Inheritance vs. lexical borrowing: some Indo-European cases.

        The difference between regular sound changes and other types of changes (most
of which are motivated by morphology) is so important that historical linguists symbolize
them differently in summaries of changes. In what follows, ">" indicates one or more
regular sound changes, while "→" indicates changes of other kinds, sometimes lumped
together as "analogical" changes. The most common of the latter is *levelling.* If one
sound occurs in some forms of an inflectional paradigm and another occurs in the same
position in other forms of the same pattern, one or the other can be generalized, or
"levelled", through the entire paradigm. (This is *not* regular sound change because it's
conditioned by morphology.) In the same way, if a word is accented on the root in some
parts of its paradigm but on the suffix in others, one accent or the other can be general-
ized; if a noun belongs to one gender in the singular but to another in the plural, one or
the other can be generalized to both numbers; and so on. Most of the analogical changes
posited below are levellings or other minor adjustments of well-known types. The major
exception is the Greek word for 'horse', on which see further below.

1. The non-Anatolian word for 'wheel'.
Reconstructable form: PIE *$k^w\acute{e}k^w$lo-s (masc.), collective *$k^w ek^w$lé-h_2 (→ neut. pl.).
Analysis: derived from *k^wel- 'turn'; pattern of derivation (reduplication + zero-grade
        root + thematic vowel) is unique (archaic?), so this word is overwhelmingly
        unlikely to have been formed more than once.
Development of attested forms in the daughter languages:
        *$k^w\acute{e}k^w$los > *$k^w\acute{e}k^w$lë (Ringe 1996:74-5, 88, 90-1) > *$k^{wy}\acute{e}k^w$lë (ibid. pp. 102-
            4) > *$k^{wy}\acute{ə}k^w$lë (ibid. pp. 124-8, 139) → Proto-Tocharian *$k^w\acute{ə}k^w$lë
            'chariot, wagon' (with adjustment of palatalization in a reduplicated
            form, ibid. pp. 143-4; or is this just straightforward assimilation?);
            > *kŭkl ~ *kŭkla- > *$k_u$käl ~ *kukla-* → Tocharian A *kukäl ~ kukla-*
                (Ringe 1998; see Kim 1999 for an important revision);
            > *$k^w ə k^w\acute{ə}$lë (Ringe 1987) > Tocharian B *kokale.*
        *$k^w\acute{e}k^w$los ~ *$k^w ek^w$léh$_2$ > *kéklos ~ *keklā́ > Proto-Indo-Iranian *čáklas ~
            *čaklā́;

>→ Vedic masc. *cakrás* (occasionally attested in the Rigveda), neut. pl.

cakrā́(ṇi) → neut. *cakrám,* pl. *cakrā́(ṇi);*

> Proto-Iranian *čaxrah > Avestan *čaxrō* (no pl. attested).

*kʷékʷlos ~ *kʷekʷléh₂ >→ Homeric Greek masc. κύκλος /kúklos/, neut. pl.

κύκλα /kúkla/;

*kʷékʷlos ~ *kʷekʷléh₂ > *hʷéhʷloz ~ *hʷegʷlā́ > Proto-Germanic masc.

*hʷehʷlaz, neut. pl. *hʷeulō (Ringe 2008:72-3, 94-6, 102-3, 108, 146-8);

>→ Old Norse *hvél* and *hjól* (both neut.);

> Proto-West Germanic *hwehl (*hwehul?), *hweul- >→ Old English

hwēol, hweowol, hweogol (all neut.), with substitution of the

productive alternation *h ~ *g for anomalous *h ~ *w (Ringe

2008:108) and various levellings of alternations.

Discussion.

Any of the sound changes peculiar to the first-order daughters would have made undetected borrowing impossible. These include the Tocharian merger of short *i, *e, and *u as *ə; the Indo-Iranian palatalization of the initial velar and the subsequent merger of nonhigh short vowels as *a;* the Greek rounding of the first vowel to (*o and then) *u,* and the consequent unrounding of the labiovelar; and Grimm's and Verner's Laws in Germanic, which radically reshaped the system of obstruents.

But the Tocharian vowel merger must have occurred far down in the independent prehistory of that subgroup, since it was preceded by more than a dozen other regular sound changes (see the chart at Ringe 1996:139). The Indo-Iranian chronology of sound changes is not much more promising: the palatalization of velars has to have been preceded not only by the merger of velars and labiovelars, but also by the resolution of *R̥H-sequences (which sometimes yielded palatalizing front vowels) and the affrication of inherited palatals (since they did *not* merge with palatalized velars). Greek and Germanic seem at first more promising, since Grimm's Law was a comparatively early Germanic sound change (see the chart at Ringe 2008:152) and the Greek vowel rounding could have occurred very early (note that the unrounding of labiovelars next to u-vowels has already occurred in the Linear B documents). But a glance at any probable cladistic tree of the Indo-European family (e.g. the first tree on p. 397 of Nakhleh et al. 2005, or any of the alternative trees in that article) will show that the divergence of Germanic,

Greek, and Indo-Iranian from one another (and from Armenian and Balto-Slavic) proba-
bly occurred fairly late in the initial diversification of the family, so being able to say that
borrowing could not have occurred after "early" changes in any of those languages is less
useful than it might be. (It's true that the divergence of Greek and/or Germanic from the
rest of the family might be as early as 3000 BCE, if the estimated dates of internal nodes
in these trees are in the right ballpark; but that would still be a good 500 years after the
probable divergence of Tocharian from the rest of the non-Anatolian branches, and a
whole millennium after the likely date of PIE.) So it looks like the recoverable relative
chronology of sound changes is not going to be very helpful in this case.

On the other hand, the pattern of shared and unique linguistic changes and the
findings of archaeology turn out to be very helpful. One of the striking things about
Proto-Tocharian is that none of the linguistic changes that characterize it can be shown to
be historically shared with any other subgroup of Indo-European; either they're "natural",
easily repeatable changes which could have occurred independently any number of times
(like the merger of palatals and velars, or the raising of word-final long *ō to *ū; see
Ringe 1990:59-105 and 1996 passim) or they're unique within the family (like the loss of
*bʰ immediately following *m, or the complex pattern of Tocharian vowel mergers). It
appears that the separation of Tocharian from the rest of the family was sharp, and that it
did not again come into contact with other IE languages (specifically, Iranian languages)
for many centuries. (The attempt to connect Tocharian B *tek-* 'touch' with Gothic *tekan*
in Ringe 1990:105-15 is tantalizing but inconclusive; there is too much likelihood that the
words resemble one another by sheer chance. The fact that the similar Romance words
—Italian *toccare,* French *toucher,* etc.—clearly do resemble the Germanic and Tocharian
words by chance (see Meyer-Lübke 1911:664) adds weight to that point.)

Moreover, there seems to be only one archaeological culture that could reflect the
pre-Tocharians, namely the Afanasievo culture. This culture, associated with horses (see
below!), appeared abruptly in the Altai around 3500 BCE and appears to represent a
migration from the lower Volga area some 2000 miles to the west. It's hard to resist the
conclusion that the Afanasievo migration represents the separation of pre-Tocharian from
the rest of the family (Anthony 2007:264-5); and if that's true, then the odd reduplicated
word for 'wheel' must already have been in existence, and have been inherited by or bor-
rowed into (or out of) pre-Tocharian, before 3500 BCE. That's later than any date that

most of us would assign to PIE, but not much later, and for the purposes of reconstructing palaeocultures it's not significantly different. The fact that the Tocharian word refers to a wheeled vehicle rather than a wheel is not problematic; words shift their meanings all the time, and this particular shift is not surprising.

So the non-Anatolian word for 'wheel' was either inherited from the last common ancestor of the non-Anatolian branches, or else it was borrowed into or from pre-Tocharian before 3500 BCE. In terms of time depth there's not much difference between those alternatives.

2. The Proto-Indo-European word for 'horse'.

Reconstructable form: PIE *éḱwos (masc.).

Analysis: apparently unanalyzable.

Development of attested forms in the daughter languages:

> *éḱwos > Proto-Anatolian *áḱḱwos (Melchert 1994:62-3, 74-5) > Proto-Luvian
> *áttˢwos (Melchert 1987, 1994:251-2);
> > Cuneiform Luvian *azzuwas,* Hieroglyphic Luvian *á-zú-wa-;*
> > *asbe > Lycian *esbe* (Melchert 1994:302, 310-1).

> *éḱwos > *ékwë (Ringe 1996, as above) > Proto-Tocharian *yə́kwë (pace Kim
> 1999:158, 163, 167);
> > *yəkʷ > *yŭk > Tocharian A *yuk;*
> > Tocharian B *yakwe.*

> *éḱwos > Proto-Indo-Iranian *áćwas;
> > Vedic *áśvas;*
> > Proto-Iranian *atˢwah > Avestan *aspō,* Old Persian *asa.*

> *éḱwos >→ Greek *íkʷkʷos (cf. Mycenaean *i-qo;* but why *i-?? contamination
> with some other word?) > *íppos (cf. compound names like Ἄλκιππος
> /Álk-ippos/, with no aspiration) → ἵππος /híppos/ (again, where does the
> /h-/ come from?); PROBLEMATIC COGNATE.

> *éḱwos > Proto-Italic *ékwos > Latin *equos.*

> *éḱwos > Proto-Celtic *ekʷos > Gaulish *Epo-* (in names), Old Irish *ech.*

> *éḱwos > Proto-Germanic *ehʷaz > Old Norse *jór,* Old English *eoh;* cf. also
> Old Saxon *ehuskalk* 'mounted retainer', Gothic *aíƕatundi* 'thornbush'

(*'horse-tooth').

*ékʷos > Proto-Balto-Slavic *éšwas; derived fem. *ešwā́ > Lithuanian *ašvà* 'mare'.

*ékʷos > *eš > Armenian *êš* 'donkey'.

Discussion.

The consonant cluster *Ḱw is rare, and we might have hoped that it would develop in some unusual way in many first-order daughters. But once again the daughters in which it underwent changes that should make loanwords detectable diverged from the rest of the family fairly late (to judge from the trees in Nakhleh et al. 2005). In Tocharian and Italo-Celtic it merely underwent the merger of palatals and velars (followed, in Celtic, by a merger with the voiceless labiovelar). The initial vowel, too, survived without change for a long time in most daughters.

But the Anatolian reflex is very distinctive, because of a bizarre sound change that replaced word-initial accented *é plus a single consonant (followed by a vowel or semi-vowel) with accented *á plus a geminate consonant ("limited Čop's Law"; see Melchert 1994, as above). That sound change can be shown to have occurred after another Anatolian sound change (loss of word-initial *h$_1$) but before a third (loss of word-initial *y when followed by an e-vowel), so it was neither among the first nor among the last pre-PA changes (Melchert 1994:90). We should be able to argue that after the limited Čop's Law change an undetectable borrowing of 'horse' into or out of Anatolian would have been impossible.

Unfortunately there is a further complication that undermines any such argument. If we could adduce a Hittite cognate "akkuwas", the argument would be ironclad. But though horses are often referred to in Hittite documents, the scribes never spell the word out (!); instead they use a logogram (word-sign), one of many adopted as part of the cuneiform writing system, which is usually transliterated with the Sumerian phrase ANŠE.KUR.RA 'donkey of the mountains'. All the actually attested Anatolian words for 'horse' are from languages of the Luvian subgroup; and in that subgroup *the initial vowels of all the relevant words are etymologically ambiguous!* Cuneiform and Hieroglyphic Luvian *a-* could in principle reflect Proto-Luvian and Proto-Anatolian *a-, *e-, or *o- (Melchert 1994:262-4). In Lycian the situation is even stranger. First Proto-Anatolian *e and *o merged as *e,* while *a remained distinct as *a.* But then an umlaut rule operated,

changing the frontness of vowels to agree with the frontness of the vowel in the next syllable, and the rule iterated from right to left through the word (Melchert 1994:310-1, 328). As a result, Lycian *esbe* could reflect earlier *asbe (with PAnat. *a-) or earlier *esbe (with PAnat. *e- or *o-). So it turns out that the Proto-Luvian form could actually have been either *áttˢwos or *éttˢwos; and since we have no other direct evidence for the Proto-Anatolian form, that could have been either *áḱḱwos or *éḱḱwos—the former if it was inherited from PIE according to the hypothesis sketched above, the latter if it was borrowed from some other IE language after the limited Čop's Law change had run its course! But what about the geminate stop, which is clearly preserved in Cuneiform Luvian? It turns out that there was yet another Anatolian sound change which geminated voiceless stops, and we don't know when it occurred—it need not have been early (Melchert 1994:62); so borrowing of *éḱwos from another IE language, followed by gemination of the stop in the Anatolian languages, is *not* impossible. Once again the linguistic evidence has left us in the lurch.

And once again cladistics and archaeology come to our rescue. The presence of this word for 'horse' in Tocharian guarantees its existence in the non-Anatolian half of the family by 3500 BCE for the reasons advanced above in the discussion of 'wheel'. The abrupt separation of Tocharian and the fact that that event can (probably) be traced archaeologically are crucial. Unfortunately the archaeological situation for Anatolian is very different. Anthony's suggestion that an expansion of the steppe culture into the Danube delta around 4200 BCE reflects the incipient separation of Anatolian from the rest of PIE (Anthony 2007:249-57) is reasonable, but any connection with Anatolia seems to rest on speculation (ibid. p. 262). Working backwards from the historical record, we know that speakers of Anatolian languages were in central Anatolia by the 19th century BCE; when and by what route they arrived there remain very unclear, though a cultural disruption in the 27th century BCE is apparently a likely candidate for an Anatolian incursion into the area (Mallory 1989:24-9). That still seems to leave many centuries for potential contact between Anatolian and other IE languages.

But the distribution of linguistic innovations tells a different story. Like Tocharian, Anatolian shares *no* distinctive innovations with any other subfamily of IE (cf. Melchert 1994:60-91, Ringe 2000); so far as we can tell, its separation from the rest of the family was reasonably "clean". Moreover, the cladistic tree tells us that that separation

must have been earlier than that of Tocharian. This reduces the viable options to two: either the well-known word for 'horse' was inherited by Proto-Anatolian, according to the scenario sketched above, or else it was borrowed into or out of pre-PA during the relatively short time when pre-PA was still in contact with related languages—and that time must have been some centuries before 3500 BCE, *and* few detectable innovations can yet have occurred in pre-PA.

This raises a methodological point that we can no longer avoid. Is there any difference between a word which is reconstructable for a protolanguage and a word which spread from dialect to dialect of the protolanguage as it was breaking up? As usual, it depends on the individual case. If the real-world separation of the daughters was genuinely abrupt—that is, one group picked up and moved within a generation or so, and subsequent contacts were infrequent and brief—then there is a clear difference between the two scenarios. But most disintegrations of speech communities don't happen like that; dialects remain in contact as they diverge, continuing to trade linguistic material until some event finally makes them lose touch altogether. (The best discussion of these processes is Ross 1997.) In such cases the "protolanguage" which we reconstruct is most unlikely to correspond to a single, completely uniform dialect that existed in the real world before its speaking population became large enough to exhibit significant linguistic diversity; it almost inevitably corresponds to a dialectally diversified speech community, still *unified* but no longer *uniform,* simply because we can't tell the difference between words and grammatical forms which had been in the language for generations and those which had arrived very recently. It is also likely that our reconstruction will be temporally "out of focus", including some inherited words and forms which were no longer characteristic of all the dialects and some new words and forms which were still spreading from dialect to dialect. There are good reasons to suspect that our reconstruction of PIE is like that.

But once again this doesn't make much difference for our reconstructions of palaeocultures. Whether the reconstructable PIE word for 'horse' was already in the common ancestor of all the IE languages in, say, 4200 BCE or spread through a rapidly diversifying IE dialect continuum around 3700 BCE can't be expected to have any impact on subsequent prehistoric and historical developments. In this case, at least, a degree of detail too fine for linguists to recover is also too fine to have any consequences

for history.

A final note about 'horse': the shape of the Greek word can't be explained by regular sound changes and plausible analogical changes. The /h-/ of the Classical form is a problem internal to the history of Greek, since it isn't there in the fossilized compounds used as personal names (thus Ἄλκιππος /Álkippos/ 'His-horses-are-his-defense', not "Ἄλχιππος /Álkʰippos/"). But the /i/ is there from the beginning of our attestation, and it's a total mystery. It's worth thinking about the possibility that the Greek word might be a loanword—if only we knew of a language in which *é- gave i- by regular sound change, or a non-Indo-European language could have borrowed the word and altered it in that way.

In both the above examples we didn't arrive at any firm conclusions by trying to exploit regular sound changes. That raises an obvious question: are there any non-obvious cases in which that approach does give good results? And if it doesn't look likely that 'wheel' or 'horse' is a Wanderwort, what would a Wanderwort look like? I hope to address those questions in a further posting.

Bibliography.

Anthony, David. 2007. *The horse, the wheel, and language.* Princeton: Princeton U. Press.

Kim, Ronald. 1999. "The development of labiovelars in Tocharian: a closer look." *Tocharian and Indo-European Studies* 8.139-87.

Melchert, H. Craig. 1987. "PIE velars in Luvian." Watkins (ed.) 1987:182-204.

—. 1994. *Anatolian historical phonology.* Amsterdam: Rodopi.

Meyer-Lübke, Wilhelm. 1911. *Romanisches etymologisches Wörterbuch.* Heidelberg: Winter.

Nakhleh, Luay, Don Ringe, and Tandy Warnow. 2005. "Perfect phylogenetic networks: a new methodology for reconstructing the evolutionary history of natural languages." *Language* 81.382-420.

Ringe, Don. 1987. "On the prehistory of Tocharian B accent." Watkins (ed.) 1987: 254-69.

—. 1990. "Evidence for the position of Tocharian in the Indo-European family?"

Die Sprache 34.59-123.

—. 1996. *On the chronology of sound changes in Tocharian. Vol. 1.* New Haven: American Oriental Society.

—. 1998. "Schwa-rounding and the chronology of sound changes in Tocharian A." Jasanoff, Jay, H. Craig Melchert, and Lisi Oliver (edd.), *Mír curad: studies in honor of Calvert Watkins* (Innsbruck: IBS) 611-8.

—. 2000. "Tocharian class II presents and subjunctives and the reconstruction of the Indo-European verb." *Tocharian and Indo-European Studies* 9.121-42.

—. 2008. *From Proto-Indo-European to Proto-Germanic. A linguistic history of English, Vol. 1.* Revised ed. Oxford: OUP.

Ross, Malcolm. 1997. "Social networks and kinds of speech-community event." Blench, Roger, and Matthew Spriggs (edd.), *Archaeology and language I: theoretical and methodological orientations* (London: Routledge) 209-61.

Watkins, Calvert (ed.). 1987. *Studies in memory of Warren Cowgill.* Berlin: de Gruyter.