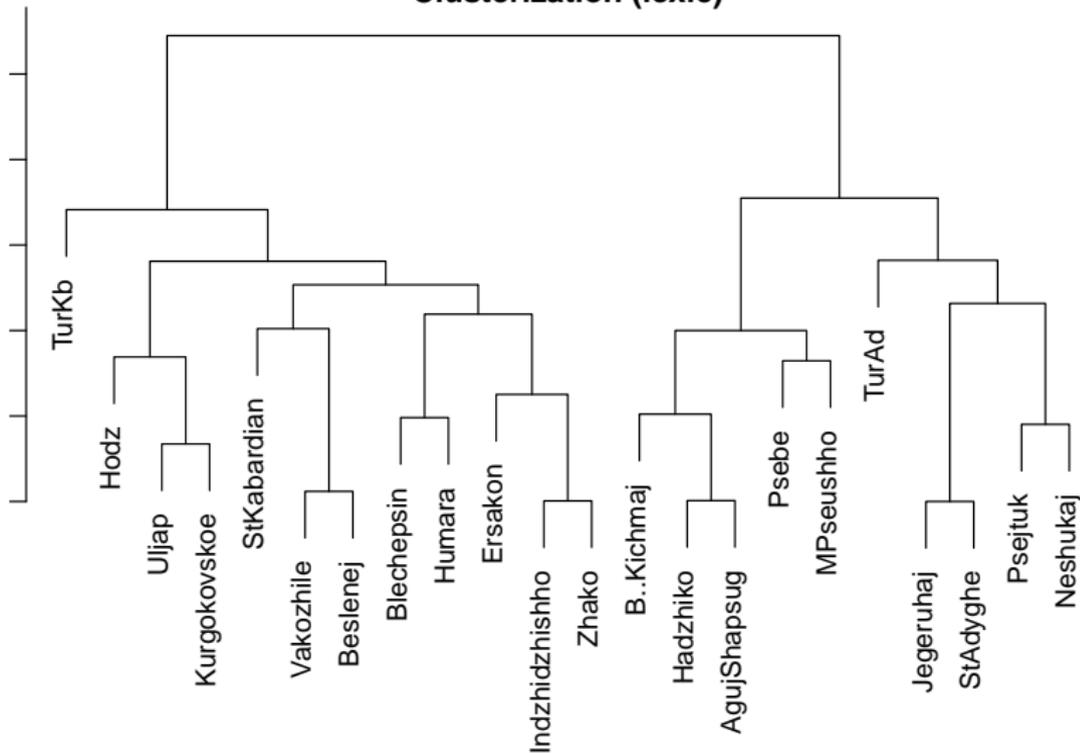
- литературный адыгейский язык
 - *темиргоевский диалект* (Республика Адыгея)
 - абадзехский диалект (РА)
 - бжедугский диалект (РА)
 - шапсугский диалект
 - прикубанские (РА)
 - причерноморские (Краснодарский Край)
 - хакучинский (КК)
- литературный кабардино-черкесский язык
 - *баксанский диалект* (Кабардино-Балкарская Республика)
 - малкинский диалект (КБР)
 - терские говоры (КБР, Ставропольский Край)
 - кубано-зеленьчукские говоры (Карачаево-Черкесская Республика)
 - кубанский диалект (РА)
 - бесленеевский диалект (РА, КК, КЧР)

Исследовались:

- литературный адыгейский язык
 - *темиргоевский диалект* (Кабехабль, Егерухай)
 - шапсугский вариант (М. Псеушхо)
 - абадзехский диалект
 - бжедугский диалект (Нешукай)
 - шапсугский диалект
 - прикубанские (Псейтук)
 - причерноморские (Агуй-Шапсуг, Псебе)
 - хакучинский (Хаджико, Б. Кичмай)
 - адыгейский идиом в Турции
- литературный кабардино-черкесский язык
 - *баксанский диалект*
 - малкинский диалект
 - терские говоры
 - кубано-зеленьчукские говоры (Хумара, Инджиджишхо, Жако, Эрсакон)
 - кубанский диалект (Ходзь, Блечепсин)
 - бесленеевский диалект (Уляп, Кургоковское, Вакожиле, Бесленей)
 - кабардинские идиом в Турции

Кластеризация по лексическим данным

Clusterization (lexic)

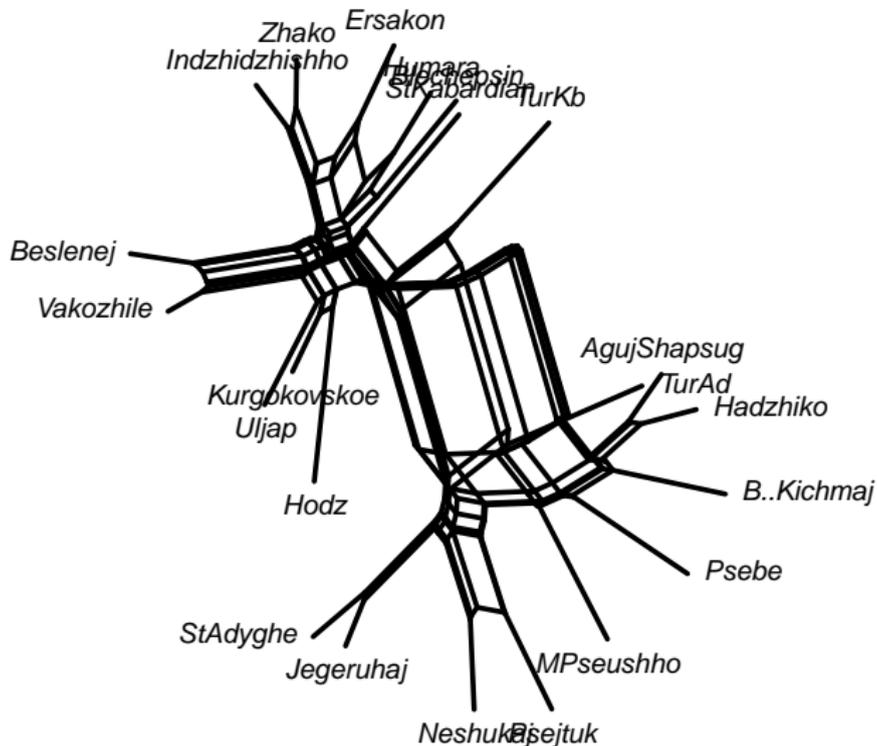


NeighborNet по лексическим данным

tree vs. wave

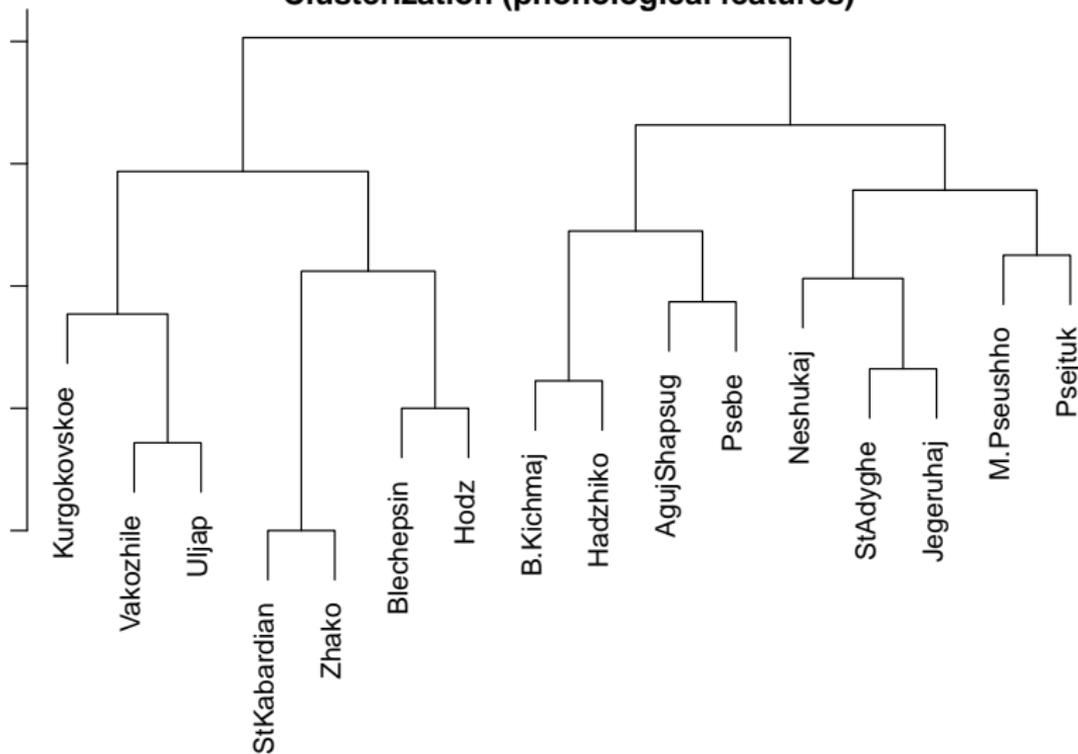
NeighborNet

адыгские
идиомы



Кластеризация по фонологическим признакам

Clusterization (phonological features)

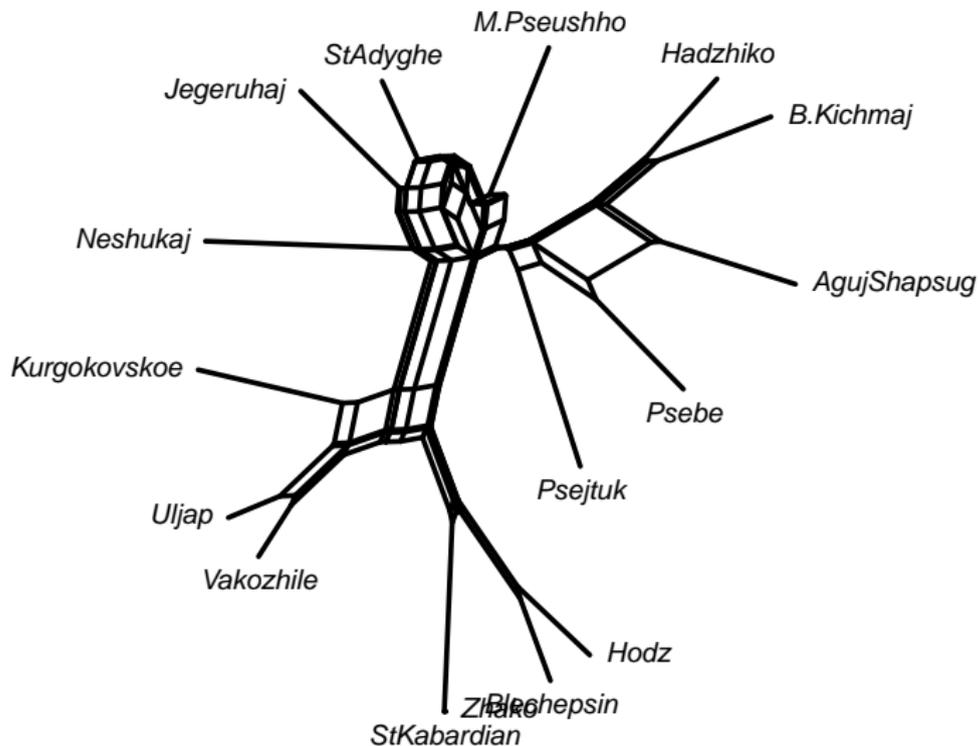


NeighborNet по фонологическим признакам

tree vs. wave

NeighborNet

адыгские
идиомы



3d-вариант
презентация доступна по адресу <http://goo.gl/hXdcJb>

tree vs. wave

NeighborNet

адыгские
идиомы

Спасибо за внимание!

Пишите письма

agricolamz@gmail.com

Список литературы

- Bryant, D., V. Moulton (2002). NeighborNet: An agglomerative method for the construction of planar phylogenetic networks. In *Algorithms in Bioinformatics*, pp. 375–391. Springer.
- Cser, A. (2003). *The typology and modelling of obstruent lenition and fortition processes*. Akadémiai Kiadó.
- Dixon, R. M. W. (1997). *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Fallon, P. D. (2013). *The synchronic and diachronic phonology of ejectives*. Routledge.
- Huson, D. H., D. Bryant (2006). Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution* 23(2), 254–267.
- Schleicher, A. (1852). *Die Formenlehre der kirchenslawischen Sprache*. Bonn: H. B. König.
- Schmidt, J. (1872). *Die verwandtschaftsverhältnisse der indogermanischen sprachen*. Weimar: Hermann Böhlau.