

Indo-European Substrates from an Anatolian Point of View

Methods, Problems, and Case Studies

1. The contribution of the Anatolian languages: what we can expect and what we cannot

1.1. The limits of the research

(a) corpus languages, extremely limited attestation (Pisidian, Sidetic); limited (Carian, Lydian, Palaic, Lycian B); quite well attested (Luwian, Lycian A); well attested (Hittite);

(b) some undeciphered signs => phonetic shape is not always clear (Sidetic, Carian, Hieroglyphic Luwian)

(c) *scriptio continua* in some of the languages => many words are unknown, although attested (Sidetic, Carian)

(d) the meaning of a huge number of words in all languages is unknown and frequently only approximately known

(f) there is no up-to-date and ready philological dictionary for any of these languages

(g) with the exception of HEG, there are no full-scale etymological dictionaries of these languages either (but see *eDiAna* for (f) and (g))

=> lack of parallels in Anatolian of an alleged substrate word does not necessarily mean that there is no Anatolian cognate of the substratum words => negative Anatolian evidence is not evidence

1.2. Possible substrates

- Anatolian only (not discussed today) vs. a common substrate of Anatolian + any other IE language (today's topic);

- the common substrates: non-areal vs. areal substrates (Greek, Phrygian, and Armenian)

- areal substrates:

1) Greek - Anatolian - Armenian:

2) Greek - Anatolian (the common one, not the alleged Anatolian substrate in Greek [not discussed here]);

3) “Mediterranean”: Anatolian - Greek - Latin

4) Anatolian - Armenian

- support from local toponyms? not yet

2) State of the art / proposals & case studies

- how to prove that we are dealing with a substrate word? See, e.g., Schrijver 1997: 294-297:

1) limited geographical distribution

2) semantic fields conducive to borrowing

3) without etymology

4) recurring phonological irregularity

5) recurring morphological irregularity

6) hint for non-IE phonology / morphology

- note: Nos. 1-3 may indicate but do not prove

1 Non-Areal Substrate

1.1. Hitt. *alanza(n)*- ‘(a tree and its wood)’: of unclear etymology (HED A: 30; HEG A-K: 15, both with refs.)

PROP: Puhvel 1977: 598, HED A: 30 “most probable” ~ Lat. *alnus*, Lith. *al̃ksnis*, East Lith. *al̃ksnis* ‘alder’, “Hitt. has metathesized **al(i)sno-* to **al̃so-*” or from **aleno-* > *alnus* - add also further Baltic forms (Lith. *el̃ksnis*); Slavic **a/elisaH* > Ru. *ol’xá* ‘alder’ (dial. *el̃xa*, *elxá*), Bulg. *elxá* ‘alder, spruce’; Germanic **aluz-* > OIc. *pl̥r*, OE *alor* ‘alder’ and **alis/zō* > OHG *elira*, MoDu *els*, Go. **alisa* > Spanish *aliso* (e.g., Friedrich 1970: 70-72 with refs. [and possible Ancient Macedonian and Celtic forms]);

DISC: (1) Baltic, Germanic, and Latin: rather non-IE loanwords, cf. *e/a*-variation in *Anlaut* & suffixes *-s-* and *-is-* (EDSIL: 371, EDBIL: 50-51, EDLI: 35; cf. also EIEC: 11 (Friedrich), Mallory – Adams 2006: 158, EDPG: 22), (2) but **al̃sn̥*° and **alsn̥*° are not fitting, (3) the

metathesis is fully *ad hoc* (rejected already by Melchert 2003: 134); (4) etymology of the Hittite word by Poetto 1973: 29: ~ Greek ἐλάτη ‘pine, fir’, now: via **elntyō-* (on PIE **e* > Hitt. *a/_RCC* see EDHIL: 95, on *-*Vntyō-* > Hitt. -*Vnza-* see Melchert 2003: 134-135).

1.2. Hitt. *kanen(iye/a)-* ‘to bow down, to crouch, to squat’: of disputed etymology (HEG A-K: 480-481, HED K: 42, EDHIL: 434, all with refs.)

PROP: Puhvel 1981: 352, HED K: 42 (“most probably”) ~ PIE **knej-* >> **knej-g^{wb}*- (Lat. *cō-nīvēō* ‘to close (the eyes)’, Goth. *hneiwan*, OE *hnīgan*, OHG *nīgan* ‘to bend down, to bow’) and **knej-b-* (ON *hnípa* ‘to be downcast’, Lith. *knìbti* ‘to collapse’) as **kn(e)ḡ-n-* (but see LIV²: 365, 366 (M. Kümmel): ?**knejb^b*- ‘hängen lassen, sinken lassen’ and **Knejg^{ub}*- ‘(sich) neigen’)

DISC: (1) some scholars: these are substrate words (cautiously EDLI: 130 [about the first group due to its limited distribution and its **T-D^H* structure]; EDHIL: 434, latter arguing that exactly that is why “a connection with the Hittite verb is rather improbable”); (2) any connection is possible only if the suffixes “-*g^{ub}*-” & “-*b^(b)*-” & “-*n-*” can independently be confirmed; (3) note the plausible etymology in EDHIL from PIE **ḡen-* ‘to bend’ (whence **ḡen-u-* ‘knee’) as **ḡ-né-n-ti*;

1.3. Hitt. *šapiya-* ‘to scrub (?)’, to rub (?)’: of unknown etymology (HEG L-S: 844-845 with refs., HED Sa: 137; not included in EDHIL);

PROP: HED Sa: 137 ~ OE *sāpe* ‘soap’, Lat. *sēbum* ‘tallow’, *sāpō(n)-* ‘soap’ (from Germanic) < PIE **sab-*;

DISC: (1) *sēbum*: of unknown origin (EDLI: 550 with refs.), Latin *sāpō* cannot be a loan from Germanic from phonological reasons (cf. OHG *seifā*, *seipfa*, *seiffā*, MLG *sēpe*, MDutch *sēpe*, OFrisian *sēpe*, and OE *sāpe* ‘soap’ <**saip^o*); and the Germanic and the Latin words from a local substrate identified by P. Schrijver and characterized by Celtic *-a-* / Germanic *-ai-* (Simon 2020); (2) the substrate form (**saippwōn-*) and the vocalism of the Germanic words in general cannot be reconciled with the Hittite word; (3) the Hittite word is derived now as *šapiyā(i)-*

from Hitt. *ša/epiya-* ‘washbowl’ (Rieken 2020) from Proto-Anatolian **seP-/sP-ié/ó-* ‘to scrub, to cleanse’ (Sasseville – Rieken 2020);

1.4. Hitt. *šumanzan-* ‘(bul)rush’: of unknown etymology (EDHIL: 781, *contra* HEG L-S: 1151-1152 following an outdated meaning)

PROP: G. Kroonen (pers. comm.): ~ PGerm. **sem-epa/ō-* ‘rush’ (> OS *semith*, MLG *sem(e)de*, OHG *semide*, etc.), PCelt. **sem-ino-* (?) (> OIr. *simin*, *sibin(n)*, *sibhean(d)*), see EDPG: 432;

DISC: (1) The Hittite word is formally ambiguous (**-m-* or **-w-*), (2) cf. ? CLuw. *šummanti*[...] ‘?’, quasi hapax (KBo 22.254, 3.11, CLL: 196), does not really fit the context semantically; (3) in case of **-m-* since *-anzan-* can be explained regularly as a suffix (Melchert 2003) => Hitt. **sum-*: PIE **sm-* > Hitt. *šum-* (e.g., Kimball 1999: 199, Rieken 2000), but see EDLI: 783-785 => these groups do not seem related

(2 Greek-Anatolian-Armenian substrate)

3 Greek-Anatolian substrate

(A) Formally possible?

3.1. Hitt. *arši-* ‘plantation’ ~ Greek *ἄρσηα* ‘meadows’: common substrate? (Simon 2018: 384-385 with refs.)

3.2. Hitt. *iškiš-* ‘back, backside, rear’ ~ Greek *ἰσχίον* ‘hips’, (Hes.) *ἰσχι* ‘loins’: *if* cognate, then loans from a common, third source (EDHIL: 403)

3.3. Hitt. *šulāi-/šūliya-* ‘lead’ ~ Greek *σόλος* ‘iron mass’: common borrowing / substrate (Gusmani 1968: 84; Egetmeyer 2010: 301-302, cautiously)

(B) Formally possible? With recurring phonological anomalies?

(a) Hitt. *-š-* : Gr. *-d-*

3.4. Hitt. *karšani(ya)-* ‘sodaplant, soapwort’: of unknown etymology (HEG A-K: 521, HED K: 106-107, both with refs.; not included in EDHIL)

PROP: Furnée 1972: 64 n. 269: ~ *κάρδαμον* / *καρδάνη* ‘nose-smart / watercress’ of unknown etymology (HEG A-K: 521: “vielleicht”)

DISC: (1) °*amon* analogically, cf. *σήσαμον*, *κίνναμον*, etc.; (2) possible problem: lax semantics;

3.5. Hitt. *paršna-* ‘leopard’: either from PIE **prs-no-* ‘dappled, having spots’ (Oettinger 1986: 22, followed by Melchert 1994: 175; EDHIL: 644) or “Wanderwort”: ~ Hatt. *ḫa-prašš-un*, Türkmen *bars*, Persian *pārs*, Greek *πάρδαλις*, *πόρδαλις* ‘leopard, panther’, etc. (cf. HEG A-K 478-479 with refs., HED Pa: 173 [PIE etymology “might be”]), common Hittite-Greek substrate: already Gusmani 1968: 85;

DISC: (1) the only geographically-chronologically fitting words are the Hattian and the Greek forms, (2) Greek and Hittite (and even Hattian??) might have borrowed from a common source (**prC-*?); (3) Problem: PIE etymology formally more plausible vs. loanword etymology semantically more plausible;

=> loans from a third, common source with a phoneme unknown to both Hittite and Greek (perhaps [θ] or [ð])?

(b) Hitt. nasal + labial stop : Gr. labial stop only

3.6. Hitt. *impa-* (also *aimpa-*) ‘weight, burden’ of unknown etymology (HEG A-K 6, HED A: 15; not included in EDHIL)

PROP: Furnée 1972: 271: ~ Greek *ἵπος* ‘weight, press’ (HED A: 15: “plausibly”, “presumably (...) borrowed from common source”, but HEG A-K: 6: “ganz unwahrscheinlich”);

3.7. Hitt. *kurimpa-* ‘dregs, sediment’ of unknown etymology (HEG A-K: 647; HED K: 265; not included in EDHIL)

PROP: Furnée 1972: 271: ~ Greek *κυρήβια* ‘husks, bran’ (HEG A-K: 647: “vielleicht”; HED K: 265: “possible”; but EDG: 806: “there seems little rea[s]on for this”).

=> loans from a third, common source with a phoneme unknown to both Hittite and Greek (perhaps nasalised vowel or prenasalised consonant)? Carian fits ...

4 “Mediterranean” substrate

4.1. Hitt. ^{TU7}*ša(m/n)pukki-* ‘(a type of stew or soup)’: of unknown origin (HEG L-S: 802; HED Sa: 106-107; not included in EDLI)

PROP: HED Sa: 106-107: “a pot-dish (ingredient)”, “typical culinary culture word” ~ Lat. *sa(m)bucus* ‘elder-tree’ (following Knobloch 1955: 8-9) as well as Gr. *σάμψ(ο)υχον* ‘marjoram’ (Knobloch and HEG L-S: 802 add Dacian *σέβα* and Gaul. *σκοβιή* ‘elder(-tree/berry)’ and prefers the Greek connection from a culinary point of view);

DISC: (1) the motivation behind the name of the soup is unknown, an ingredient is of course possible; (2) but no hint about the ingredients; (3) formal assonance only with the Latin word; (4) elderberries are not used for soups, but for drinks or marmalades, attestations from ritual texts, thus it may refer to such potions;

4.2. Hitt. *ḫaš(š)ik(ka)-* ‘(a tree and its fruit, perhaps ‘nut)’: of unknown etymology (HEG A-K: 200-201; HED H: 230; HW² H: 423; cf. also EDHIL: 325)

PROP: Hoffner 1967: 43 n. 58: *ḫa-š(š)ik(ka)-* ‘a kind of fig’ (cf. *maršigga-* and *šiggašigga-*) ~ Greek *σῦκον*, *τῦκον*, but see also Lat. *figus*, Arm. *t’owz* ‘fig’ (HED H: 230: “at best conjectural”);

DISC (1) *šiggašigga-* is a noun or an adjective (CHD Š: 359 with refs.), *maršikka-* is a tree or its fruits (CHD L-N: 200), their etymology is unknown (HED M: 89, HEG L-S: 147, 1035-1036, both with refs., not included in EDLI);

(2) an old Indo-European etymology (e.g., the ref. in HEG A-K: 201, EDAIL: 398 (“perhaps”), cf. also EIEC: 32 (P. Friedrich), Mallory – Adams 2006: 158-159): ~ Proto-Germanic **aska-* ‘ash’ < **Hh₃esko-* and Alb. *ab* ‘beech’, Arm. *ḫac^ci* ‘ash-tree’ < **Hh₃esk-*, but to be excluded: **-sk-* does not receive an epenthetic vowel in Hittite in this environment (EDHIL: 60-61); a comparison with Baltic (**Heh₃-s-io-* > Lith. *úosis* ‘ash-tree’, etc.), Slavic (**Heh₃-s-en-* > SCr. *jäsēn*, etc. ‘ash-tree’), Latin (*ornus* ‘a kind of ash-tree’), and Celtic (**Hh₃-es-no-* > Mlr. *onn* ‘pine tree’, etc.): derivation unparalleled, spelling variation unexplained;

(3) spelling variation => loanword, similar to adapted Hattian toponyms (type *Nerik(ka)*) + *ḫa-* typical nominal prefix in Hattian (Soysal 2004: 217) => Hattian loan in Hittite;

4.3. Hitt. *ḫalenzu*- ‘surface growth of stationary water (algae, duckweed, foliage, etc.)’: of unknown etymology (HEG A-K: 128, HED H: 20, both with refs.; not included in EDHIL)

PROP: Pisani 1967: 403: OHG *linsa*, Lat. *lēns* ‘lentil’, also OCS *lęsta* (HED H: 20: “unwarranted association”)

DISC: (1) ? Hattian origin for the Hittite word (Güterbock 1974: 309-310; HW² H: 27): formally plausible, *ḫa*= prefix + *=u* (regular Hattian morpheme, Soysal 2004: 259); (2) semantically no problem: duckweed = Hung. *békalencse*, lit. ‘frog-lentil’;

=> Hattian **lenz*, **sik*: is there more than superficial similarity?

3) Instead of conclusions

a) generally low overall number of possible words belonging to a common substrate of Anatolian with any other Indo-European language is not surprising considering (1) the limited knowledge of the Anatolian vocabulary in many languages as well as the (2) prehistory of the Anatolian branch

b) there seems to be no non-areal substrate words in Anatolian, which fits our prehistorical expectation, but it cannot be used as a prehistoric argument due to the limited availability of the Anatolian lexicon

c) there seems to be no exclusive Anatolian-Armenian substrate, which might be explained by historical-geographical factors, but rather to be attributed to the known limits

d) common Anatolian - Greek substrate: relatively numerous, there might be even recurring phonological irregularities (a connection to Carian ???) - the possibilities are seriously limited by the fact that we hardly know the vocabulary of the Anatolian languages of the western coast

e) some hints to a Mediterranean substrate including Latin, but currently it is hard to prove

f) a remarkable superficial Hattian connection - only a Fata Morgana?

References (beyond the usual abbreviations of lexica)

- Egetmeyer, Markus (2010): *Le dialecte grec ancien de Chypre*. Berlin – New York
- Friedrich, Paul (1970): *Proto-Indo-European Trees. The Arboreal System of a Prehistoric People*. Chicago – London
- Furnée, Edzard J. (1972): *Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen. Mit einem Appendix über den Vokalismus*. Den Haag
- Gusmani, Roberto (1968): *Il lessico ittito*. Napoli
- Güterbock, Hans Gustav (1974): The Hittite Palace. In: Paul Garelli (ed.): *Le palais et la royauté (Archéologie et civilisation)*. XIX^e Rencontre Assyriologique Internationale. Paris, 305-321.
- Hoffner, Harry A. (1967): An English-Hittite Glossary. *RHA* 25: 5-99.
- Kimball, Sara E. (1999): *Hittite Historical Phonology*. Innsbruck.
- Knobloch, Johann (1955): Holunder im Hethitischen? *PHMA* 1: 5-10.
- Mallory, James P. – Adams, Douglas Q. (2006): *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford
- Melchert, H. Craig (1994): *Anatolian Historical Phonology*. Amsterdam – Atlanta
- Melchert, H. Craig (2003): Hittite Nominal Stems in -*anzan-*. In: Eva Tichy – Dagmar S. Wodtko – Britta Irslinger (eds.): *Indogermanisches Nomen. Derivation, Flexion und Ablaut. Akten der Arbeitstagung der Indogermanischen Gesellschaft. Freiburg, 19. bis 22. September 2001*. Bremen, 129-139.
- Neumann, Günter (1971): Review of StBoT 7-13. *IF* 76: 260-274.
- Oettinger, Norbert (1986): *“Indo-Hittite“-Hypothese und Wortbildung*. Innsbruck
- Pisani, Vittore (1967): Review of *HW*. 3. *Ergänzungsheft*. *Paideia* 22: 402-404.
- Poetto, Massimo (1973): Su alcuni termini botanici etei. *RIL* 107: 25-32.
- Puhvel, Jaan (1977): Review of *HW*². *Lieferung 1*. *JAOS* 97: 596-599.
- Puhvel, Jaan (1981): Review of *HEG*. *Lieferung 3*. *BiOr* 38: cols. 349-354.
- Rieken, Elisabeth (2000): Hethitisch *šumumahh-*. *HS* 113: 171-175.
- Rieken, Elisabeth (2020): Cuneiform Luwian *šapiyā(i)*-(^{di}) ‘to scrub, rub, cleanse’. In: *eDiAna*. <http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=1288>
- Sasseville, David – Rieken, Elisabeth (2020): Proto-Anatolian **seP-/sP-yé/ó-* ‘to scrub, to cleanse’. In: *eDiAna*. <http://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=1290>
- Schrijver, Peter (1997): Animal, vegetable and mineral: some Western European substratum words. In: Alexander Lubotsky (ed.): *Sound law and analogy. Papers in honor of Robert S. P. Beekes on the occasion of his 60th birthday*. Amsterdam – Atlanta, 293-316.
- Simon, Zsolt (2018): Anatolian influences on Greek. In: Łukasz Niesiołowski-Spanò – Marek Węcowski (eds.): *Change, Continuity, and Connectivity. North-eastern Mediterranean at the turn of the Bronze Age and in the early Iron Age*. Wiesbaden, 376-418.
- Simon, Zsolt (2020): Latin *sāpō* ‘hair-dye, soap’, the Germanic Words for Soap, and the Common Substrate of Celtic and Germanic. *Amsterdamer Beiträge zur älteren Germanistik* 80: 391-400.
- Soysal, Oğuz (2004): *Hattischer Wortschatz in hethitischer Textüberlieferung*. Leiden – Boston