

Linguistics 051

Proto-Indo-European Language and Society

Rolf Noyer

- The **comparative method** involves demonstrating that two languages are related through **systematic sound correspondences**.
- Forms which (on the basis of their meanings) are presumed to be descended from a common form in an ancestral language are collected into **correspondence sets**.

English	German	Spanish	French	Latin	Greek	Sanskrit	Russian
father	Vater	padre	père	pāter	patē'r	pitár-	(otéts)
foot	Fuss	pie	pied	ped-	pod-	pad-	(nogá)
for	für	para	pour	per	para	pra	pered
fish	Fisch	pescado	poisson	piscis	(ik <sup>h</sup> t <sup>h</sup> u-)	—	(rýba)
salt	Salz	sal	sel	sal-	hal-s	sal-ilā	sol <sup>j</sup>
six	sechs	seis	six	sex	hex	ṣáṣ	šest <sup>j</sup>
seven	sieben	siete	sept	septem	heptá	saptá	sem <sup>j</sup>

The above data can be used to establish the following correspondences (at least for initial segments of the forms):

NE	NHG	Sp	Fr	L	Gk	Skt	R
f	f	p	p	p	p	p	p
s	z	s	s	s	h	s/ş	s/š

Note that NHG initial *S-* is pronounced /z/.

Although we often represent linguistic expressions using **conventionalized spelling (orthography)**, this is a convenience and can be misleading, because **what is really being compared** is the sounds and not the letters used to represent them.

To circumvent this problem and also to be able to represent pronunciations in languages with no standard orthography, phoneticians have developed standard uses of letters and symbols to **unambiguously** represent pronunciations.

This is known as **phonetic** (or **phonological**) **transcription** and employs a special set of symbols.

There are several variant transcriptional types in use. One common one is that of the **International Phonetics Association** or **IPA**.

Indo-Europeanists often do not use the IPA, preferring to use either:

- (a) standard orthographies of languages (this entails, though, that the reader knows how to interpret the spelling systems of many languages), *or*
- (b) transcriptional systems that predate the IPA and are still in use because they are 'traditional' in their own way.

Often, because the languages in question are known only from written records, the ways in which they conventionally written using Roman letters merely **substitutes** a unique Roman letter for each symbol of the original writing system.

Since this is not a direct representation of pronunciation (a **transcription**), it is called a **transliteration**.

To understand sound changes you need to have some awareness of how transliteration works and how it differs from transcription.

All of the following ancient languages are written in non-Roman alphabets but are normally transliterated into a modified form of the Roman alphabet (but **not in IPA**). In most cases the descendants of these languages continue to use a non-Roman alphabet either identical to the one used in the ancient language or a variant of it.

Sanskrit (written in **devanagari**)

Avestan (written in the **Avestan** alphabet)

Hittite and other Anatolian languages, and Old Persian (written using various types of **cuneiform** writing)

Old Church Slavonic (written in **cyrillic** characters)

Armenian (written in the **Armenian alphabet**)

Runic, a dialect of early Germanic (written in **runes**)

Old Italic (Oscan, Umbrian, Faliscan etc.) were written in an early variant of the Roman alphabet

Greek (written in the **Greek alphabet**)

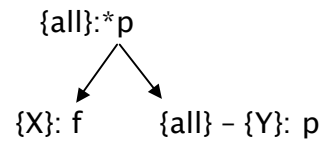
If [f] in NE and NHG corresponds systematically to [p] in cognates in the other languages, then we can consider three scenarios:

- |     | proto-form  | NE, NHG | the rest |
|-----|---|---------|----------|
| I   | *p  | f       | p        |
|     | (implies *p > f in NE, NHG, so they are the <b>innovating</b> lgg. /p/ then is an <b>archaism</b> while /f/ is an <b>innovation</b> )                         |         |          |
| II  | *f  | f       | p        |
|     | (implies *f > p in the other languages, so they are innovating. NE, NHG are <b>archaic</b> (wrt this property, at least), and the other lgg. are innovating.) |         |          |
| III | *# (something else)   | f       | p        |
|     | (implies that both /f/ and /p/ are innovations, derived from some third and different proto-form)   |         |          |

**Simplicity** disfavors scenario III since it requires postulating two changes instead of one.

Scenarios I and II are equal with respect to simplicity, provided certain assumptions are made.

Scenario I:



If NE and NHG both descend from some **common intermediate** language 'X' then we need to postulate only one change: \*p > \*f (in X). The remaining languages are archaic and don't have any change.



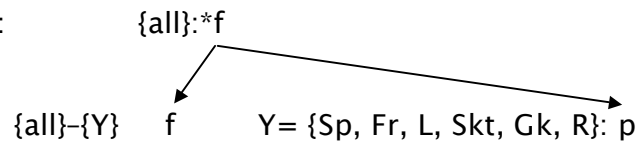
- Could NE and NHG have **individually** changed \*p > f without having descended from a common ancestor?

In principle, **yes**, although this would be a more complicated hypothesis. But such cases of **accidentally parallel development** are not impossible, and when they go unrecognized they can create serious problems in reconstruction.

Thus **identical traits** do not guarantee a common ancestor.

This is well-known in biology: the complex image-forming **eye** developed independently in a number of clades (e.g. the octopus eye developed independently of the vertebrate eye although they currently have very similar appearances).

Scenario II:



Scenario II is equally simple, **provided that** all the language we have examined **besides** NE and NHG descend from an ancestor they share but NE and NHG do not have.

Reconstruction entails hypotheses about **subgroupings** based on **shared archaisms** and **shared innovations**.

Ideally, a complete reconstruction of a proto-language will involve sound changes that are consistent with a branching 'tree', where each branching is defined by a set of sound changes.

In practice, though, language contact and borrowing often obscure this 'regular' tree to some extent.

Systematic sound correspondences hold between **cognate** words (or parts of words).

Two forms in different languages A and B are **cognate** if they 'descend' from the same form in a language ancestral to A and B.

- What does this really mean?

In applying the comparative method we must be careful in choosing the things to be compared — the **comparanda** (singular = comparandum). The comparanda need to be cognate, but cognates don't always have exactly the same meaning.

The reason for this is because meanings change over time: **semantic drift**.

Old English	NE	OE	NE
cniht 'boy' >	knight	mōdig 'courageous' >	moody
gār 'spear' >	gore	rēc 'smoke' >	reek
brond 'fire' >	brand	sacu 'conflict, strife' >	sake
burh 'stronghold' >	borough	scacan 'to hasten' >	shake
cēap 'purchase' >	cheap	swāt 'blood' >	sweat
drēam 'joy, revelry' >	dream	winnan 'to strive' >	win
gīlp 'boast' >	yelp		
lār 'advice' >	lore		
mǣl 'time, occasion' >	meal		

In order to determine what forms are comparanda one must study the history of meanings.

Linguists are typically interested in finding the oldest attested meanings available, since these will be:

- a. more reliable as indicators of what forms can be cognate
- b. better approximations of the meaning of the proto-form

- Knowledge of language consists of two parts:
  - a. a **rule system** expressing systematic facts and relationships. This is also known as the **grammar**.
  - b. a list of **arbitrary information that must be memorized**, commonly called the **lexicon**.
  
- Change in language can consist of change in the grammar or change in the lexicon.

Parts of the grammar of a human language:

- a. **Phonology** — permitted sounds and their distributions; ways in which sounds change depending on their environment of occurrence
- b. **Morphology** — ways in which elements in the **lexicon** (**morphemes**: roots and affixes) can combine to produce words
- c. **Syntax** — ways in which words can combine to form sentences.
- d. **Semantics** — ways in which the meaning of a sentence is derived from the meanings of its parts (morphemes and words)
- e. **Pragmatics** — how expressions of language are **used** to achieve various communicative functions

- Why do languages change?
- How does **your** language differ from that of your parents?
- Why would **you** change your way of speaking?



- What happens when speakers of different language come into contact with each other?
  - what kind of contact?
  - bilingualism
  - pidgin, trade language, creole, lingua franca
  - borrowing
  
- What aspects of language are typically borrowed?

## Proto-Indo-European Society

- domesticated cattle, pigs, goats and sheep and so had an advanced notion of **property**, particularly **movable property** (\*peku)
- hierarchical and patriarchal
- patrilocal extended family
- kings (L rēx), priests and poets distinct from common people
- king ruled over a 'people' (\*teutā)
- warrior band (Männerbund or L comitātus) under a chief

self-designation	<b>*h<sub>2</sub>ar(i)-ós</b> Skt ā́rya ARYAN, EIRE Olr aire 'freeman' cf. H āra 'what is fitting', Gk áristos 'best' ARISTOCRACY
(free) people	<b>*h<sub>1</sub>leud<sup>h</sup>-os</b> OE lēod , NHG Leute, OCS ljudīje 'people' < *h <sub>1</sub> leud <sup>h</sup> - 'to grow'
servants, slaves	<b>*h<sub>2</sub>antb<sup>hi</sup>-k<sup>w</sup>ol-os</b> L anculus, Gk amp <sup>h</sup> ípolos Skt abhi-cara-ḥ 'servant'  <b>*h<sub>2</sub>upo-sth<sub>2</sub>-i/o-s</b> Skt úpasti-ḥ 'servant'

1. \*h<sub>3</sub>rēǵ-s      L rēx, Skt rāj-aḥ 'king'  
    < \*h<sub>3</sub>reǵ- 'stretch out the arm, direct'  
    hence 'king' < 'director'  
    kings could have priestly functions  
    L rēx sacrōrum
  
2. \*u<sub>2</sub>nákt-s      Gk ánaks 'leader, lord', Toch A nātāk
  
3. \*u<sub>2</sub>iĵ-pot-s    Skt víś-páti-s 'master of the clan'  
    clan: \*u<sub>2</sub>eiĵ-/uoiĵ-o-s    ECONOMY
  
4. \*tag-ós      Gk tāg-ós 'leader', Toch A tāsíi 'leaders'  
    < \*tag- 'put in order, arrange'    TACTICS

## Proto-Indo-European Sound System

1	2	3	4	5		
p	t	k̑	k	k <sup>w</sup>	voicless stop	1 labial
b	d	ǵ	g	g <sup>w</sup>	voiced stop	2 dental
b <sup>h</sup>	d <sup>h</sup>	ǵ <sup>h</sup>	g <sup>h</sup>	g <sup>hw</sup>	voiced aspirated stop	3 palatal
	s	h <sub>1</sub>	h <sub>2</sub>	h <sub>3</sub>	fricative	4 velar
m	n				nasal	5 labialized
	l				lateral	velar
	r				rhotic	
6	7					
i	u	high vocoid				6 front
e	o	mid vowel				7 back
a		low vowel			{nasal, lateral, rhotic}	= sonorant

- The 'laryngeals' are a category of segments whose precise phonetic properties are quite uncertain
- The laryngeals were first reconstructed by Ferdinand de Saussure, who called them 'coefficientes sonantiques' and wrote them \*ə<sub>1</sub> \*ə<sub>2</sub> \*ə<sub>3</sub>
- In terms of their behavior in syllables and because of the possible 'gaps' in the segment inventory, they seem to fit into the system as fricatives.
- Moreover, when Hittite was deciphered in the 1930s some of the hypothesized laryngeals appeared in Hittite words, spelled with a symbol corresponding to a /x/ like sound in Semitic languages
- This has suggested to some a phonetic realization of /ç x x<sup>w</sup>/; others believe some may have been pharyngeals like /ħ ʕ/
- Some evidence suggests that \*h<sub>3</sub> was voiced

- The laryngeals have 4 principal effects
  - a. Coloring:  $h_2$  causes an adjacent /e/ to become /a/  
 $h_3$  causes an adjacent /e/ to become /o/
  - b. Lengthening:  $*VH > *V:$
  - c. Vocalization: If a laryngeal becomes *syllabic* it has a vowel as a reflex in many languages.  
 Famously, only *Greek* distinguishes the vocalized laryngeals:  $*h_1 > e$ ,  $*h_2 > a$ ,  $*h_3 > o$
  - d. Tone: A laryngeal after a vowel can affect the tone or 'accent' of words in Balto-Slavic languages

Syllables are subgroupings of segments; generally the groups that are formed can be predicted based on the properties the segments have.

The key concept for syllable structure is **sonority**, an approximate measure of the amount of energy required in the production of a type of segment.

In PIE the sonority scale is:

**most sonorous**

**least sonorous**

non-high vowel > high vocoid > sonorant > fricative > stop

The **Sonority Sequencing Principle** requires that syllables contain a **single sonority peak — the ‘nucleus’**, and that they **do not fall in sonority before the peak — the ‘onset’** and **do not rise in sonority after the peak — the ‘rhyme’**.

An unusual property of PIE syllables is that **all segments except stops** (and possibly /s/) **can be syllable sonority peaks**.



* r   sēn	'male'	> Gk ársēn, Skt ṛsabhá-
*ph <sub>2</sub>   tēr	'father'	> Gk patēr, L pater
*ǵe   nh <sub>1</sub>   tōr	'father, creator'	> Gk genétōr, L genitor, Skt janitár-
*ǵ h <sub>3</sub>   ūōs	'husband's sister'	> L glōs, Grk gálōs, Skt girí-

