

Sémi összehasonlító nyelvészet

10. A „legvalószínűbb történet” és Hetzron modellje.
Kvantitatív módszerek. Az arab besorolása.

Biró Tamás

BMA-HEBD-303

2016. május 11.



Arab helye a nyugati sémi nyelvek között

- Érvek a déli besorolás mellett, például:
 - Földrajz
 - Tört többes szám megléte.
 - [p] > [f]
 - L-igetörzs: *qātala* (v.ö. klasszikus arab III/fā^cala törzs)
- Érvek az északi/centrális besorolás mellett, pl.:
 - Emfatikus hangok képzése: dél–glottalizáció *vs.* centrális–faringalizáció.
 - Szuffixkonjugáció homogenizált mássalhangzója: dél–[k] *vs.* centrális–[t].
 - Második gyökms h geminációja prefixragozásban: dél–igen *vs.* centr.–nem.
- Melyik érv mennyire győz meg? Tipológiai érv? Történeti érv?

Arab helye a nyugati sémi nyelvek között

- Látszólagos érvek déli mellett: areális hatás (hullámmodell)?
- Érvek centrális mellett: közös innováció (családfamodell)?
 - Szuffixkonjugáció homogenizált mássalhangzója: dél-[k] *vs.* centrális-[t].
 - Prefixragozás: **yaqattal* (pl. akkád, geez, mehri) *vs.* centr. sémi **yaqtulu*.
- További érvek centrális mellett **vagy** areális hatás?
 - A nőnem *-t* végződése lekopik.
 - Triftongusok monoftongizációja: *-aya, -awa* < *-ā*
 - Emfatikus hangok képzése: dél-glottalizáció *vs.* centrális-faringalizáció.
 - Főnév határozottságának jelölése
(**hal-* > héber *haG-*, arab *?al-~?aG-* eredménye?)



Arab helye a nyugati sémi nyelvek között

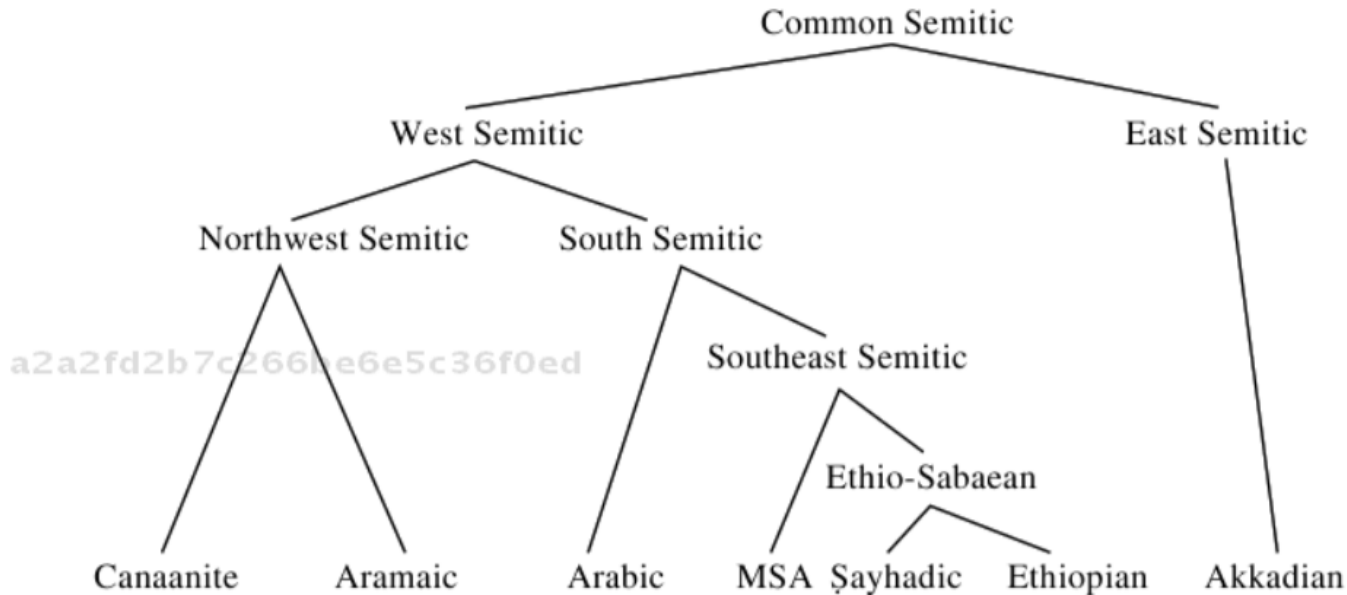


Figure 9.1: Traditional Subgrouping (after Faber 1997)

J. Huehnergard and A. Rubin. Phyla and Waves: Models of classification of the Semitic languages. In S. Weninger (ed.), *Semitic Languages: An International Handbook* (de Gruyter Mouton, 2011), pp. 259–278.

Arab helye a nyugati sémi nyelvek között

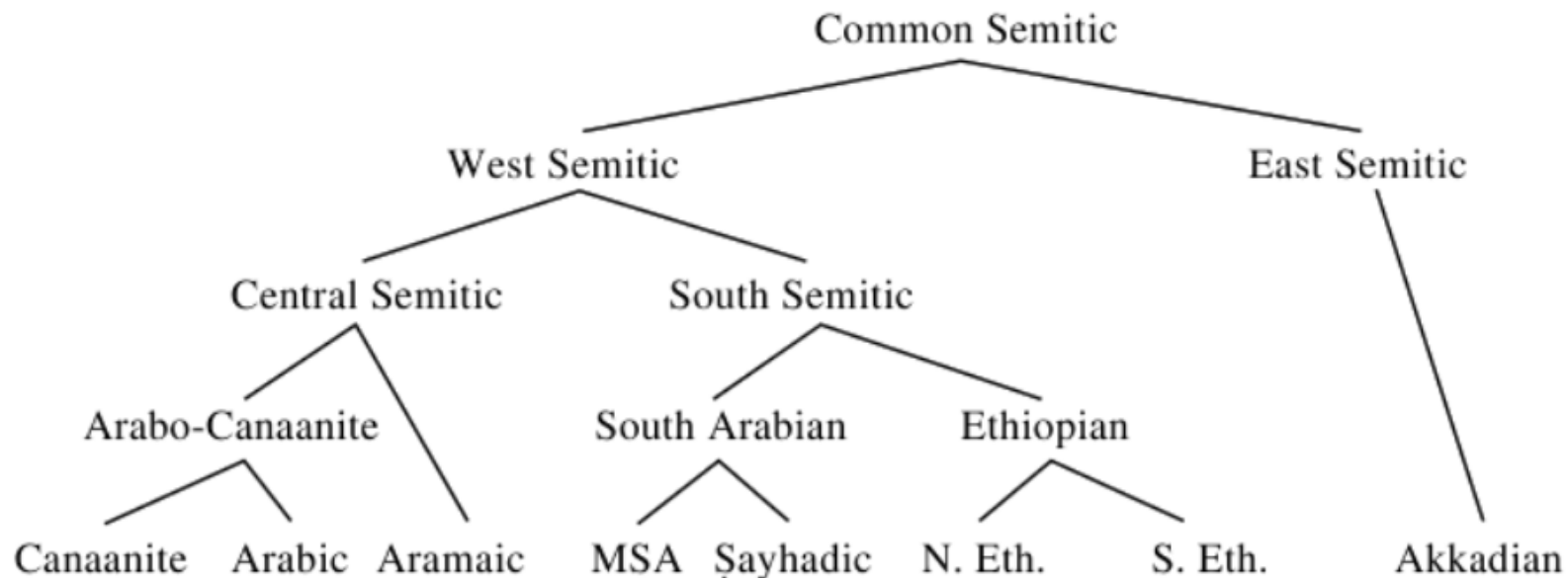


Figure 9.2: R. Hetzron's Model Classification Scheme (Hetzron 1976, etc.)

J. Huehnergard and A. Rubin. Phyla and Waves: Models of classification of the Semitic languages. In S. Weninger (ed.), *Semitic Languages: An International Handbook* (de Gruyter Mouton, 2011), pp. 259–278.

Arab helye a nyugati sémi nyelvek között

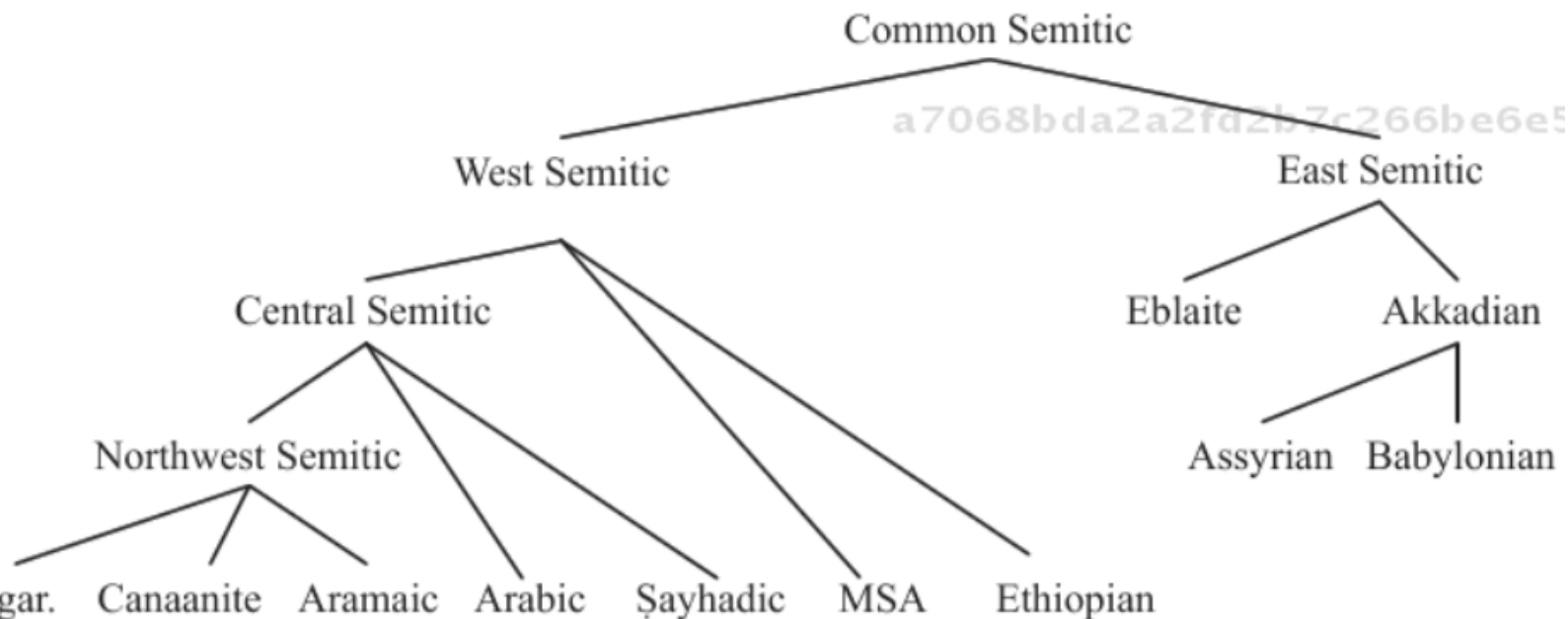


Figure 9.3: Modifications to Hetzron's Model

J. Huehnergard and A. Rubin. Phyla and Waves: Models of classification of the Semitic languages. In S. Weninger (ed.), *Semitic Languages: An International Handbook* (de Gruyter Mouton, 2011), pp. 259–278.

Hetzron két elve

Robert Hetzron: Two principles of genetic reconstruction. *Lingua* 38.2 (1976): 89-108.

Principle of archaic heterogeneity: „when cognate systems (i.e. paradigms) in related languages are compared, the system that exhibits the most inner heterogeneity is likely to be the closest to the ancestor-system. ” (P. 89.)

Principle of shared morpholexical innovations:

Hetzron két elve

Robert Hetzron: Two principles of genetic reconstruction. *Lingua* 38.2 (1976): 89-108.

Principle of archaic heterogeneity: „unless one can find a clear conditioning factor for differentiation – the relatively most heterogeneous system might be considered the most archaic, the closest to the ancestor, and the more homogeneous ones might be assumed to have arisen as a result of simplification. Furthermore, if the more homogeneous related systems, which exhibit more internal consistency, differ from one another, the chances of the heterogeneous one to be more archaic are even better.” (P. 93.)

Principle of shared morpholexical innovations:

Hetzron két elve

Robert Hetzron: Two principles of genetic reconstruction. *Lingua* 38.2 (1976): 89-108.

Principle of archaic heterogeneity:

Példa: szuffixkonjugáció (p. 93.)

	Akkád	Arab	Geez
Sg. 1c.	-(ā)ku	-tu	-ku
Sg. 2m.	-(ā)ta	-ta	-ka

Principle of shared morpholexical innovations:

Hetzron két elve

Robert Hetzron: Two principles of genetic reconstruction. *Lingua* 38.2 (1976): 89-108.

Principle of archaic heterogeneity:

Principle of shared morpholexical innovations:

„The most arbitrary elements of language are the phonetic shape of morphological and lexical items (the requirement of arbitrariness safeguards against possible developments due to general tendencies), and the phonetic shape of morphological items is the least likely to be borrowed (as against lexical items).”

(P. 89.)

Hetzron két elve

Robert Hetzron: Two principles of genetic reconstruction. *Lingua* 38.2 (1976): 89-108.

Principle of archaic heterogeneity:

Principle of shared morpholexical innovations:

Példa: prefixkonjugáció (p. 103.)

	Akkád	Geez	Arámi	Héber	Arab
Pl. 2/3f.	-ā	-ā	-ān	-nā	-na

Hetzron családfája

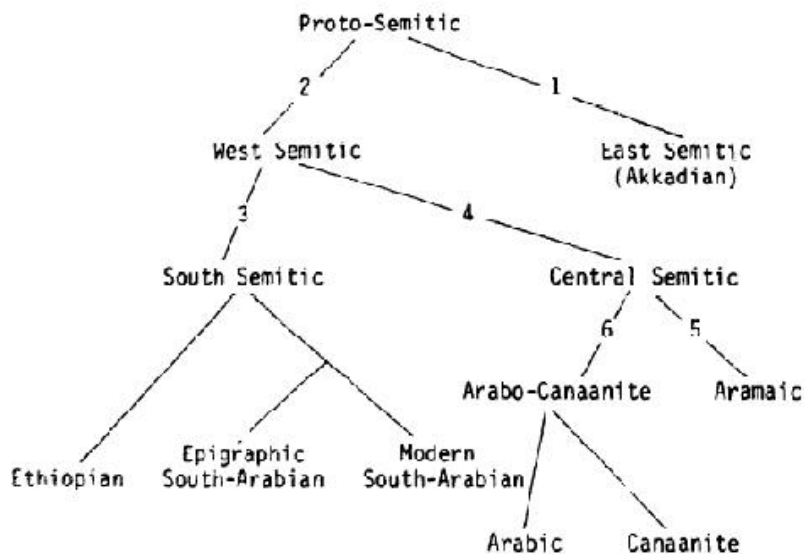


Fig. 2. The genealogical tree of Semitic. 1: Akkadian – archaic in many respects. 2: West-Semitic – innovation: adoption of a suffix-conjugation for the past tense (4.2.2). 3: South-Semitic – innovation: Generalization of *-k-* in the past tense (2.1), generalization of a closed vowel in verbal prefixes (2.2). 4: Central Semitic – innovation: adoption of Jussive *+u* for the nonpast (4.2.3), generalization of *-t-* in the past tense (2.1), generalization of the same vowel for verbal prefixes in one verb (2.2). 5: Aramaic – archaic feminine plural (4.1). 6: Arabo-Canaanite – innovation: adoption of *-na* as a feminine plural ending in verbs (4.1).

A legvalószínű narratíva

- $\mathcal{D} = \{d_1, d_2, \dots\}$: megfigyelt adatok.
 \mathcal{H}_1 : első/egyik hipotézis (esetünkben: narratíva)
 \mathcal{H}_2 : második/másik hipotézis (esetünkben: narratíva)
- Adott adatok mellett, melyik hipotézis valószínűsége nagyobb?
 $P(\mathcal{H}_1|\mathcal{D})$ vagy $P(\mathcal{H}_2|\mathcal{D})$?
- De: *Mit jelent / hogyan határozható meg „hipotézis valószínűsége, adott adatok mellett”?* Fordítsuk meg:
- Melyik hipotézis mellett nagyobb az adatok valószínűsége?
 $P(\mathcal{D}|\mathcal{H}_1)$ vagy $P(\mathcal{D}|\mathcal{H}_2)$?
- [+]
Melyik hipotézis *a priori* valószínűsége nagyobb? $P(\mathcal{H}_1)$ vagy $P(\mathcal{H}_2)$?

Lexiko-statisztika [és dialektometria]

- Adva nyelvek [dialektusok] egy csoportja (pl. sémi nyelvek).
- Minden nyelvpárra [dialektuspárra] számítsuk hasonlóságukat:
 - Hány rokon szó van?
 - Rokon szavak hangtani hasonlósága?
 - Morfológiai, szintaktikai, stb. hasonlóság?
- Statisztikai módszerekkel klaszterekbe rendezzük a nyelveket [dialektusokat].

Lexiko-statisztika [és dialektometria]

- **Swadesh-lista:** az *alapszókincs*ből vett 120 / 200 / stb. elemű minta. (Pl. legfontosabb testrészek, alapvető természeti jelenségek, névmások, kérdőszók, alapvető életfunkciókat és mozgást kifejező igék: ld. Bennett, p. 40.)
- A szókincs összefüggése a természeti és kulturális környezettel (hol és milyen típusú társadalomban éltek az őseink?).

Glottokronológia

- Radioaktív bomlás: konstans annak valószínűsége, hogy adott atom a következő másodpercben / napban / évezredben elbomlik.
- Exponenciális bomlás: $m(t) = m_0 \cdot e^{-C \cdot t} = m_0 \cdot 2^{-T_{1/2} \cdot t}$.
- Glottokronológia feltételezése: a szókincs egy elemének a lecserélése szintén konstans valószínűségű.
- Ezért az ősnyelvből megőrzött szavak száma (a Swadesh-listán) időben exponenciálisan csökken.
- Ekkor, ha a $T_{1/2}$ felezési időt tudjuk, akkor két nyelv közös eredetű szavainak számából a szétvállás ideje kiszámolható.

Filogenetikai módszerek

- Evolúciós családfák a biológiában.
- Elágazás: mutáció valószínűsége.
- A $\mathcal{D} = \{d_1, d_2, \dots\}$ megfigyelt adatokra különböző családfák $(\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_i, \dots)$ illeszthetők.
- Matematikai modellek: mi az adatok valószínűsége adott családfa mellett a családfa által feltételezett mutációk számából, stb.?
- Keressük a legvalószínűbb [„legtakarékosabb”] családfát.
- A biológiai evolúció analógiája a nyelvészetben, mint a családfamodell legmatematizáltabb verziója.

Viszlát holnap!

