

# The comparative method in historical linguistics

Gerhard Jäger

ESLLI 2016

# The comparative method

(from Ross and Durie 1996)

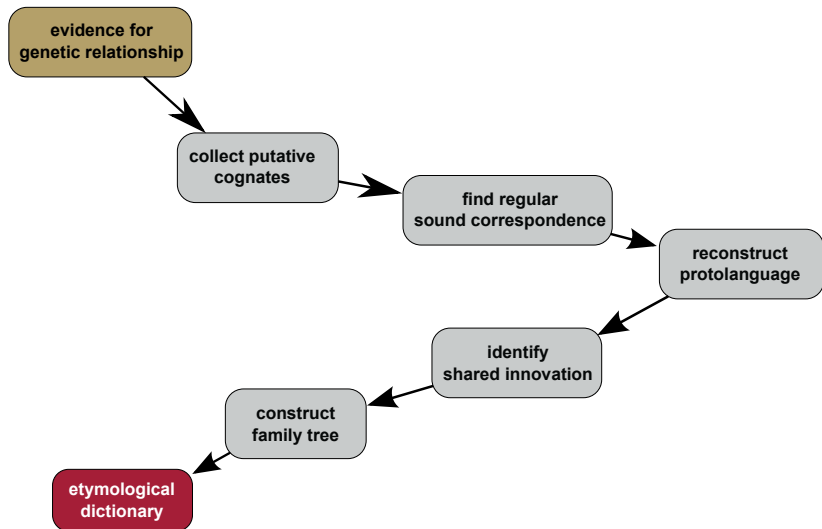
- dominant paradigm in (non-computational) historical linguistics
- developed during the 19th century
- originally applied mostly to Indo-European, but applicable to all language families
- central axiom:  
*Neogrammarian Hypothesis* Sound laws apply without exception.

## ● Workflow

- 1 Determine on the strength of diagnostic evidence that a set of languages are genetically related, that is, that they constitute a 'family';
- 2 Collect putative cognate sets for the family (both morphological paradigms and lexical items).
- 3 Work out the sound correspondences from the cognate sets, putting 'irregular' cognate sets on one side;
- 4 Reconstruct the protolanguage of the family as follows:
  - a. Reconstruct the protophonology from the sound correspondences worked out in (3), using conventional wisdom regarding the directions of sound changes.
  - b. Reconstruct protomorphemes (both morphological paradigms and lexical items) from the cognate sets collected in (2), using the protophonology reconstructed in (4a).

- 5 Establish innovations (phonological, lexical, semantic, morphological, morphosyntactic) shared by groups of languages within the family relative to the reconstructed protolanguage.
- 6 Tabulate the innovations established in (5) to arrive at an internal classification of the family, a 'family tree'.
- 7 Construct an etymological dictionary, tracing borrowings, semantic change, and so forth, for the lexicon of the family (or of one language of the family).

# Workflow



# Diagnostic evidence for genetic relatedness

- sometimes self-evident (e.g. Slavic)
- similarities in morphological paradigms (example from Clackson 2007, 124)

PIE	Sanskrit	Greek	Latin	Gothic	Lith.	O.C.S.
1. <i>*h<sub>1</sub>és-mi</i>	<i>ásmi</i>	<i>eimí</i>	<i>sum</i>	<i>im</i>	<i>esmì</i>	<i>jesmǐ</i>
2. <i>*h<sub>1</sub>és-si</i> (or <i>*h<sub>1</sub>ésti</i> )	<i>ási</i>	<i>eí</i>	<i>ess, es</i>	<i>is</i>	<i>esi</i>	<i>jesǐ</i>
3. <i>*h<sub>1</sub>és-ti</i>	<i>asti</i>	<i>estí</i>	<i>est</i>	<i>ist</i>	<i>ēsti</i>	<i>jestǔ</i>
4. <i>*h<sub>1</sub>s-mé</i>	<i>smás</i>	<i>esmén</i>	<i>sumus</i>	<i>sijum</i>	<i>esme</i>	<i>jesmǔ</i>
5. <i>*h<sub>1</sub>s-té</i>	<i>sthá</i>	<i>éste</i>	<i>estis</i>	<i>sijuþ</i>	<i>este</i>	<i>jeste</i>
6. <i>*h<sub>1</sub>s-énti</i>	<i>sánti</i>	<i>eisí</i>	<i>sunt</i>	<i>sind</i>		<i>sputǐ</i>

- overwhelming lexical similarities ... ⇒

# Diagnostic evidence for genetic relatedness

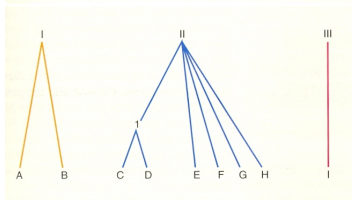
Eine Klassifikationsübung nach der vergleichenden Methode à la Merritt Ruhlen:

Sprache	zwei	drei	ich	du	wer?	nicht	Mutter	Vater	Zahn	Herz	Fuß	Maus	er trägt
<b>A</b>	ʔiθn-	θalāθ-	-ni	-ka	man	lā	ʔumm-	abū	sinn	lubb	rijl-	fār	yaħmil-
<b>B</b>	ʃn-	šaloš	-ni	-ka	mi	lo	ʔem	aβ	šen	leβ	regel	ʃaħbər	nošeh
<b>C</b>	duvā	trāyas	mām	tuvām	kás	ná	mātár	pitár-	dant-	ħʔd-	pád	muš-	bhárati
<b>D</b>	duva	θrāyō	mām	tuvəm	čiš	naē-	mātar-	pitár-	dantan-	zərəd	paiðya		baraiti
<b>E</b>	duo	treis	eme	sú	tis	ou(k)	māter	pater	odón	kardiā	pod-	mūs	phérei
<b>F</b>	duo	trēs	mē	tū	kwis	ne-	māter	pater	dent-	kord-	ped-	mūs	fert
<b>G</b>	twai	θreis	mik	θu	hwas	ni	aiθei	faðar	tunθus	ħaírtō	fōt		baíriθ
<b>H</b>	dó	trí	-m	tú	kía	ní-	máθir	aθir	dēt	kríde	traig	lux	berid
<b>I</b>	iki	üč	ben-i	sen	kim	deyil	anne	baba	diš	kalp	ayak	sičan	tašiyor

# Diagnostic evidence for genetic relatedness

Eine Klassifikationsübung nach der vergleichenden Methode à la Merritt Ruhlen:

Sprache	zwei	drei	ich	du	wer?	nicht	Mutter	Vater	Zahn	Herz	Fuß	Maus	er trägt
A	ʔiθn-	θalāθ-	-ni	-ka	man	lā	ʔumm-	abū	sinn	iubb	rijl-	fār	yaθmil-
B	fɲ-	šaloš	-ni	-ka	mi	lo	ʔem	aβ	šen	leβ	regel	ʕaʕbər	nošeh
C	duvā	trāyas	mām	tuvām	kás	ná	mātár	pitár-	dant-	hʔd-	pád	muš-	bhárati
D	duva	θrāyō	mām	tuvəm	čiš	naē-	mātár-	pitár-	dantan-	zərəd	paiðya		baraiti
E	duo	treis	eme	sú	tfs	ou(k)	mäter	pater	odón	kardiā	pod-	müs	phérei
F	duo	trēs	mē	tū	kwis	ne-	mäter	pater	dent-	kord-	ped-	müs	fert
G	twai	θreis	mik	θu	hwas	ni	aíθei	faðar	tunθus	haírtō	fōt		baíriθ
H	dó	trí	-m	tú	kía	ní-	máθir	aθir	dēt	kríde	traig	lux	berid
I	iki	üč	ben-i	sen	kim	deyil	anne	baba	diš	kalp	ayak	sičan	tašiyor



Klassifizieren Sie die angegebenen neun Sprachen (von A bis I) in Familien und Unterfamilien und vergleichen Sie den Wortschatz für die 13 Wörter, die hier in phonetischer Umschrift geboten werden. Lösung: Sprache A und B (Arabisch und Hebräisch) gehören zur Familie der semitischen Sprachen. Die sechs Sprachen C bis H (Sanskrit, Awestisch, Altgriechisch, Latein, Gotisch und Altirisch) sind indogermanische Sprachen. I (Türkisch) läßt sich keiner Familie zuordnen. Mit einer längeren Wortliste kann man nach demselben Verfahren die Familien wieder in Überfamilien einteilen usw. Der Stammbaum, den man so erhält, würde dann beweisen, daß alle Sprachen von einer Muttersprache abstammen.

chisch, Latein, Gotisch und Altirisch) sind indogermanische Sprachen. I (Türkisch) läßt sich keiner Familie zuordnen. Mit einer längeren Wortliste kann man nach demselben Verfahren die Familien wieder in Überfamilien einteilen usw. Der Stammbaum, den man so erhält, würde dann beweisen, daß alle Sprachen von einer Muttersprache abstammen.



# Example

	<i>two</i>	<i>three</i>	<i>tooth</i>	<i>foot</i>	<i>heart</i>	<i>skin</i>
<b>Ancient Greek</b>	dýo	treis	odús	pu:s	kardía:	dérma
<b>Dutch</b>	twe	dri	tant	vut	hart	hœyt
<b>Latin</b>	'duo	treis	dens	pe:s	kor	'kutis
<b>Old Church Slavonic</b>	dŭva	trĭje	zŏbŭ	noga	sŕj'dits'je	kŏza
<b>Old Norse</b>	tveir	þri:r	tŏn:	fo:tr	'yarta	hu:ð
<b>Russian</b>	dva	tri	zub	noga	s'jerdtsɛ	'kŏzɔ

# Example

- identify cognates

	<i>two</i>	<i>three</i>	<i>tooth</i>	<i>foot</i>	<i>heart</i>	<i>skin</i>
<b>Ancient Greek</b>	dýo	treis	odús	pu:s	kardía:	dérma
<b>Dutch</b>	twe	dri	tant	vut	hart	hœyt
<b>Latin</b>	'duo	treis	dens	pe:s	kor	'kutis
<b>Old Church Slavonic</b>	dŭva	trĭje	zŏbŭ	noga	sŕj'dits'je	kŏza
<b>Old Norse</b>	tveir	þri:r	tŏn:	fo:tr	'yarta	hu:ð
<b>Russian</b>	dva	tri	zub	noga	s'jerdtsɛ	'kŏza

# Example

- identify cognates

	<i>two</i>	<i>three</i>	<i>tooth</i>	<i>foot</i>	<i>heart</i>	<i>skin</i>
<b>Ancient Greek</b>	dýo	treis	odús	pu:s	kardía:	dérma
<b>Dutch</b>	twe	dri	tant	vut	hart	hoeyt
<b>Latin</b>	'duo	treis	dens	pe:s	kor	'kutis
<b>Old Church Slavonic</b>	dŭva	trĭje	zŏbŭ	noga	sŕj'ditsĭe	kŏza
<b>Old Norse</b>	tveir	þri:r	tŏn:	fŏ:tr	'yjarta	hu:ð
<b>Russian</b>	dva	tri	zub	noga	sĭerdtse	'kŏza

# Example

- identify cognates

	<i>two</i>	<i>three</i>	<i>tooth</i>	<i>foot</i>	<i>heart</i>	<i>skin</i>
<b>Ancient Greek</b>	dýo	treis	odús	pu:s	kardía:	dérma
<b>Dutch</b>	twe	dri	tant	vut	hart	hœyt
<b>Latin</b>	'duo	treis	dens	pe:s	kor	'kutis
<b>Old Church Slavonic</b>	dŭva	trĭje	zŏbŭ	noga	sŕj'ditsĭe	koza
<b>Old Norse</b>	tveir	þri:r	tŏn:	fo:tr	'yjarta	hu:ð
<b>Russian</b>	dva	tri	zub	noga	sĭerdtse	'koza

- establish **regular**, i.e. recurrent, sound correspondences

# Example

- identify cognates

	<i>two</i>	<i>three</i>	<i>tooth</i>	<i>foot</i>	<i>heart</i>	<i>skin</i>
<b>Ancient Greek</b>	dýo	treis	odús	pu:s	kardía:	dérma
<b>Dutch</b>	twe	dri	tant	vut	hart	hœyt
<b>Latin</b>	'duo	treis	dens	pe:s	kor	'kutis
<b>Old Church Slavonic</b>	dŭva	trĭje	zŏbŭ	noga	sŕj'dits'je	koza
<b>Old Norse</b>	tveir	þri:r	tŏn:	fo:tr	'yjarta	hu:ð
<b>Russian</b>	dva	tri	zub	noga	sjerdtse	'koza

- establish **regular**, i.e. recurrent, sound correspondences  
Greek, Latin, OCS, Russian [d] ~ Dutch, Norse [t]

# Example

- identify cognates

	<i>two</i>	<i>three</i>	<i>tooth</i>	<i>foot</i>	<i>heart</i>	<i>skin</i>
<b>Ancient Greek</b>	dýo	treis	odús	pu:s	kardía:	dérma
<b>Dutch</b>	twe	dri	tant	vut	hart	hœyt
<b>Latin</b>	'duo	treis	dens	pe:s	kor	'kutis
<b>Old Church Slavonic</b>	dŭva	trĭje	zŏbŭ	noga	sŕj'dits'je	koza
<b>Old Norse</b>	tveir	þri:r	tɔn:	fo:tr	'yjarta	hu:ð
<b>Russian</b>	dva	tri	zub	noga	s'jerdtsɛ	'koza

- establish **regular**, i.e. recurrent, sound correspondences

Greek, Latin, OCS, Russian [d] ~ Dutch, Norse [t]

Greek, Latin, OCS, Russian [t] ~ Dutch [d] ~ Norse [θ]

# Example

- identify cognates

	<i>two</i>	<i>three</i>	<i>tooth</i>	<i>foot</i>	<i>heart</i>	<i>skin</i>
<b>Ancient Greek</b>	dýo	treis	odús	pu:s	kardía:	dérma
<b>Dutch</b>	twe	dri	tant	vut	hart	hœyt
<b>Latin</b>	'duo	treis	dens	pe:s	kor	'kutis
<b>Old Church Slavonic</b>	dŭva	trĭje	zŏbŭ	noga	sŕj'dits'je	koza
<b>Old Norse</b>	tveir	þri:r	tŏn:	fŏ:tr	'yjarta	hu:ð
<b>Russian</b>	dva	tri	zub	noga	s'jerdtsɛ	'koza

- establish **regular**, i.e. recurrent, sound correspondences

Greek, Latin, OCS, Russian [d] ~ Dutch, Norse [t]

Greek, Latin, OCS, Russian [t] ~ Dutch [d] ~ Norse [θ]

Greek, Latin [k] ~ Dutch, Norse (?) [h] ~ OCS, Russian [s]

# Example

- reconstruct proto-forms and directionality of changes



# Example

- reconstruct proto-forms and directionality of changes
  - PIE [*\*d*] → Germanic [*t*]

# Example

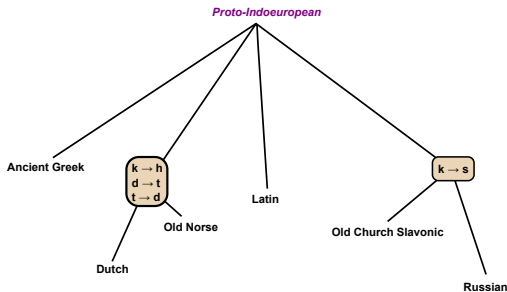
- reconstruct proto-forms and directionality of changes
  - PIE [ $*d$ ] → Germanic [t]
  - PIE [ $*t$ ] → Germanic [d/θ]

# Example

- reconstruct proto-forms and directionality of changes
  - PIE [*\*d*] → Germanic [t]
  - PIE [*\*t*] → Germanic [d/θ]
  - PIE [*\*k*] → Germanic [h], PIE [*\*k*] → Slavic [s]

# Example

- reconstruct proto-forms and directionality of changes
  - PIE [ $*d$ ]  $\rightarrow$  Germanic [ $t$ ]
  - PIE [ $*t$ ]  $\rightarrow$  Germanic [ $d/\theta$ ]
  - PIE [ $*k$ ]  $\rightarrow$  Germanic [ $h$ ], PIE [ $*k$ ]  $\rightarrow$  Slavic [ $s$ ]
- construct family tree based on *shared innovations*



# Example

## ● compile etymological dictionary

here: Köbler, Gerhard, Indogermanisches Wörterbuch, (5. Auflage) 2014, <http://www.koeblergerhard.de/idgwbhin.html>

**penk<sup>e</sup>\*\*\***, idg., Num. Kard.: nhd. fünf; ne. five; RB.: Pokorny 808 (1398/40), ind., iran., arm., gr., alb., ital., kelt., germ., balt., slaw., toch., heth.; Hw.: s. \*penk<sup>h</sup>tos, \*penk<sup>h</sup>ēkōnta, \*pñksti-? (?), \*penk<sup>h</sup>tos; W.: vgl. gr. πεντηκοστη (pentekoste), Num. Ord., fünfzigster Tag; ae. pent-e-costen, M., Pfingsten; W.: vgl. gr. πεντηκοστη (pentekoste), Num. Ord., fünfzigster Tag; afries. pink-ost-r-a 1 und häufiger?, pinxt-e-r-a, pinst-er, Sb. (Pl.), Pfingsten; W.: s. gr. πεντηκοστη (pentekoste), Num. Ord., fünfzigster Tag; as. pinkoston\* 2, sw. F. (n), Pfingsten; W.: vgl. gr. πεντηκοστη (pentekoste), Num. Ord., fünfzigster Tag; mlat. pentecoste; mnd. pinkestēn, pinxten: an. pik-is-dag-r, pikk-is-dag-r, st. M. (a), „Pfingsten“, Weißer Sonntag; W.: lat. quīnque, Num. Kard., fünf; W.: s. lat. quīntus, quīntus, Num. Ord., fünfte; W.: vgl. lat. Quīnticius, M.=PN, Quīnticius (Name einer römischen Gens); W.: germ. \*femf, \*femfe, Num. Kard., fünf; got. fimf 23, kringot. fynf\*, fyuf, Num. Kard., indekl., fünf (, Lehmann F55); W.: germ. \*femf, \*femfe, Num. Kard., fünf; an. fim, fim-m, Num. Kard., fünf; W.: germ. \*femf, \*femfe, Num. Kard., fünf; ae. fif, Num. Kard., fünf; W.: germ. \*femf, \*femfe, Num. Kard., fünf; afries. fif 14, Num. Kard., fünf; W.: germ. \*femf, \*femfe, Num. Kard., fünf; as. fif 17, Num. Kard., fünf; mnd. vif, Num. Kard.; W.: germ. \*femf, \*femfe, Num. Kard., fünf; ahd. fimf 90, Num. Kard., fünf; mhd. vūnf, Num. Kard., fünf; nhd. fünf, Num. Kard., fünf, DW 4, 548; W.: s. germ. \*femftō-, \*femftōn, \*femfta-, \*femftan, Num. Ord., fünfte; got. \*fimf-t-a, Num. Ord., fünfte; W.: s. germ. \*femftō-, \*femftōn, \*femfta-, \*femftan, Num. Ord., fünfte; an. fim-t-i, Num. Ord., fünfte; W.: s. germ. \*femftō-, \*femftōn, \*femfta-, \*femftan, Num. Ord., fünfte; ae. fif-t-a, Num. Ord., fünfte; W.: s. germ. \*femftō-, \*femftōn, \*femfta-, \*femftan, Num. Ord., fünfte; afries. fif-ta 14, Num. Ord., fünfte; W.: s. germ. \*femftō-, \*femftōn, \*femfta-, \*femftan, Num. Ord., fünfte; as. fif-to\* 1, Num. Ord., fünfte; mnd. vifte, vifte, Num. Ord.; W.: s. germ. \*femftō-, \*femftōn, \*femfta-, \*femftan, Num. Ord., fünfte; ahd. fimfto 14, Num. Ord., fünfte; mhd. vūnfte, Num. Ord., fünfte; nhd. fünfte, Num. Ord., fünfte, DW 4, 572; W.: vgl. germ. \*femti-, \*femtiz, st. F. (i), Fünzfahl; an. fim-t, st. F. (i), Fünzfahl; W.: vgl. germ. \*fengra-, \*fengraz, \*fingra-, \*fingraz, st. M. (a), Finger; got. figg-r-s\* 1, st. M. (a), Finger (, Lehmann F47); W.: vgl. germ. \*fengra-, \*fengraz, \*fingra-, \*fingraz, st. M. (a), Finger; an. fing-r, st. M. (a), später st. N. (a), Finger; W.: vgl. germ. \*fengra-, \*fengraz, \*fingra-, \*fingraz, st. M. (a), Finger; ae. fing-er, st. M. (a), Finger; W.: vgl. germ. \*fengra-

# Exercise

(from Crowley and Bower 2010, 80)

	Tongan	Samoaan	Rarotongan	Hawaiian	
1.	tapu	tapu	tapu	kapu	<i>forbidden</i>
2.	pito	pute	pito	piko	<i>navel</i>
3.	puhi	feula	puʔi	puhi	<i>blow</i>
4.	tafaʔaki	tafa	taʔa	kaha	<i>side</i>
5.	taʔe	tae	tae	kae	<i>feces</i>
6.	taŋata	taŋata	taŋata	kanaka	<i>man</i>
7.	tahi	tai	tai	kai	<i>sea</i>
8.	malohi	malosi	kaʔa	ʔaha	<i>strong</i>
9.	kalo	ʔalo	karo	ʔalo	<i>dodge</i>
10.	aka	aʔa	aka	aʔa	
11.	ʔahu	au	au	au	<i>gall</i>
12.	ʔulu	ulu	uru	poʔo	<i>head</i>
13.	ʔufi	ufi	uʔi	uhi	<i>yam</i>
14.	afi	afi	aʔi	ahi	<i>fire</i>
15.	faa	faa	ʔaa	haa	<i>four</i>
16.	feke	feʔe	ʔeke	heʔe	<i>octopus</i>
17.	ika	iʔa	ika	iʔa	<i>fish</i>
18.	ihu	isu	putaŋio	ihu	<i>nose</i>

	Tongan	Samoaan	Rarotongan	Hawaiian	
19.	hau	asu	ʔau	hau	<i>dew</i>
20.	tafuafi	siʔa	ʔika	iʔa	<i>firemaking</i>
21.	hiku	siʔu	ʔiku	hiʔu	<i>tail</i>
22.	hake	aʔe	ake	aʔe	<i>up</i>
23.	huu	ulu	uru	komo	<i>enter</i>
24.	maŋa	maŋa	maŋa	mana	<i>branch</i>
25.	maʔu	mau	mau	mau	<i>constant</i>
26.	maa	mala	mara	mala	<i>fermented</i>
27.	naʔa	faʔaŋa	maninia	naa	<i>quieten</i>
28.	nofo	nofo	noʔo	noho	<i>sit</i>
29.	ŋalu	ŋalu	ŋaru	nalu	<i>wave</i>
30.	ŋutu	ŋutu	ŋutu	nuku	<i>mouth</i>
31.	vaka	vaʔa	vaka	waʔa	<i>canoe</i>
32.	vaʔe	vae	vae	wae	<i>leg</i>
33.	laho	laso	raʔo	laho	<i>scrotum</i>
34.	lohu	lou	rou	lou	<i>fruit-picking pole</i>
35.	oŋo	loŋo	roŋo	lono	<i>hear</i>
36.	ua	lua	rua	lua	<i>two</i>

- ① Where do we find non-cognate words within the same row?
- ② Which regular sound correspondences do we observe?
- ③ How do you reconstruct the proto-sounds?
- ④ What family tree best explains the observed patterns?

- Clackson, J. (2007). *Indo-European Linguistics. An Introduction*. Cambridge University Press, Cambridge, UK.
- Crowley, T. and C. Bowerman (2010). *An introduction to historical linguistics*. Oxford University Press, Oxford.
- Ross, M. and M. Durie (1996). Introduction. In M. Durie and M. Ross, eds., *The Comparative Method Reviewed. Regularity and Irregularity in Language Change*, pp. 3–38. Oxford University Press, New York and Oxford.