Historical Linguistics—the study of language change

All languages change (a fact of life, not something that grammarians like Swift or Safir can stop)

Articulation—[intrəstin]

Analogy—sing/sang, wring/wrang??, dive/dived??, stride/strided??

Change takes place at every level—Phonology, Morphology, Syntax, Semantics

Modern linguistics began with the hypothesis that languages change

Sir William Jones—1786 address to Royal Asiatic Society

Discovery of systematic relations between different languages made linguistics a science

First time anyone proposed ‘rules’ for language

How would you visualize the effect of language change?

August Schleicher proposed family tree model of language change in 1871

Johannes Schmidt proposed wave model of language change in 1872

Influenced by Darwin and contemporary theories of biological evolution

Difference between language change and biological evolution

Still need to develop better ways to display language change

Four explanations for similarities across languages:

1. genetic relationship (a historical relationship)
2. borrowing (substratum/superstratum influences)
3. universal tendencies
4. chance

Genetic relationships differ from the other causes due to presence of regular correspondences

A change is regular if the change spreads throughout the vocabulary

Best if regular correspondences occur in phonology, morphology and syntax

Linguists use comparative method to establish genetic relationships

Compare words of similar form and meaning across languages

Comparative method rests on two assumptions:

1. Changes are regular—can be gradual or abrupt, but the end result is the same
2. Assume an arbitrary relation between form and meaning—Why?

1. Compile cognate sets, eliminate any borrowings
cognates are words that are genetically related

<table>
<thead>
<tr>
<th>gloss</th>
<th>Spanish</th>
<th>Sardinian</th>
<th>French</th>
<th>Portuguese</th>
<th>Rumanian</th>
</tr>
</thead>
<tbody>
<tr>
<td>embankment</td>
<td>[riβa]</td>
<td>[ripa]</td>
<td>[ʁiv]</td>
<td>[riba]</td>
<td>[ripə]</td>
</tr>
</tbody>
</table>

Throw out any “oddballs”, i.e. borrowings (internal or external)
2. Determine the sound correspondences
   1. r r θ r r
   2. i i i i i
   3. B p v b p
   4. a a θ a θ

3. Reconstruct a sound for each position
   a. Total correspondence *[_i_ _]
   b. Most natural development—use knowledge of common language changes
      voiceless -> voiced / between vowels
      stop -> fricative / between vowels *[_i_p_]
      unstressed vowel -> θ *[_i_p_a]
   c. Majority rules—Why? *[ripa]

Sardinian is the most conservative language since it preserves more of the original sounds
Where is Sardinian spoken in relation to the other languages? Why would it be the most conservative?

4. Check for regularity of sound changes in other cognate sets
   strawberry [siza] [sesa] [siza] *[sïsa] ??
   but pitchfork [sizu] [sïsu] [siza] therefore must reconstruct *[sesa] ‘strawberry’

5. Use changes to reconstruct proto-forms (mark with asterisk) proto-language
   e.g., Proto-Indo-European *bhrâter ‘brother’ *bher ‘carry, bear’

6. Reconstruct family tree:

   PIE *bher- ‘carry, bear’
   
   Germanic Sanskrit Greek Latin
   bh>b (Grimm’s Law) e>a bh>ph bh>f
   | | | |
   bear bhar- pher- fer-

Group languages into families by the number of changes the language share
Indic—short e>a
Germanic—Grimm’s Law—innovations common to Germanic language family
   1. separate from other Indo-European languages
   2. Germanic languages remained united before breaking up—Why?
Jacob Grimm—1st to point to systematic character of phonological change
collected fairy tales with bother to compile language cognates
Grimm’s Law

<table>
<thead>
<tr>
<th>Original</th>
<th>Indo-European</th>
<th>Germanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>b-&gt;p</td>
<td>lūbricus</td>
<td>slippery</td>
</tr>
<tr>
<td>d&gt;t</td>
<td>decem</td>
<td>ten</td>
</tr>
<tr>
<td>g&gt;k</td>
<td>iugum</td>
<td>yoke</td>
</tr>
<tr>
<td>p&gt;f</td>
<td>pater</td>
<td>father</td>
</tr>
<tr>
<td>t&gt;θ</td>
<td>trēs</td>
<td>three</td>
</tr>
<tr>
<td>k&gt;x (&gt;h)</td>
<td>cornū</td>
<td>horn</td>
</tr>
<tr>
<td>bh&gt;b</td>
<td>bhrātar</td>
<td>brother</td>
</tr>
<tr>
<td>dh&gt;d</td>
<td>bandh</td>
<td>bind</td>
</tr>
<tr>
<td>gh&gt;g</td>
<td>hostis (‘enemy’)</td>
<td>guest</td>
</tr>
</tbody>
</table>

Borrowing occurs when languages come into contact
Borrowing can also occur at different levels; obscures genetic relationship
Commonly find borrowed words
- Japanese -> English: hibachi, karate, sushi
- English -> Japanese: beer, computer

borrowed words often come from a common semantic domain
- French -> English: veal, venison, mutton, beef

Various features can spread to different languages
- apico -> uvular r in Europe /b/
- NW U.S. Makah, Quileute, Salish
- Wakashan, Chemakuan, Salishan
  - m->b, n->d, ñ->g
surprising because nasals are near universals across the world’s languages
  - Just an areal feature in the NW

Areal features indicate languages were in contact over long periods—sprachbund
  - e.g., India, Mesoamerica & NW U.S.
Although areal features are widespread, they do not produce systematic correspondences

Universal tendencies limit the arbitrariness of the sound meaning correspondence
- Some sounds appear in most languages: /p, t, k, m, n i, u, a/
- Onomatopoeia is fairly common
  - words for frog and frog sounds have velar stop and /r/, e.g., croak
  - words for mother have a nasal and /a/, e.g., Navajo má; K’iche’ nan

Chance can also produce apparent cognates—chance cognates are few and scattered across various semantic domains—not systematic!
<table>
<thead>
<tr>
<th>gloss</th>
<th>Algonkian</th>
<th>Scots Gaelic</th>
</tr>
</thead>
<tbody>
<tr>
<td>woman</td>
<td>bhanem</td>
<td>ban</td>
</tr>
<tr>
<td>person</td>
<td>alnoba</td>
<td>allaban-‘immigrant’</td>
</tr>
<tr>
<td>netting</td>
<td>lhab</td>
<td>lion-obhair</td>
</tr>
<tr>
<td>town</td>
<td>odana</td>
<td>dun</td>
</tr>
<tr>
<td>everywhere</td>
<td>ha?lwiwi</td>
<td>na h-uile</td>
</tr>
</tbody>
</table>

Random processes don’t produce systematic correspondences.
Requiring cognate sets to have similar meanings reduces the chance of false cognates.
But it is also possible to have semantic change—Latin hostis ‘enemy’ -> guest.
How similar must meanings be to be cognate? Lummi [mæn] ‘father’

There are limits to linguistic reconstruction—10,000 years beyond that time, historical changes obscure systematic correspondences.

Hasn’t stopped some linguists from trying to group languages further.
Nostratic *majrV ‘young male’ > mlarrā ‘marry a man’ (Altaic), mer-lo ‘young man’ IE
originated in Soviet Union in 1963 by Illich-Svitych & Dolgopolovsky
spoken around 12,000 BCE; reconstructed vocabulary of ~500 words.

Pinker (Words and Rules, p. 212)

```
?  Nostratic (15,000 years?)  Sino-Tibetan  New Guinea
Afro-Asiatic  Dravidian  Eurasiatic (10,000 years?)
Indo-European (5000-3500 B.C.)  Altaic  Uralic
  Romance  Germanic  Korean  Japanese  Hungarian
```

Joseph Greenburg—method of mass comparison (not accepted by most linguists)
Languages in the Americas derived from three protolanguages:
  Eskimo-Aleut (Nostratic?), Na-Dene & Amerind
Johanna Nichols—morphological distributions

Also some attempts to ‘time’ the regularity of language change
Morris Swadesh—glottochronology (also not accepted by most linguists)
Swadesh List: all, ashes, bark, belly, big, bird, bite, black, blood, bone, burn, cloud, cold
81-86% of common basic vocabulary remains after 1,000 years
Applied to Romance languages suggests initial divergence around 1,200 BCE
actually began diverging around 2,200 BCE according to Latin texts.
What does **Proto-Indo-European** look like?

Calvert Watkins ([American Heritage Dictionary](https://www.americanheritagedictionary.com/))

<table>
<thead>
<tr>
<th>Phonology</th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>kʷ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>d</td>
<td>g</td>
<td>gʷ</td>
</tr>
<tr>
<td></td>
<td>bh</td>
<td>dh</td>
<td>gh</td>
<td>gʰ h</td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>r</td>
<td>l</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>w</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Morphology**  
PIE had ablaut (vowel alternation), e.g., write/wrote

**Syntax**  
highly inflected words, case marking, person, voice, tense

**Culture**  
*deiw-os ‘god’ from root deiw ‘to shine’ (Latin diēs ‘day’)  
dyeu-pōter ‘chief god’ (> Latin Jupiter)  
society of gods was patriarchal ~ Latin pater familias  
kred-dhə ‘to put heart’ (> Latin crēdō ‘I believe’) heart-put  
sengʰh- ‘to prophesy, sing, make incantations  
gʷere- ‘to praise aloud’ (> Latin grātia ‘grace’)  
sak- ‘sacred’

many of these concepts were absorbed into Christianity

words related to time, weather, seasons, and natural surroundings
-> infer what PIE homeland was like; different conceptions of time  
yēr- ‘year’ related to words denoting activity  
wet- ‘year as a measure of domestic animal growth  
at (> Latin annum ‘to go’ > annual ---year as a passage of time

*aus ‘to shine’ East (related directions to the sun)  
also ‘to dawn’ (Latin Aurora) a Greek religious concept

*welt ‘forest or uncultivated land’ > wild  
But no word for ‘sea’; therefore inland area

*bhāgo ‘beech’ used to pinpoint PIE, but ranges could change over millennia  
bherəg ‘birch’ as well as the word’s meaning  
abel ‘apple’  
ker ‘cherry’  
bher ‘beaver’  
mūs ‘mouse’  
lūs ‘louse’ rhymed since PIE!  
knid ‘nit’ (louse egg)
Gamkrelidze & Ivanov ‘Early history of Indo-European languages’ (Sci. Am., March 1990) put original IE homeland near the Caucasus mountains about 6,000 years ago

around 4,000 BCE invade Anatolia -> Hittite kingdom
cuneiform tablets from library at capital Hattusas ~ Ankara
also find tablets from two related languages: Luwian & Palaic
therefore Anatolian split from IE by at least 6,000 BCE, possibly earlier

around 6,000 BCE Greek-Armenian-Indo-Iranian split from IE
have evidence of Indo-Iranian and Greek-Armenian by 5,500 BCE

Tocharian also diverged early from IE; first recognized in texts from Chinese Turkestan
easy to decipher since they were written in a Brahmi script and were mainly translations from known Buddhist writings
May be the Gutians mentioned in Babylonian cuneiform inscriptions
~ 5,000 BCE ~ King Sargon
Tocharian is similar to Italo-Celtic; so the languages were together before splitting off

IE has vocabulary for agricultural technology—developed around 7,000 BCE
Landscape was mountainous—IE has words for high mountains, mountain lakes, rapid rivers
~ East Anatolia
mountain oak, birch, beech, ash, willow, yew, pine, leopard, lion, monkey, elephant

Also had words for wheeled transport: wheel (*rotā), axle, yoke (*yugo), horse (*ekwo), foal

The neat aspect about the comparative method is that it is possible to apply it to unwritten languages
proto-Siouan probably originated in northern part of Mississippi Valley
The Siouan family includes Crow, Mandan, Dakotah and Kanza among other languages.

proto-Siouan had word for ‘gourd’ (correlates with archeological evidence ~ 3,000 BCE)
aquired words for ‘squash’ and ‘pumpkin’ after initial break up
borrowed from Algonkian in northern Siouan
tq in southern Siouan

The introduction of corn produced a whole set of words associated with processing corn
Siouan had already split up
get compound for corn  kō+ ‘grass’ = ‘gourd-grass’
‘gourd’
~ 1,800 BCE find a few grains; cultivated a few plants
1,300-900 BCE cultivated intensely (after Mississippi Valley split up)
find impact on degree of sedentism (~ 1,000 BCE according to bone samples)
happa ‘ear of corn’ is similar across Mississippi Valley languages
also refers to cattail seed pod, pigweed, goosefoot, amaranth
originally had word for the seed pod of edible plants
derived from ha-apa ‘grows a skin/covering
outside covering-grow
wa-ha-apa ‘thing that grows a covering’ descriptive compound common to Siouan
typical for introduced items, e.g. ‘horseless carriage’

Olmec civilization (~3,500 BCE) correlates with reconstruction of Mixe-Zoquean language family
MZ loan words:
PMZ *kakawa ‘cacao’—pan-Mayan kakaw  PMZ *tsima ‘gourd’—pan-Mayan tsima
PMi *tsiʔwa ‘squash’—Huastec tsiw  PMZ *koya ‘tomato’—Chol koya?
PMZ *sʔok ‘bean’—none  Zoque ʔotso ‘papaya’—Nahua očonih-tli

When did language originate?  What was it like?
See Prof. Gene Buckley’s page on the Evolution of Language.