Historical Linguistics

You know that our ancestors made good use of the sounds of iota and delta...but nowadays people change iota to eta or epsilon, and delta to zeta, thinking they have a grander sound...For instance, in the earliest times they called 'day' *himéra*, others said *heméra*, and now they say *hêméra*.

--Plato, *Cratylus*

One of the morals of this class:

- language is complicated.

One of the morals of this class:

- language is complicated.

    downward entailing contexts,
    binding theory,
    X-bar theory......

One of the morals of this class:

- language is complicated.

  downward entailing contexts,
  binding theory,
  X-bar theory......

how do we figure stuff like this out?

One of the morals of this class:

- language is complicated.

      downward entailing contexts,
      binding theory,
      X-bar theory......

how do we figure stuff like this out?
   -->claim: in many cases, it's <u>innate</u>.

Innateness Hypothesis:

we don't start with a blank slate, but rather with a rich body of linguistic knowledge.

Innateness Hypothesis:

we don't start with a blank slate, but rather with a rich body of linguistic knowledge.

as a result, we don't have to figure some things out...and for things that we do, we have help.

on the other hand...

...clearly, not everything is innate.

/kʰæt/

...and, in fact, kids do make mistakes about things like this.

...and, in fact, kids do make mistakes about things like this.

- what does 'livid' mean?

...and, in fact, kids do make mistakes about things like this.

- what does 'livid' mean?
  white? red? angry?

...and, in fact, kids do make mistakes about things like this.

- what does 'livid' mean?
  white? red? angry

- the verb 'misle': I used to believe in this verb...

...and, in fact, kids do make mistakes about things like this.

- what does 'livid' mean?
  white? red? angry

- the verb 'misle': I used to believe in this verb...but I'd been **misle**d.

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift
    OE *(ge)*bed 'prayer'
    (cf. German *beten* 'pray')
    > ModE *bead*
  ("I'm counting my *beads* on this rosary")

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift

  OE *(ge)bed* 'prayer' > ModE *bead*

  OE *steorfan* 'die'
  (cf. German *sterben*)
     > ModE *starve*

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift
    OE *(ge)bed* 'prayer' > ModE *bead*
    OE *steorfan* 'die' > ModE *starve*

  Fr. *nègre* 'black man' >
    Haitian Creole *nèg* 'man'

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift
  - OE *(ge)bed* 'prayer' > ModE *bead*
  - OE *steorfan* 'die' > ModE *starve*
  - Fr. *nègre* 'black man' >HC *nèg* 'man'

OE *cniht* 'boy, servant'
(German *Knecht*) > ModE *knight*

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift

OE *(ge)bed* 'prayer' > ModE *bead*
OE *steorfan* 'die' > ModE *starve*
Fr. *nègre* 'black man' >HC *nèg* 'man'
OE *cniht* 'servant' > ModE *knight*

OE *huswif* 'housewife'
   > ModE *hussy*

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift
    OE *(ge)bed* 'prayer' > ModE *bead*
    OE *steorfan* 'die' > ModE *starve*
    Fr. *nègre* 'black man' >HC *nèg* 'man'
    OE *cniht* 'servant' > ModE *knight*
    OE *huswif* 'housewife' > ModE *hussy*
    PAN **wada* 'there is'
        > Tagalog *wala* 'there isn't'

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift

  OE *(ge)bed* 'prayer' > ModE *bead*

  OE *steorfan* 'die' > ModE *starve*

  Fr. *nègre* 'black man' >HC *nèg* 'man'

  OE *cniht* 'servant' > ModE *knight*

  OE *huswif* 'housewife' > ModE *hussy*

  PAN **wada* 'there is'

reconstructed > Tagalog *wala* 'there isn't'

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift
- recuttings (like *misled > misle-d*)

ME a**n e**kename >
ModE a **ni**ckname

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift
- recuttings (like *misled > misle-d*)

ME a**n e**kename >
   ModE a **ni**ckname
ME pease (mass noun) >
   ModE pea-s (plural count noun)

Sometimes these 'mistakes' catch on,
in the form of:

- various kinds of semantic drift
- recuttings (like *misled > misle-d*)

OE nēah 'near'

OE nēarra 'nearer'

OE nēahsta 'nearest'

Sometimes these 'mistakes' catch on, in the form of:
- various kinds of semantic drift
- recuttings (like *misled > misle-d*)

OE nēah 'near'   >ModE nigh

OE nēarra 'nearer'   >ModE near

OE nēahsta 'nearest' >ModE next

in the form of:
- various kinds of semantic drift
- recuttings (like *misled > misle-d*)

OE nēah 'near'   >ModE nigh

OE nēarra 'nearer'   >ModE near

OE nēahsta 'nearest' >ModE next
  doesn't look much like a comparative...

in the form of:
- various kinds of semantic drift
- recuttings (like *misled > misle-d*)

OE nēah 'near'   >ModE nigh

OE nēarra 'nearer'   >ModE near

OE nēahsta 'nearest' >ModE next

reanalysis: **near**, **near-er**, **near-est**

Sometimes these 'mistakes' catch on, in the form of:

- various kinds of semantic drift
- recuttings (like *misled > misle-d*)
- sound changes!

(side note:
   "catch on" is a relative term...)

positive 'anymore'

I used to walk to work,
   but **anymore**, I usually take the T.

"A language is a dialect with an army and a navy"
    --Uriel Weinreich

- Danish, Norwegian, Swedish
- Portuguese, Spanish
- Mandarin, Cantonese
- Serbo-Croatian >
    Serbian, Croatian

if you've got two populations with different sets of changes spreading among them, you might eventually want to call them different languages...

some numbers:

	Skt.	Greek	Latin	Gthc.	O.Ir	Lith.	OCS*	Bsque	Tkish
1.	ékas	hei:s	u:nus	ains	oín	víenas	jedinŭ	bat	bir
2.	dvaú	dúo:	duo	twai	da	dù	dŭva	bi	iki
3.	tráyas	trei:s	tre:s	θreis	tri	try:s	trĭje	hiru	üç

*Old Church Slavonic

some numbers:

	Skt.	Greek	Latin	Gthc.	O.Ir	Lith.	OCS*	Bsque	Tkish
1	ékas	hei:s	u:nus	ains	oín	víenas	jedinŭ	bat	bir
2	dvaú	dúo:	duo	twai	da	dù	dŭva	bi	iki
3	tráyas	trei:s	tre:s	θreis	tri	try:s	trĭje	hiru	üç

cognates

in fact, we can be more systematic than this:

Grimm's Law (Rasmus Rask, Jakob Grimm)

	Latin	**Greek**	**English**
d-t	**d**uo	**d**úo	**t**wo
	e**d**-o	é**d**-o	ea**t**
	decem	**d**éka	**t**en
g-k	**g**enus	**g**onu	**k**in
	a**g**er	a**g**rós	a**c**re
b-p	la**b**ium	--	li**p**
	canna**b**is	kánna**b**is	hem**p**
	lu**b**ricus	--	sli**pp**ery

Grimm's Law (Rasmus Rask, Jakob Grimm)
[+stop, +voice] --> [-voice]

	Latin	**Greek**	**English**
d-t	**d**uo	**d**úo	**t**wo
	e**d**-o	é**d**-o	ea**t**
	decem	**d**éka	**t**en
g-k	**g**enus	**g**onu	**k**in
	a**g**er	a**g**rós	a**c**re
b-p	la**b**ium	--	li**p**
	canna**b**is	kánna**b**is	hem**p**
	lu**b**ricus	--	sli**pp**ery

Once we've figured out all the sound laws we need for a bunch of related languages, we can posit the 'underlying forms' that underwent the sound changes: <u>protolanguage</u>

Sanskrit	ad-
Latin	ed-
English	eat

....

Sanskrit	a**d**-
Latin	e**d**-
English	ea**t**

....

Sanskrit	a**d**-
Latin	e**d**-
English	ea**t**

Grimm's Law (Germanic):
   [+stop, +voice]-->[-voice]

Sanskrit	**a**d-
Latin	**e**d-
English	eat

Sanskrit	**a**d-
Latin	**e**d-
English	eat

Sanskrit	Latin	
ad-	**e**d-	'eat'
d**a**nta	d**e**nt-	'tooth'
avi-	**o**vi-	'sheep'
dv**a**-	du**o**	'two'
ajr**a**	**a**g**e**r	'field'

Proto-Indo-European: *ed-   'eat'

Sanskrit   (*e>a)   ad-
Latin   ed-
English   (G.L...)   eat

Proto-Indo-European: *ed-   'eat'

Sanskrit	(*e>a)	ad-
Latin		ed-
English	(G.L...)	eat

careful! The proto-form doesn't have to be the same as any daughter form.

w-->gw in Chamorro:

Tagalog	**Chamorro**
asa**w**a	asa**gw**a 'spouse'
dala**w**a	hu**gw**a 'two'
wala 'there isn't'	**gw**aha 'there is'

...

w-->gw in Chamorro, and...

Tagalog	**Chamorro**
asa**w**a	asa**gw**a 'spouse'
dala**w**a	hu**gw**a 'two'
wala 'there isn't'	**gw**aha 'there is'

PIE	**Welsh**
***w**ir	**gw**ir 'man'

Proto-Germ.	**Late Latin**
***w**erra	***gw**erra 'war'
***w**ard-	***gw**ard- 'guard'

Hawaiian	Maori	Tongan	Samoan	
kalo	taro	talo	talo	'taro'
piko	pito	pito	pito	'navel'
moko	moto	moto	moto	'punch'
aka	ata	ata	ata	'dawn'
kai	tai	tahi	tai	'sea'
nuku	ŋutu	ŋutu	ŋutu	'beak'

Hawaiian	Maori	Tongan	Samoan	
kalo	**t**aro	**t**alo	**t**alo	'taro'
pi**k**o	pi**t**o	pi**t**o	pi**t**o	'navel'
mo**k**o	mo**t**o	mo**t**o	mo**t**o	'punch'
a**k**a	a**t**a	a**t**a	a**t**a	'dawn'
kai	**t**ai	**t**ahi	**t**ai	'sea'
nu**k**u	ŋu**t**u	ŋu**t**u	ŋu**t**u	'beak'

Hawaiian	Maori	Tongan	Samoan	
kalo	**t**aro	**t**alo	**t**alo	'taro'
pi**k**o	pi**t**o	pi**t**o	pi**t**o	'navel'
mo**k**o	mo**t**o	mo**t**o	mo**t**o	'punch'
a**k**a	a**t**a	a**t**a	a**t**a	'dawn'
kai	**t**ai	**t**ahi	**t**ai	'sea'
nu**k**u	ŋu**t**u	ŋu**t**u	ŋu**t**u	'beak'

Hawaiian:

[+stop, +alveolar]-->[+velar]

Hawaiian	Maori	Tongan	Samoan	P-Pol
kalo	**t**aro	**t**alo	**t**alo	*talo 'taro'
pi**k**o	pi**t**o	pi**t**o	pi**t**o	*pito 'navel'
mo**k**o	mo**t**o	mo**t**o	mo**t**o	*moto 'punch'
a**k**a	a**t**a	a**t**a	a**t**a	*ata 'dawn'
kai	**t**ai	**t**ahi	**t**ai	*tahi 'sea'
nu**k**u	ŋu**t**u	ŋu**t**u	ŋu**t**u	*ŋutu 'beak'

Hawaiian:

[+stop, +alveolar]-->[+velar]   (t-->k)

Hawaiian	Maori	Tongan	Samoan	P-Pol	
kalo	taro	talo	talo	*talo	'taro'
piko	pito	pito	pito	*pito	'navel'
ʔele	kere	kele	ʔele		'black'
ʔula	kura	kula	ʔula		'red'
aʔe	ake	hake	aʔe		'up'
ʔapo	kapo	--	ʔapo		'grasp'

Hawaiian:

[+stop, +alveolar]-->[+velar]   (t-->k)

Hawaiian	Maori	Tongan	Samoan	P-Pol
kalo	taro	talo	talo	*talo 'taro'
piko	pito	pito	pito	*pito 'navel'
ʔele	**k**ere	**k**ele	ʔele	*kele 'black'
ʔula	**k**ura	**k**ula	ʔula	*kula 'red'
aʔe	a**k**e	ha**k**e	aʔe	*hake 'up'
ʔapo	**k**apo	--	ʔapo	*kapo 'grasp'

Hawaiian:

[+stop, +alveolar]-->[+velar]   (t-->k)

[+stop, +velar]-->[+glottal]   (k-->ʔ)

Hawaiian	Maori	Tongan	Samoan	P-Pol
kalo	taro	talo	talo	*talo 'taro'
piko	pito	pito	pito	*pito 'navel'
ʔele	**k**ere	**k**ele	**ʔ**ele	*kele 'black'
ʔula	**k**ura	**k**ula	**ʔ**ula	*kula 'red'
a**ʔ**e	a**k**e	ha**k**e	a**ʔ**e	*hake 'up'
ʔapo	**k**apo	--	**ʔ**apo	*hapo 'grasp'

Hawaiian:

[+stop, +velar]-->[+glottal]   (k-->ʔ)

[+stop, +alveolar]-->[+velar]   (t-->k)

Hawaiian	Maori	Tongan	Samoan	P-Pol	
kalo	taro	talo	talo	*talo	'taro'
ʔele	kere	kele	ʔele	*kele	'black'
aka	ata	ata	ata		'dawn'
ihu	ihu	ihu	isu		'nose'
ao	ao	ʔaho	ao		'day'
aloha	aroha	ʔaloʔofa	alofa		'love'
wae	wae	vaʔe	vae		'leg'
leo	reo	leʔo	leo		'voice'
hau	hau	hau	sau		'dew'
wai	wai	vai	vai		'water'

Hawaiian	Maori	Tongan	Samoan	P-Pol
kalo	taro	talo	talo	*talo 'taro'
ʔele	kere	kele	ʔele	*kele 'black'
aka	ata	ata	ata	*ata 'dawn'
ihu	ihu	ihu	isu	*isu 'nose'
ao	ao	ʔaho	ao	*ʔaho 'day'
aloha	aroha	ʔaloʔofa	alofa	*ʔaloʔofa 'love'
wae	wae	vaʔe	vae	*vaʔe 'leg'
leo	reo	leʔo	leo	*leʔo 'voice'
hau	hau	hau	sau	*sau 'dew'
wai	wai	vai	vai	*vai 'water'

Hawaiian	Maori	Tongan	Samoan	P-Pol
kalo	taro	talo	talo	*talo 'taro'
ʔele	kere	kele	ʔele	*kele 'black'
aka	ata	ata	ata	*ata 'dawn'
ao	ao	ʔaho	ao	*ʔaho 'day'

Hawaiian:

[+stop, +velar]-->[+glottal]   (k-->ʔ)

[+stop, +alveolar]-->[+velar]   (t-->k)

[+stop, +glottal]-->Ø   (ʔ-->Ø)

Hawaiian	Maori	Tongan	Samoan	P-Pol
kalo	taro	talo	talo	*talo 'taro'
ʔele	kere	kele	ʔele	*kele 'black'
aka	ata	ata	ata	*ata 'dawn'
ao	ao	ʔaho	ao	*ʔaho 'day'

Hawaiian:

[+stop, +glottal]-->Ø   (ʔ-->Ø)

[+stop, +velar]-->[+glottal]   (k-->ʔ)

[+stop, +alveolar]-->[+velar]   (t-->k)

big discovery:
   sound change is <u>regular</u>.

(Neogrammarian Hypothesis)

big discovery:
   sound change is <u>regular</u>.

--->shifts emphasis away from looking for lists of words that 'look similar'; now what we're looking for is lists of words that can be related by regular sound laws.

"looking similar" is not **<u>necessary</u>** to prove relationship:

	<u>A</u>	**<u>B</u>**	**<u>C</u>**
'two'	er	erku	duo

"looking similar" is not **<u>necessary</u>** to prove relationship:

	<u>Mandarin</u>	**<u>Armenian</u>**	**<u>Greek</u>**
'two'	er	erku	duo

"looking similar" is not **<u>necessary</u>** to prove relationship:

	<u>Mandarin</u>	**<u>Armenian</u>**	**<u>Greek</u>**
'two'	er	erku	duo
'fear'		erki-	dwi-
'long'		erkar	dwa:ron

PIE *dw* > Armenian *erk*

"looking similar" is not **<u>sufficient</u>** to prove relationship:

<u>Mbabaram</u> **<u>English</u>**

"looking similar" is not **<u>sufficient</u>** to prove relationship:

<u>Mbabaram</u>   **<u>English</u>**
                        dog

"looking similar" is not **<u>sufficient</u>** to prove relationship:

<u>Mbabaram</u>
dog

<u>English</u>
dog

"looking similar" is not **<u>sufficient</u>** to prove relationship:

<u>Mbabaram</u>	**<u>English</u>**
dog	dog
(<*gudaga:	(<OE docga
Yidiñ gudaga,	'mastiff'
Dyirbal guda)	

"looking similar" is not **sufficient** to prove relationship:

Mbabaram
dog

English
dog

Persian
bad

English
bad

Malay
mata 'eye'

Greek
mati 'eye'

"looking similar" is not **<u>sufficient</u>** to prove relationship:

<u>English</u>	**<u>Kaqchikel</u>**
mess	mes

"looking similar" is not **<u>sufficient</u>** to prove relationship:

<u>English</u>	**<u>Kaqchikel</u>**
mess	mes
man	ači
mouse	č'oy
moon	qatiʔt

two bad ideas:
- glottochronology
- megalocomparison

glottochronology (Swadesh, 1950s):
"carbon dating" of language splits.

take a list of 'basic vocabulary':

I	fish
you	kill
we	swim
this	hot
that	good
man	name
.....	

figure out how many cognates the two languages share on the list

(e.g., English and Danish share 59%, English and Albanian share 13%)

Assume that 'cognate loss' happens at a constant rate.

(14% every 1000 years)

Do some math.

$$t = \log C / 2 \log r$$

t=time depth in millennia
C=percentage of cognates
r=constant (.86)

problem:

- 'cognate loss' does not in fact happen at a constant rate.

(Icelandic's retention rate: 97%, English's retention rate: 68%)

problem:

- 'cognate loss' does not in fact happen at a constant rate.

  - language contact
  - taboos (Australia,
       PIE bears...)

another bad idea:
  megalocomparison (Greenberg,
                          Ruhlen)

megalocomparison (Greenberg)
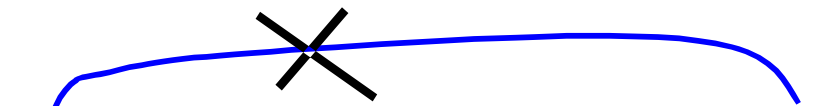## Proto-World *maliq'a* 'swallow'

Arabic	m-l-j 'suck a breast'	(Afro-Asiatic)
English	milk	(Indo-European)
Saami	mielga 'breast'	(Finno-Ugric)
Tamil	melku 'chew'	(Dravidian)
Yupik	melug- 'suck'	(Eskimo-Aleut)
Kutenai	u'mqolh 'swallow'	(Almosan)
Tfaltik	milq 'swallow'	(Penutian)
Akwa'ala	milqi 'neck'	(Hokan)
Cuna	murki 'swallow'	(Chibchan)
Quechua	malq'a 'throat'	(Andean)

This method is guaranteed to give you false positives:

English *have*   Latin *habere* 'have'
(German *haben*)

This method is guaranteed to give you false positives:

English *have* ✕ Latin *habere* 'have'

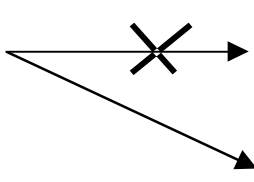|	
horn	**c**ornu
heart	**c**ord-em
hemp	**c**annabis
hundred	**c**entum

This method is guaranteed to give you false positives:

English *have* ⟶✗⟶ Latin *habere* 'have'

⟶ Latin *capire* 'take'

...and false negatives:

Armenian *erku*, English *two*

Hindi *pa:ntʃ*, English *five*