Historical linguistics – lecture 2

the Comparative Method – the principal method that linguists have developed in order to establish genetic relatedness of languages and recover their histories.

Some lexical similarities (lexical correspondences) among Indo-European languages: Klein (1971)

*IE	English	Sanskrit	Greek	Latin	Old Irish
**māter*	mother	mātár-	māter-	māter	máthair
**g^{w}ōu*	cow	gāus	bous	bōs	bó
**k^{w}on*	hound	śvān-	kuōn	canis	con-
**genu*	knee	jānu	gonu	genū	glún
**neu(i)o*	new	návas	ne(w)os	novus	nue
**esti*	is	ásti	esti	est	is
**bher-*	bears	bhárati	pherei	ferit	berid
**duo*	two	dvā(u)	duo	duo	dā
**treies*	three	tráyas	treis	trēs	trí
**k^{w}etwor*	four	catvāras	tettares	quattuor	cethir

Some grammatical similarities among Indo-European languages
Singular o-stem declension in various IE languages (after Bammesberger (1992:49))

sg.	Sanskrit	Greek	Latin	Lith.	Gothic	PrGerm	IE
nom	vŕkas	lúkos	lupus	viľkas	wulfs	*wulfaz	*wĺk^{w}os
voc	vŕka	lúke	lupe	vilkè	wulf	*wulfe	*wĺk^{w}e
acc	vŕkam	lúkon	lupum	viľką	wulf	*wulfa^n	*wĺk^{w}om
gen	vŕkasya	lúkoio	lupī		wulfis	*wulfas(a)	*wĺk^{w}osyo
abl	vŕkād		lupō(d)	viľko			*wĺk^{w}ōd
dat	vŕkāya	lúkō	lupō(i)	viľkui		*wulfai	*wĺk^{w}ōi
loc	vŕke			vilkè			*wĺk^{w}ei
inst	vŕkā, -eṇa			vilkù	wulfa	*wulfē	*wĺk^{w}ō/ē

It is surprising that such structural similarities passed so long unnoticed. No evidence of any serious observations of this kind in ancient times. Also medieval and early modern periods simply ignored them. The idea of a historical development was not embraced by scholars. Aristotle explains certain linguistic forms found in Homer (used 500 years before his times) as produced from the forms of contemporary Greek. For him language is unchangeable and whatever differences appear they are attributed to arbitrary poetic license.

The situation changes at the end of the 18th and the beginning of the 19th c. Two circumstances are important:

1) the idea of comparison had established itself in other disciplines e.g. comparative anatomy.

In 1787 Christian Jakob Kraus ‘each language may be compared with each other in terms of general features of structure’

In 1781 Johann Christoph Adelung laid down precise criteria for different degrees of language relationship: dialects, related languages, distinct languages.

2) the intensive study of Sanskrit and other languages (William Jones 1746-94, Friedrich von Schlegel 1772-1829, Franz Bopp 1791-1867). And the true founder of historical linguistics Jacob Grimm (1785-1863) *Deutsche Grammatik*

The comparative method relies on certain characteristics of language and language change in order to work

1) arbitrariness of the relationship between phonological form and meaning
2) regularity of sound change.

Most linguists believe that change in articulation begins as a geographically and/or socially limited but **regular, unconscious and purely phonetic process** which then spreads by several different mechanisms (dialect borrowing, both social and otherwise, rule formation during acquisition) until regularity over a greater area is achieved.

3) uniformitarianism

How then is the task of language comparison to be undertaken? What can be compared and how?

In principle, similarities between languages could be due to:

- chance: this happens sometimes (cf. Mbabaram *dog* ‘dog’, Persian *bad* ‘bad’); our knowledge of the earlier history of these languages helps. These are not due to genetic relationship, or borrowing, but pure chance. Anyway these are easily eliminated by the requirement that similarities are not limited to a few lexical items but recur in a large set of forms
- borrowing: to control for this, we stick to "core" vocabulary and look at grammatical systems too;

a)

English	*French*	*German*
calf		Kalb
veal	veau	
cow		Kuh
beef	boeuf	
swine		Schwein
pork	porc	

b)

English	*French*	*German*
to	a	zu
two	deux	zwei
eat	manger	essen
bite	mordre	beissen

It might look from (a) as if English was simultaneously related to two quite distinct languages. In comparative linguistics this is considered suspect and indicative of secondary contact. Thus we look at (b) and quickly observe that English and German are closer, while French generally differs from both. The German/English correspondences pervade the whole vocabulary (including basic ‘core’ vocabulary) while the English/French correspondences are restricted to certain limited spheres of vocabulary, which is typical of borrowing

- necessity: perhaps words are intrinsically connected to their meanings. Aside from onomatopoeia, this isn't the case (the arbitrary nature of the linguistic sign);

- language universals: certain structural properties may be instances of these, but not the relation between a word and its meaning (the arbitrary nature of the linguistic sign again)

Evidence for possible **genetic relationship** should consist of correspondences which pervade the vocabulary and include most of the basic vocabulary. Having reduced the likelihood of chance similarities and similarities due to linguistic contact, we can strengthen our case for genetic relationship by showing that the correspondences in our putative cognates recur in a systematic fashion.
Identifying genetically related languages is one aspect of the comparative methodology. The second major goal of this method is the **reconstruction of ancestral forms** of genetically related languages. This lies at the heart of the technique known as **comparative reconstruction**.
The thinking behind reconstruction is: (a) many similarities among languages are due to a common origin and (b) sound changes are regular. Therefore, it should be possible to reconstruct earlier forms by running sound change backwards. Recognition of the regularity of sound change has been one of the major contributions of the Neogrammarians.

Take the words for 'heart' in a couple of languages:

English	*heart*	[ha:t]
German	*Herz*	[herts]
Greek	*kard-iú*	[kard-iú]
French	*càur*	[kàr]
Latin	*cord-is*	[kord-is]
Russian	*serd-ce*	[s'erd-cy]
Hindi	*xrid-aj*	[hrid-aj]
reconstruction ??		*[h/k/s]Vr[d/t/c]

Reconstruction involves answering a couple of questions:

1. How do we explain -*is* in Latin and -*ce* in Russian?
   Latin case ending; Russian diminutive
2. How do we explain the absence of a stop word-finally in French?
   Consonant deletion (more generally, Lenition)
3. How would we explain the relation of /d/: /t/ between English, Latin, and Russian?
   Grimm's Law
4. How do we explain the /c/ (/ts/) in German?
   OHG consonant shift
5. How to explain the relation of /k/ : /h/ between Latin and Germanic?
   Grimm's Law
6. How do we explain the relation of /k/ in Latin and the /s/ in Russian?
   centum/satem

7. How do we explain the different positions of the vowel and the liquid in Hindi? m*etathesis*

Hypothesized proto-form could be: *kord-/kerd-/krd- "heart"

The above relies on two related methodological principles: “majority rule” (the commonest reflex represents the original sound) and phonetic relatedness k/h/s, t/d/Ø. However, these principles can be supplemented by other considerations:

a. *Naturalness of sound changes: k > s* is a frequent change and *s > k* is very rare, and so we posit a velar in the word for “heart”, supposing, therefore, that Russian has diverged from the other languages in undergoing *k > s*.

b. *Relative age of daughter languages*: Latin is much older than English or German, and so – other things being equal – should be closer to the parent language. So older forms might “weigh more” than those of other languages in working out the reconstructions. (However, other things are not always equal – cf. point (a); also it’s known for example that Greek conserves the Indo-European vowel system better than Sanskrit does).

c. *Typological “fit”*: the reconstructed system should look like a normal language. It is debatable whether Indo-European, with its 20 reconstructed stop consonants and single fricative, really does.

Applying this kind of technique, we can develop a reasonably detailed idea of many features of the protolanguage. We also see shared innovations, which define subgroups - for example, all the Germanic languages have an initial voiceless fricative in the word for *heart*; what defines this subgroup is known as **Grimm’s Law**.

Grimm's Law describes some very general correspondences all across the consonant system, viz.:

	labial	***alveolar***	***velar***	***labio-velar***
VOICELESS:				
Indo-European	p (Lat. *piscis*)	t (Lat *tenuis*)	k (Lat *centum*)	k^w (Lat *quod*)
Germanic	f (*fish*)	θ (*thin*)	x (*hundred*) (In PGmc this word began with fricative /x/)	x^w (OE *hwæt*) (In PGmc this word began with fricative /x^w/)
VOICED:				
Indo-European	b (Lith. *dubùs*)	d (Lat *decem*)	g (Lat *genus*)	g^w (Gk. *gun*)
Germanic	p (*deep*)	t (*ten*)	k (*kind*)	k^w (OHG *quena*)
VOICED ASPIRATED:				
Indo-European	bh (Skt. *bhárāmi*)	dh (Skt. *mádhu*	gh (Skt. *stighnóti*)	g^wh (Skt. *gharmás*)

		"honey")		
Germanic	b (*bear*)	d (*mead*)	g (*steigen*)	g/w (*warm*)

These correspondences can be stated like this:

- An IE voiceless unaspirated stop (p, t, k, k^w) corresponds to a Germanic voiceless fricative (f, θ, x, x^w) at the same place of articulation.

- An IE voiced unaspirated stop (b,d,g,g^w) corresponds to a Germanic voiceless stop (p, t, k, k^w) at the same place of articulation.
- An IE voiced aspirated stop (bh,dh,gh,gh^w) corresponds to a Germanic voiced unaspirated stop (b,d,g,g^w) at the same place of articulation.

Laws such as Grimm's Law are essential to comparative reconstruction. They encapsulate regular sound changes. Observed similarities lead to the postulation of **protolanguages** - reconstructed, hypothesised ancestors of existing languages and groups of languages.

There are thousands of correspondences of the kind illustrated above, and so we postulate **(Proto-) Indo-European** (or IE) as the hypothetical ancestor language of all those above (and many more).

- **Indo-European**

Indo-European is a remarkable discovery. It is a language spoken in the very distant past in an uncertain location by an unknown people, and we have no written records of it. Our knowledge of it comes entirely from the technique of **comparative reconstruction.**

IE was spoken around 3500-3000BC, and began breaking up into its subgroups - presumably through processes of migration and fragmentation of the original 'Indo-European' people - between 3000 and 2000BC.

There's some debate concerning the Indo-Europeans:
a) The Indo-Europeans were the bearers of the Kurgan culture. Kurgans are prehistoric burial mounds found in the Southern Russian steppes. Gimbutas (1970) argued that the Indo-Europeans spread from this area by conquest (west into Europe, south to Persia and India). More specifically, the Yamna/Yamnaya culture occupied the Pontic/Caspian steppes around 3500BC, moving to the Danube basin around 3000BC. The Yamna probably rode horses, and are linked clearly to the Andronovo culture, which appeared around 2200BC in Northern Kazakhstan, and is known to be Indo-Iranian. The precursors of the Yamna is Sredny Stog culture, north of Black Sea around the Dneiper, ca 4500-3500BC.
b) Renfrew (1987) argues that the Indo-European homeland was Eastern Turkey, and that the language spread very gradually into Europe along with the introduction of agriculture. Cf. also Gamkrelidze & Ivanov (1995), who argued for the neighbouring part of the Caucasus as the homeland.

Comparative reconstruction is probably the most successful methodology ever developed in the history of linguistics. It depends for its success on the assumption that sound changes are exceptionless. On this basis, the huge edifice of knowledge about the reconstructed states of IE, its subfamilies and many other languages can be deduced.