- **Genetically related languages**
- **Comparative reconstruction**

Background reading:

- *CL* Ch 8, sec 1
- *CL* Ch 8, §7 (especially §7.1, §7.2)

1. Genetically related languages

- Why might two languages have words (morphemes) that are similar in sound and meaning?

English	**French**
escargot	*escargot*
chair	*chaise*
six [sɪks]	*six* [sis]

English	**Mbabaram**
dog [dɔg]	[dɔk]

English	**(many lgs)**
mama	[mama]

1. Genetically related languages

- Why might two languages have words (morphemes) that are similar in sound and meaning?

English	**French**	
escargot	*escargot*	borrowing
chair	*chaise*	(old) borrowing
six [sɪks]	*six* [sis]	common ancestor!

English	**Mbabaram**	
dog [dɔg]	[dɔk]	chance resemblance

English	**(many lgs)**	
mama	[mama]	sound symbolism

1. Genetically related languages

- When two languages are descended from a **common ancestor language**, linguists say that they are **genetically related**
  - This doesn't have anything to do with DNA or human biology — it's a metaphor
- For example, we can say that Spanish, French, and Italian are *genetically related* because they all descend from a common ancestor language
  - This does not mean that all current speakers of these languages are closely genetically related to each other in the biological sense!

1. Genetically related languages

- People love lists of words from different languages that resemble each other…
- But finding "similar" words is not proof of **genetic relationship** between languages
  - Borrowing? (from each other/both from a third lg)
  - Chance resemblance?
  - Sound symbolism?

→ Which English word is **genetically related** to the Greek word ***δεκα*** [ðeka] 'ten': ***decade***, or ***ten***?

1. Genetically related languages

- Etymologies (word origins) from the *Oxford English Dictionary*:

  decade < French *decade* …, < Latin *decas, decad-em,* < Greek *δεκάς, δεκάδα,* a group of ten, < *δέκα* ten

  ten < Old English *tíen, -e,* Anglian *tén, -e,* Common Germanic, = Old Low German **tehan,* … Old Saxon *tehan (tîan, tein)*, … < Old Germanic **teχan*, beside **teχun* [<] pre-Germanic **ˈdekm*

1. Genetically related languages

- What these eymologies are telling us
  - ***decade***: A Greek word (*δεκάδα*) was borrowed into Latin, then changed over time into French, and then was borrowed into English
  - ***ten***: A Pre-Germanic word (**ˈdekm*) changed into an Old Germanic word that changed into a word in descendant languages, including Old English

→ Which English word is **genetically related** to the Greek word ***δεκα*** [ðeka] 'ten': ***decade***, or ***ten***?

1. Genetically related languages

- What these eymologies are telling us
  - ***decade***: A Greek word (*δεκάδα*) was borrowed into Latin, then changed over time into French, and then was **borrowed** into English
  - ***ten***: A Pre-Germanic word (**ˈdekm*) changed into an Old Germanic word that changed into a word in descendant languages, including Old English
    | gradually changed over time

→ Which English word is **genetically related** to the Greek word ***δεκα*** [ðeka] 'ten': ***decade***, or ***ten***?

- The less-similar-looking one is related!

1. Genetically related languages

- Genetically related languages:
  "languages descended [through processes of language change] from a common parent" (*CL*, p 329)
  - I.e., language variation taken "to the extreme" — varieties that diverge until they **cease to be mutually intelligible**
- We've just seen: genetically related words/ morphemes may look *less* alike than borrowings
- So how do we **identify** genetic relationships?
  - Look for ***systematic* sound correspondences** that *recur* in many morphemes

2. Systematic sound correspondences

- Observe the following (forms given in orthography)

Sanskrit	*Greek*	*Latin*	*Gothic*	*English*
pad-	pod-	ped-	fōtus	foot
pra-	pro-	pro-	fra-	fro
nápāt- ‘descendant’		nepōs ‘nephew, grandson’	(OHG nefo)	ne[f]ew (OE nefa)
trī-/tráyas	treĩs/tría	trēs	þrija	[θ]ree
tv-am	tū (Doric)	tv-am	þu	thou (formerly [θ])
śatám	(he-)katón	[k]entum	hunda (pl.)	hundred
dáśa	déka	de[k]em	taíhun [tɛxun]	ten

- What are the **systematic** patterns?

2. Systematic sound correspondences

- Sound correspondences (part of Grimm's Law)

Sanskrit	*Greek*	*Latin*	*Gothic*	*English*
p	p	p	f	f
t	t	t	θ	θ
ʃ	k	k	h (x)	h

- Can we state any generalizations here?
  - Sanskrit/Greek/Latin ___ : Germanic ___

    (Sanskrit [ʃ] is the result of a separate sound change)

2. Systematic sound correspondences

- Sound correspondences (part of Grimm's Law)

Sanskrit	*Greek*	*Latin*	*Gothic*	*English*
p	p	p	f	f
t	t	t	θ	θ
ʃ	k	k	h (x)	h

- Can we state any generalizations here?
  - S/Gk/L voiceless stops : Gmc voiceless fricatives

    (Sanskrit [ʃ] is the result of a separate sound change)
- Be able to discuss a simple example like this:
  Which sounds/sound classes correspond?

2. Systematic sound correspondences

- Grimm's Law: Sound changes in Germanic

Proto-Indo-European	>	Germanic
[p]		[f]
[t]		[θ]
[k]		[x] (>[h])
[b]		[p]
[d]		[t]
[g]		[k]
[b^h]		[b]
[d^h]		[d]
[g^h]		[g]

- What English word is **genetically related** to Latin *card(ium)* [**k**ar**d**-]?

2. Systematic sound correspondences

- Grimm's Law: Sound changes in Germanic

Proto-Indo-European	>	Germanic
[p]		[f]
[t]		[θ]
[k]		[x] (>[h])
[b]		[p]
[d]		[t]
[g]		[k]
[b^h]		[b]
[d^h]		[d]
[g^h]		[g]

- What English word is **genetically related** to Latin *card(ium)* [**k**ar**d**-]? | *heart* [**h**ɑɹ**t**] (*cardio* is borrowed!)

3. The Indo-European language family

- Once we find **systematic sound correspondences**, they can be used to:
  - identify genetically related languages
  - develop hypotheses about the structure of the ancestor language

- One very successful example of this technique: The **Indo-European** language family
  - Here are some images of the Indo-European family tree: From SDSU | From Wikipedia

3. The Indo-European language family

FYI only…

- Most languages of **Europe** are Indo-European
  - Exception: Basque — no known relatives
  - Exception: Finnish, Estonian, Hungarian, Saami — belong to the Finno-Ugric family (which may also be related to the Turkic languages, but this is controversial)
- Some of the languages of **South Asia** are Indo-European: Hindi, Urdu, Bengali/Bangla, Gujarati, Marathi…
  - Others belong to the Dravidian family: Tamil, Malayalam, Telugu, Kannada…

4. Reconstructing a proto-language

- From a group of genetically related languages, we can use the systematic correspondences and our understanding of language change to **reconstruct** the ancestor language
- This technique is **comparative reconstruction**
  - It can be done with any aspect of linguistic structure: phonology, morphology, syntax, semantics, ...
- A reconstructed ancestor language is called a **proto-language**

4. Reconstructing a proto-language

- Example:
  The ancestor consonant of the [p]:[p]:[p]:[f]:[f] correspondence set shown above has been reconstructed as *[p] for Proto-Indo-European

- Warning! In historical linguistics *only*, * means RECONSTRUCTED/HYPOTHETICAL, *not* UNGRAMMATICAL

5. Comparative reconstruction: Examples

- Can we reconstruct these Middle Chinese forms?

gloss	Mandarin (Beijing)	Hakka (Huizhou)	Reconstructed proto-form
'spicy hot'	[la]	[lat]	
'basket'	[lan]	[lam]	
'lazy'	[lan]	[lan]	
'fear'	[pa]	[pa]	

5. Comparative reconstruction: Example

- Can we reconstruct these Middle Chinese forms?

gloss	Mandarin (Beijing)	Hakka (Huizhou)	Reconstructed proto-form
'spicy hot'	[la]	[lat]	*[lat]
'basket'	[lan]	[lam]	*[lam]
'lazy'	[lan]	[lan]	*[lan]
'fear'	[pa]	[pa]	*[pa]

- 'spicy hot' is *[lat] because if *[la], we can't explain why Hakka 'fear' is [pa] and not [pat]
- 'basket' is *[lam] because if *[lan], we can't explain why Hakka 'lazy' is [lan] and not [lam]

5. Comparative reconstruction: Example

- Can we reconstruct these Middle Chinese forms?

gloss	Mandarin (Beijing)	Hakka (Huizhou)	Reconstructed proto-form
'spicy hot'	[la]	[lat]	*[lat]
'basket'	[lan]	[lam]	*[lam]
'lazy'	[lan]	[lan]	*[lan]
'fear'	[pa]	[pa]	*[pa]

- To think about: What **sound-change rules** would we need, assuming our proposed reconstructions?
  - What sounds, if any, changed from the proto-language to Mandarin? From the proto-language to Hakka?

5. Comparative reconstruction: Examples

- Polynesian cognate sets (Crowley 1992) — a well-accepted case of related languages

Tongan	***Samoan***	***Rarotongan***	***Hawai'ian***	***Māori***	*gloss*
kafa	ʔafa	kaʔa	ʔaha	kaha	'strong'
ʔufi	ufi	uʔi	uhi	uhi	'yam'
afi	afi	aʔi	ahi	ahi	'fire'
faa	faa	ʔaa	haa	ɸaa	'four'
feke	feʔe	ʔeke	heʔe	ɸeke	'octopus'

- What are the sound correspondences?

5. Comparative reconstruction: Examples

- Polynesian sound correspondences

Tong	***Sam***	***Rar***	***Haw***	***Māori***	*consonant sound correspondences*
kafa	ʔafa	kaʔa	ʔaha	kaha	k:ʔ:k:ʔ:k f:f:ʔ:h:h
ʔufi	ufi	uʔi	uhi	uhi	ʔ:Ø:Ø:Ø:Ø f:f:ʔ:h:h
afi	afi	aʔi	ahi	ahi	Ø:Ø:Ø:Ø:Ø f:f:ʔ:h:h
faa	faa	ʔaa	haa	ɸaa	f:f:ʔ:h:ɸ
feke	feʔe	ʔeke	heʔe	ɸeke	f:f:ʔ:h:ɸ

- How to reconstruct ‘yam’ — *[ʔufi] or *[ufi]? Why?

6. What not to do

- The key to a plausible reconstruction is to ensure that the languages actually are genetically related
- This means that relationships between the (proposed) related forms in the languages are **systematic**
- It is important to **exclude borrowings** when doing comparative reconstruction (why?)
  - This can be very difficult to do, in practice

7. Historical linguistics: Implications

Techniques of historical linguistics allow us to:

- Identify **related languages**
  - This may in turn shed light on prehistoric population movements, etc.
- **Reconstruct** extinct languages
- Test hypotheses about **mental grammar**
  - Do our current models of mental grammar correctly predict the ways languages change?
- Test hypotheses about **language acquisition** and socially motivated **language variation**