Comparative Linguistics

Languages are complex dynamic systems that change constantly. Similarities between languages provide the primary evidence of genetic relationships. Comparative linguistics attempts to identify the regular correspondences between languages that result from historical changes. Languages may have irregular correspondences due to chance or borrowing. Two principles to follow in establishing regular correspondences are:

1. Majority rules: Choose the sound that appears in the majority of languages.
2. Most natural change: Choose the sound that results in the most natural change:

Common sound changes
- Voiceless sounds become voiced between vowels and before voiced consonants
- Stops become fricatives between vowels
- Consonants become palatalized before front vowels
- Consonants become voiceless at the ends of words
- Difficult consonant clusters are simplified
- Difficult consonants are made easier
- Oral vowels become nasalized before nasals
- Fricatives change to [h]
- [h] deletes between vowels
- Clusters of vowels are broken up by consonants

Steps in reconstruction:
1. Determine the sound correspondences in the set of cognates
2. Reconstruct a sound for each position
3. Check for the regularity of sound change

A. Checking the regularity of sound change

1. Proto-Quechua

For each word specify the sound change(s) between Proto-Quechua and Tena, one of its daughter languages. Determine whether each sound change is conditioned or unconditioned, and formulate a rule for the conditioned sound changes.

Proto-Quechua	Tena	Gloss
tʃumpi	tʃumbi	'belt'
timpu	timbu	'boil'
nutku	nuktu	'brains'
akla	agla	'choose'
wakli	wagli	'damage'
utka	ukta	'fast'
kunka	kunga	'neck'
ljantu	ljandu	'shade'

mutki	mukti	'smell'
pukju	pugju	'spring'
inti	indi	'sun'
sanku	sangu	'thick'
hampatu	hambatu	'toad'

2. Proto-Slavic

Specify the changes between Proto-Slavic and Bulgarian. Classify each change and conditioned or unconditioned, and write a rule for each change. Finally, state the order in which these rules apply.

PROTO-SLAVIC	BULGARIAN	GLOSS
gladuka	glatkə	'smooth'
kratuka	kratkə	'short'
blizuka	bliskə	'near'
ʒeʒika	ʒeʃkə	'scorching'
lovuka	lofkə	'adroit'

3. Old Indic

Determine the sound changes that took place in the development of Marathi from Old Indic. Classify the changes as conditioned or unconditioned, and write a rule for each change.

OLD INDIC	MARATHI	GLOSS
aŋka	aŋka	'hook'
arka	akka	'sun'
bʰakti	bʰatti	'devotion'
catwa:ri	catta:ri	'four'
kalpa	kappa	'rule'
kardama	kaddama	'mud'
kaʈaka	kaɖa:	'bracelet'
kaʈaka	kaɖa:	'crow'
mudgara	muggara	'mallet'
pita:	pia:	'father'
rudra	rudda	'terrible'
sapatni:	savatti:	'co-wife'
supta	sutta	'asleep'
ʃabda	sadda	'sound'
ʃata	sa:	'hundred'

utkaɳʈaː	ukkaɳʈaː	'desire'
vikrama	vikkama	'strength'
viʈapa	viʈava	'branch'

<bʰ> represents a murmured (breathy voiced) bilabial stop: <ʈ, ɖ, ɳ> represent retroflex stops.

3. Latin (Langacker 1972)

CLASSICAL	VULGAR	
ásinus	ásnos	'ass'
kálidum	káldo	'hot' (accusative)
kéera	kéra	'wax'
kínerem	kénre	'ash' (accusative)
díikere	díkre	'say'
fúrnus	fórnos	'oven'
kolóorem	kolóre	'color' (accusative)
kírkulus	kérklos	'circle'
dúukere	dúkre	'lead'
pílus	pélos	'hair'
stáare	estáre	'stand'
stríktus	estréktos	'close'
spíina	espína	'thorn'
skúutum	eskúto	'shield'
plakéere	plakére	'please'

B. Syntactic Change

Old English (ca. 1100)

Fæder ure þu þe eart on heofonum, si þin nama gehalgod. Tobecume þin rice. Gewurþe þin willa on eorðan swa swa on heofonum. Urne gedœghwamlican hlaf syle us to dæg. And forgyf us ure gyltas, swa swa we forgyfaδ urum gyltedum. And ne gelæd þu us on costnungen ac alys us of yfele. Soδlice.

Middle English (ca. 1400)

Oure fadir that art in heuenes halowid be thi name, thi kyngdom come to, be thi wille don in erthe es in heuene, yeue to us this day oure bread ouir other substance, & foryeue to us oure dettis, as we forgeuen to oure dettouris, & lede us not in to temptacion: but delyuer us from yuel, amen.

Early Modern English (ca. 1611)

Our father which art in heaven, hallowed be thy Name. Thy kingdome come. Thy will be done, in earth, as it is in heaven. Giue vs this day our dayly bread. And forgiue vs our debts, as we forgiue our debters. And leade vs not into temptation, but deliuer vs from euill: For thine is the kingdome, and the power, and the glory, for euer, Amen.

2. Yaqui (Langacker 1972)

Seventeenth Century Yaqui

hu oʔoo hume haamuchim benasya bʷeʔu
this man these women like big
'This man is as big as these women.'

hu oʔoo cheʔa hume haamuchim beppa bʷeʔu
this man more these women over big
'This man is bigger than these women.'

hu oʔoo hume haamuchim beppa bʷeʔu
this man these women over big
'This man is bigger than these women.'

Modern Yaqui

hu oʔoo bʷeʔu ke hume haamuchim benasya
this man big than these women like
'This man is as big as these women.'

hu oʔoo cheʔa bʷeʔu ke hume haamuchim beppa
this man more big than these women over
'This man is bigger than these women.'

hu oʔoo bʷeʔu ke hume haamuchim beppa
this man big than these women over
'This man is bigger than these women.'

hu oʔoo cheʔa bʷeʔu ke hume haamuchim
this man more big than these women
'This man is bigger than these women.'

C. Reconstruction

For each of these reconstruction problems:
  a. Set up the sound correspondences for each cognate set, and reconstruct the earlier form for the word.
  b. Establish the sound changes that have affected each language.
  c. Construct a cladigram that shows the changes that define each branch.

1. Middle Chinese

Mandarin (Beijing)	Hakka (Huizhou)	Gloss
a. tɕin	kim	'zither'
b. la	lat	'spicy hot'
c. mɔ	mɔk	'lonesome'
d. lan	lam	'basket'
e. tɕi	gip	'worry'
f. lan	lan	'lazy'
g. pa	pa	'fear'

2. Proto-Peninsular Spanish

CASTILIAN	ANDALUSIAN	GLOSS
a. majo	majo	'May'
b. kaʎe	kaje	'street'
c. poʎo	pojo	'chicken'
d. pojo	pojo	'stone bench'
e. dos	dos	'two'
f. dieθ	dies	'ten'
g. θiŋko	siŋko	'five'
h. si	si	'yes'
i. kasa	kasa	'house'
j. kaθa	kasa	'a hunt'
k. θißiliθaθion	sißilisasion	'civilization'

<ʎ> represents a palatal lateral; <ß> represents a voiced bilabial fricative

3. Proto-Numic

YERINTON PAVIOTSO	NORTHFORK MONACHI	GLOSS
a. mupi	mupi	'nose'
b. tama	tawa	'tooth'
c. piwɪ	piwɪ	'heart'
d. soŋo	sono	'lungs'
e. sawaʔpono	sawaʔpono	'proper name (fem.)'
f. nɪwɪ	nɪwɪ	'liver'
g. tamano	tawano	'springtime'
h. pahwa	pahwa	'aunt'

i. kuma	kuwa	'husband'	
j. wowaʔa	wowaʔa	'Indians to the West'	
k. mɪhɪ	mɪhɪ	'porcupine'	
l. noto	noto	'throat'	
m. tapa	tape	'sun'	
n. ʔatapɪ	ʔatapɪ	'jaw'	
o. papiʔi	papiʔi	'older brother'	
p. patɪ	petɪ	'daughter'	
q. nana	nana	'man'	
r. ʔati	ʔeti	'bow, gun'	

4. Maya (England, 1994, Autonomia de los idiomas Mayas)

K'ICHEE'	Q'ANJOB'AL	POPTI'	GLOSS
a. ja	na	ŋa	'house'
c. jee'	ne	ŋe	'tail'
e. oj	on	oŋ	'avocado'
h. jun	jun	jun	'one'
i. q'an	q'an	q'an	'yellow'
j. naj	naj(at)	naj(at)	'far'
k. jolomaj	jolomej	jolomej	'head'

5. Proto-Central Pacific

	Maori	Hawaiian	Samoan	Fijian
1. post	pou	pou	pou	bou
2. forbidden	tapu	kapu	tapu	tabu
3. cry	taŋi	kani	taŋi	taŋi
4. keel	takere	kaʔele	taʔele	takele
5. stay, sit	hono	hono	fono	vono
6. light, moon	marama	malama	malama	malama
7. thatch	kaho	ʔaho	ʔaso	kaso

6. Proto-Uto-Aztecan

		Papago	Hopi	Huichol	Tarahumara
1.	name	tʃɨɨgig	tɨŋʷa	-tewa	riwa
2.	tooth	taatam	tama	ta:me	rame
3.	heavy	wɨɨtʃ	pɨti	hete	bite
4.	red	hɨt	sɨta	ʃeta	sita
5.	reed	waapk	pa:qa	haka	baka
6.	hair	kuup	ko:pa	kɨ:pa	kupa
7.	road	woog	pöhɨ	huye	----
8.	husband	kun	ko:ŋya	kɨna	kuna
9.	ear	naak	na:qa	naka	naka
10.	leg	kahi	qa:si	----	kasi

7. Proto-Uto-Aztecan

	SHOSHONE	UTE	NORTHERN PAIUTE	GLOSS
a.	tuhu	tuu	tuhu	'black'
b.	nika	ni̥ka	nika	'dance'
c.	kasa	kḁsi	kasa	'feather'
d.	tuku	tu̥ku	tuku	'flesh'
e.	juhu	juu	juhu	'grease'
f.	pida	pida	pita	'arm'
g.	kadi	kadi	kati	'sit'
h.	kwasi	kwḁsi	kwasi	'tail'
i.	kwida	—	kwita	'excrement'

8. Proto-Western Turkic

	TURKISH	AZERBAIJANI	CRIMEAN TARTAR	KAZAN TARTAR	GLOSS
a.	burun	burun	burun	bɪrɪn	'nose'
b.	kabuk	gabɪx	—	kabɪk	'bark'
c.	bojun	bojun	mojun	mujɪn	'neck'
d.	toprak	torpax	toprak	tufrak	'earth'
e.	kujruk	gujruk	kujruk	kɪjrɪk	'tail'
f.	japrak	jarpak	dʒaprak	jafrak	'leaf'

9. Arabic

Data from M. Y. Van Wagoner, Spoken Iraqi Arabic (New York 1949) and Introduction to the Spoken Arabic of Lebanon (Washington, DC 1960).

	IRAQI	LEBANESE
1. arm	ðraa9	draa9
2. black	sooda	sooda
3. cheek	xad	xad
4. cold	baarid	baarid
5. corn	ðurra	durra
6. figs	tiin	tiin
7. girl	binit	bint
8. gold	ðahab	dahab
9. heavy	θgiil	t'iil
10. house	beet	beet
11. like	miθil	mitil
12. medicine	dawa	dawa
13. oil	zeet	zeet
14. second	θaani	taani
15. snow	θiliǰ	tilž
16. he studied	diras	daras
17. he takes	yaaxuð	yaaxud
18. third	θaaliθ	taalit
19. this	haaða	hayda
20. he wrote	kitab	katab

Reconstruct dental stops and spirants. Do not look for conditioned changes. /9/ is a 'voiced pharyngeal spirant', /'/ a glottal stop.

10. Italian (Langacker 1972)

Listed below are corresponding forms from three Italian dialects; I is the standard dialect, II is the regional variety of northern Italy, and III is the Lombard dialect (Wanner 1970). The underlying representations, which can be assumed to be identical for the three dialects, are listed at the left. Determine precisely how the three dialects differ.

	DIALECT I	DIALECT II	DIALECT III	GLOSS
/fisso/	fisso	fiso	fis	'fixed'
/kassa/	kassa	kasa	kasə	'cabinet'
/kasa/	kaasa	kaaza	kaazə	'house'
/kosa/	koosa	kooza	koozə	'thing'

11. Papago (Langacker 1972)

The rules (A)-(C)) below are needed for the most insightful description of Papago phonology:

(A) /ḍ/ becomes [l], and /ṣ/ becomes [s], when followed by [i].
(B) A dental consonant becomes palatal before a high vowel.
(C) A word-final high vowel is deleted after an alveolar or palatal consonant.

All three rules function in the derivation of the forms in column I below, which represent one dialect of Papago, Columns II and III contain corresponding forms from other Papago dialects (Hale 1965). Analyze this data and determine how dialects II and III differ from each other and from dialect I. Assume the phonological representations, given at the left, to be the same for all three dialects.

	DIALECT I	DIALECT II	DIALECT III	GLOSS
/ʔuuṣi/	ʔuus	ʔuus	ʔuus	'stick'
/tini/	tʃiñ	tʃiñi	tin	'mouth'
/hɨwɨḍi/	hɨwɨl	hɨwɨli	hɨwɨ	'wind'
/ṣuudagi/	ṣuudagi	ṣuudagi	ṣuudag	'water'
/tɨhani/	tʃɨhañ	tʃɨhañi	tɨhan	'hire'
/naadi/	naadʒ	naadʒi	naad	'something kindled'
/duuki/	dʒuuki	dʒuuki	duuk	'rain'

12. Indo-European (Langacker 1972)

Listed below are sets of related forms from four Indo-European languages which represent four separate branches of the Indo-European family. Consider just the italicized consonants and state the systematic sound correspondences that they reveal. The letters have the value of the identical phonetic symbols with the following exceptions: Latin c = [k], qu = [kʷ]; Sanskrit ś = [ʃ]; English th = [θ], wh = [h].

LATIN	GREEK	SANSKRIT	ENGLISH	GLOSS
*p*ater	*p*atēr	*p*itā	*f*ather	
*c*entum	he*k*aton	*ś*atam	*h*undred	
*p*ēs	*p*ōs	*p*āt	*f*oot	
*qu*od	*p*oteros	*k*as	*wh*o	
*t*rēs	*t*reis	*t*rayas	*th*ree	
*c*or	*k*ardiā	—	*h*eart	
*t*enuis	*t*anaos	*t*anuḥ	*th*in	
de*c*em	de*k*a	da*ś*a	—	'ten'
lin*qu*ō	lei*p*ō	ri*ṇ*akti	—	'let'

13. Cupan (Langacker 1972)

Examine the Cupan cognate sets below (Bright & Hill 1967) and state the sound correspondences that account for the vowels. Reconstruct a proto Cupan segment for each correspondence; reconstruct only one proto segment for two or more correspondences that are environmentally conditioned variants of one another. State the sound changes that must have occurred in the evolution of the daughter languages based on the proto segments you reconstruct. Ignore the difference between long and short vowels; treat long vowels as if they were short vowels of the same quality. Ignore the elements enclosed in parentheses.

CAHUILLA	CUPEÑO	LUISEÑO	GLOSS
kut	kut	kut	'fire'
ne(ʔ)	nə	noo	'I'
waaviš	waviš	waaviš	'foxtail grass'
nexiš	nixiš	neexiš	'gourd'
pe(ʔ)	pə	po	'he'
ʔis	ʔis	ʔes	'teardrop'
haal	hal	hal	'look for'
tew	təw	toow	'see'
mex(an)	mix(ən)	—	'possession'
wih	wih	weh	'two'
nit	nit	net	'pregnant woman'
yaw	yaw	yaaw	'bring'
puul	puul	puul(a)	'doctor'
pi(ʔ)	pi	pi(ʔ)	'breast'
mu	mu	muu(vi)	'nose'
ma	ma	maa	'hand'
ʔamu	ʔamu	ʔaamu	'hunt'
čaʔiš	čaʔiš	čaaʔiš	'bluebird'
pit	pit	pet	'road'
pal	pal	paal(a)	'water'
teʔ(e)	təʔ(e)	tooʔ	'borrow'
wex(et)	—	wex(eʔtut)	'pine'

14. Cupan (Langacker 1972)

Reconstruct the proto Cupan lexeme that underlies each cognate set in problem 11. Do not consider parenthesized segments and ignore the difference between long and short vowels. Your first step will be to establish correspondences for the consonants and to reconstruct proto segments that underlie them.

15. Uto-Aztecan (Langacker 1972)

Below are cognate sets from six Uto-Aztecan languages, including reconstructed proto Cupan.
Consider only the initial consonant and vowel of each form (excluding segments in parentheses),
and ignore vowel length. State the correspondences for the initial consonants and vowels. Select
a proto segment to underlie each correspondence, and reconstruct the initial consonant and vowel
of the proto Uto-Aztecan root that underlies each cognate set. State the sound changes that have
affected each daughter language, and evaluate whatever evidence there may be for grouping any
of these daughters into subfamilies.

CUPAN	CORA	HOPI	PAPAGO	YAQUI	HUICHOL	GLOSS
*puʃ	hɨʔɨsi	poosi	wuhi	puusi	hɨʃi	'eye'
*pet	huye	pőhɨ	woog	booʔo	huye	'road'
*nəma	neemʷa	nɨɨma	ñɨm	—	nema	'liver'
*muk	mɨʔɨtʃi	mooki	muuki	muuke	mɨki	'kill, die'
—	(ti)hete	pɨtɨ	wɨɨtʃ	bette	hete	'heavy'
*ʔeŋ	ʔunah	ʔőŋa	ʔon	ʔoona	ʔuna	'salt'
*tama	tame	tama	taatam	tammin	taame	'tooth'
*tʃivu	tsihivi	tsiivo	siw	tʃiibu	tsiɨ	'bitter'
*piva	—	piiva	wiw	biiva	—	'tobacco'
*naqa	—	naaqa	naak	nakka	naka	'ear'

16. Athabaskan

Examine the following verb and noun stems in four Athabaskan languages, extract the stem-initial consonant sets only and try to arrange them in order of their place of articulation. Assume that there are no consonant clusters in the data, i.e., all apparent combinations of consonants are to be considered single phonological units. Cʷ are labialized; Cʸ are palatalized; C' are glottalized, while /tθ/ and /dð/ are interdental affricates. Ɣ with the rightward hook beneath are nasal; accented vowels have high pitch; los pitch is unmarked. Make the simplifying assumption that set lacking cognates may fill in data from other partial sets and be collapsed with them, and you should be able to isolate about 15 different sets. Then reconstruct a likely proto-phoneme for each set. Do not look for conditioning factors in this problem.

GLOSS	HUPA	CHIPEWYAN	SARSI	NAVAJO
1. afraid	-gʸid	-ʤɛr	-ʤí'	-dzíd
2. bark	-sits'	θəð	--	sis
3. basket	-ts'a	tθ'áí	--	ts'aa'
4. big	-kʸoh	-tʃoɣ	-tʃóó	-tso
5. black	-hʷin	-zən	--	-ʒin
6. blanket	--	ts'ɛrɛ	ts'idí	-tʃ'idí
7. blow	-yooɬ	-yúɬ	-yuɬ	-yóóɬ
8. body	--	-zí	-zi'	-ʒi'
9. bone	-ts'iŋ'	-tθ'ən	-ts'ín	-ts'in
10. day	ʤiŋ-	-dzį	dzín-	-ʤį́
11. do	tʃʷe	-tsi	-tsi'	-tʃí
12. firewood	tʃʷiǰ	tsɛz	--	tʃíž
13. food	--	ʃi-	ʃih	--
14. here	-gʸaŋ	ʤạ	--	dzą́ą́
15. I	hʷe	si	sí	ší
16. liver	-sit'	-ðər	-zi'	zid
17. meat	-tsiŋ'	-tθə́n	--	-tsį'
18. navel	ts'eek'	-tθ'əɣɛ	-ts'ák'-	-ts'éé'