

Lesson 1: Introduction to Albanology¹

1. The Albanian speaking area and the Illyrian question

1.1. Where is Albanian (= Alb.) spoken today (Pic. 1)?

1.1.1. Albania and Kosovo (official language).

1.1.2. Serbia, Macedonia, Greece, Italy, Ukraine, etc. (minority language).

1.1.3. Alb. spoken in Greece → **Arvanitic**; Alb. spoken in (Southern) Italy → **Arbëresh**.

1.2. In the 1st mill. BCE, today Albania was part of the region inhabited by the **Illyrians** (Pic. 2), on whose language little is known (Lippert & Matzinger 2022).

1.3. For political reasons, in the 20th cent. Albania developed a state doctrine, according to which Alb. is a direct descendant of Illyrian.

1.4. This can be disproved through linguistic evidence (Matzinger 2009; Schumacher 2009).

1.4.1. We know that PIE $*sk/\hat{k}$ became ($*hk > *kh > *k^h >$) h in Alb., cf. *hije* f. ‘shadow’ ~ Gk. *σκιᾶ* f. ‘id.’ (DPEWA s.v.).

1.4.2. However, Pre-Roman—i.e., Illyrian—toponyms in today Albanian exhibit an (initial) sequence *sk*: e.g., Ill. *Skodra* → Gk. *Σκόδρα* → Alb. *Shkodra* (city in Northern Albania, Pic. 3).

1.4.3. This shows that *sk* was preserved in Illyrian and did not become *h*: accordingly, Illyrian cannot be the direct ancestor of Albanian.

1.4.4. Toponyms like *Skodra* / *Σκόδρα* must have entered Alb. at a relatively late stage, when the change $*sk/\hat{k} > h$ had already been completed.

1.5. At the time when the Illyrians inhabited the Adriatic coast of today Albania, speakers of Pre-Albanian must have lived elsewhere: in the mountainous **inland of the Balkans** (§3.2.1.3 below).

¹ Cf. Sh. Demiraj 1998; Fiedler 2003; Matzinger 2006, esp. 1-46.

2. Albanian dialectology

2.1. Alb. is divided into two main dialects: **Gheg** (= Gh.) in the North vs. **Tosk** (= To.) in the South.

- Linguistic border approximately coincides with the course of the river Shkumbin (Pic. 1).

2.2. Standard Alb. is based on Tosk for political reasons (dictator Enver Hoxha, 1908-1985).

2.2.1. Arvanitic (Greece) and Arbëresh (Southern Italy) also belongs to Tosk.

2.2.2. Kosovo Alb. belongs to Gheg.

2.3. Some distinctive isoglosses between Gheg and Tosk.

2.3.1. Phonological isoglosses.

2.3.1.1. **Tosk rhotacism:** Proto-Alb. (= PALb.) $*\hat{V}nV > \text{Gh. } \hat{V}nV \sim \text{To. } VrV$ ($\langle r \rangle = [\text{ɾ}]$ vs. $\langle rr \rangle = [\text{r}]$).

- (PIE $*\mu\acute{o}i(h_1)\text{-}n\text{-}eh_2\text{-}$) PALb. $*\mu\acute{e}na > \text{Gh. } v\acute{e}n\acute{e} \sim \text{To. } (= \text{St.-Alb.}) v\acute{e}r\acute{e}$ f. 'wine' ($\langle \acute{e} \rangle = [\text{ə}]$).

2.3.1.2. **Tosk denasalization of vowels:** PALb. $*\hat{V}n > \text{Gh. } \hat{V}n \sim \text{To. } Vr$ ($\hat{V}$ = nasal vowel).

- (Lat. *Diāna* →) PALb. $*D^{\text{̃}}\acute{a}na > \text{Gh. } z\acute{a}n\acute{e} \sim \text{To. } z\acute{e}r\acute{e}$ f. 'fairy' ($\langle z \rangle = [\text{z}]$).

2.3.2. Morphological isoglosses.

2.3.2.1. **Infinitive:**

- Present in Gheg, with structure [*me* 'with' + ptc.]: e.g., *ai fillon me punue* 'he begins to work', lit. 'with worked'.
- Absent in Tosk (and St.-Alb.), where it is replaced by the subjunctive: *ai fillon të punojë* 'he begins to work', lit. 'that [he] works'—typical feature of the Balkan sprachbund.

2.3.2.2. **Future:**

- "Romance" construction [HAVE + inf.] in Gheg: e.g., *kam me punue* 'I will work', lit. 'I have to work'.
- "Balkan" construction [WANT + subj.pres.] in Tosk (and St.-Alb., but not in Arbëresh [§4.5.5 in Less. 4]): *dua të punoj* 'I will work', lit. 'I want that [I] work'—cf. Mod.-Gk. $\theta\acute{\alpha}\gamma\rho\acute{\alpha}\psi\omega$ 'I will write' ($< \theta\acute{\epsilon}[\lambda\omega\text{'}\acute{\iota}\nu]\alpha\gamma\rho\acute{\alpha}\psi\omega$ lett. 'I want that [I] write').

3. The lexicon of Albanian (cf. Matzinger 2017)

- Multi-layered, reflecting the different (pre)historic phases of the language.

3.1. **Inherited words:** directly inherited from PIE.

3.1.1. Alb. *derë* f. 'door' < PIE $*d^h\mu\acute{o}r\text{-}$, cf. PDE *door*, Gk. $\theta\acute{\upsilon}\rho\alpha$ f. 'door', Lat. *forēs* pl.f. 'id.', Ved. *dvārā* du.f. 'id.', Russ. *dver'* f. 'id.', Toch.B *twere* 'id.', etc. (DPEWA s.v.).

- 3.1.2. Alb. *bie* ‘bring’ < PIE **b^her-*, cf. Lat. *ferō* ‘bring’, Gk. *φέρω* ‘id.’, Ved. *bhárati* ‘brings’, NHG *gebären* ‘give birth’, etc. (DPEWA s.v.).
- 3.1.3. Often mono- or disyllabic words (phonetically eroded).
- 3.2. **Greek loanwords** (Thumb 1909; Ölberg 1972; di Giovine 1988, 1993, 2021, 2025; Spaans 2024).
- 3.2.1. **Loanwords from Ancient Greek.**
- 3.2.1.1. They spread from the Greek (Doric) colony of Ἐπίδαμνος or Δυρράχιον (Lat. *Dyrrachium*, Alb. *Durrës*, It. *Durazzo*, SCr. *Drač*—Pic. 4), starting from the 7th cent. BCE.
- 3.2.1.2. They are few (max. 20) and mostly restricted to the semantic fields of agricultural products and artifacts:
- Alb. *lakër* (Gh. *lakën*) f. ‘cabbage’ ← Gk. *λάχανον* (esp. pl. *λάχανα*) n. ‘vegetables’ (<χ> = [k^h]).
 - Alb. *shtyllë* f. (OAlb. m.) ‘column’ (<ll> = [ɫ]) ← Gk. *στῦλος* m. ‘id.’ (<ũ> = [uː]).
- 3.2.1.3. Pre-Alb. speakers lived in the Balkan inland (§1.5) and were therefore not in direct contact with the Greek settlements on the Adriatic coast (Pic. 5).
- 3.2.2. **Loanwords from Byzantine and Modern Greek.**
- 3.2.2.1. They are more numerous than the loanwords from Ancient Greek.
- 3.2.2.2. They entered Albanian after Alb. speakers migrated to their historical settlements (i.e., today Albania) starting from the Late Antiquity (5th/6th cent. CE).
- E.g., Alb. *zavë* f. ‘armor’ ← Biz.-Gk. *ζάβα* f. ‘id.’ (<β> = [v]).
- 3.3. **Latin loanwords** (Haarman 1972; Landi 1989; di Giovine 1993, 2012; Bonnet 1998).
- 3.3.1. Much more numerous than the (Ancient) Greek loanwords and occurring in virtually all semantic fields / word classes:
- Body parts: Alb. *kofshë* f. ‘thigh’ ← Lat. *coxa* f. ‘id.’.
 - Kinship terms: Alb. *emtë* f. ‘aunt’ ← Lat. *amita* f. ‘(paternal) aunt’.
 - Everyday lexicon: Alb. *bukë* f. ‘bread’ ← Lat. *bucca* f. ‘cheek, mouth’.
 - Adjectives / verbs / adverbs: Alb. *i aftë* ‘suitable’ ← Lat. *aptus* ‘id.’ / Alb. *kujton* ‘think’ ← Lat. *cōgitō*, *-āre* ‘id.’ / Alb. *shum* ‘very’ ← Lat. *summus* ‘highest’, etc.
- 3.3.2. Roman colonization of the Balkans (2nd cent. BCE - 4th/5th cent. CE) was much deeper and capillary than the Greek one.
- Romans built fortified cities (*castra*) in the Balkan inland, cf. Sirmio, Singidunum (Belgrad), Naissus (Niš), Scupi (Skopje), etc.
 - Romans built streets, esp. the *Via Egnatia* (Durrës - Thessaloniki - Byzantium, Pic. 6).

3.3.3. Despite the intensive contacts with Latin, Alb. did not become a Romance language!

3.4. Slavic loanwords (Svane 1992; Ylli 1997).

3.4.1. Starting from the 6th/7th cent. CE, Slavic tribes invaded the Balkans and got in contact with the speakers of Alb. → common lifestyle based on agriculture and pastoralism.

3.4.2. Many Slavic loanwords in Alb. belong to the semantic fields of pastoral life and nature / landscape:

➤ Alb. *kosë* f. 'sickle' ← Serb. *kòsa* ~ Mac. *kosá* f. 'id.' (cf. Russ. *kosá* 'id.', Pol. *kosa* 'id.', etc.).

➤ Alb. *breg* 'shore' ← Serb. *brêg* ~ Mac. *breg* m. 'id.' (cf. Russ. *béreg* m. 'id.').

3.4.3. Slavic donor languages are usually (Old) Serbian and/or (Old) Macedonian.

3.5. Italian loanwords (di Giovine 2008).

3.5.1. Loanwords from (Old) Venetian, between 12th and 18th cent. CE:

➤ Alb. *lexon* 'read' (<x> = [d̥z]) ← Venet. *lèzer* 'id.' (cf. It. *leggere* 'id.').

➤ Alb. *karrige* f. 'chair' ← Venet. *cariega* 'id.' (cf. Mod.-Gk. *καρέκλα* 'id.').

3.5.2. Loanwords from Modern Italian (through television), in the second half of the 20th cent.:

➤ Alb. *mafioz* m. 'criminal' ← It. *mafioso* 'id.' (lit. 'member of the Mafia').

➤ Alb. *lavatriçe* f. 'wash machine' (<ç> = [tʃ]) ← It. *lavatrice* 'id.'.

3.6. Turkish loanwords (Boretzky 1975; Friedman & Joseph 2025).

3.6.1. 1479-1912: Albania (like most other Balkan regions) was part of the Ottoman Empire (Pic. 7).

3.6.2. Turkisms are widespread in other Balkan languages too (Mod.-Gk., Serb., Mac., Rum.) and occur in virtually all semantic fields / word classes:

➤ Culture / entertainment: Alb. *pi* 'drink' ~ 'smoke', calque on Turk. *içmek* 'id.'.

➤ Onomastics: Alb. *Hoxha* (surname, <xh> = [d̥ʒ]) ← Turk. *hoca* '(Muslim) teacher'.

➤ Function words: Alb. *gjoja* (epistemic particle expressing doubt) ← Turk. *göya* (id.).

➤ Denigratory words: Alb. *budalla* 'stupid' ← Turk. *budala* 'id.'.

3.6.3. Many Turkish loanwords were abandoned for political reasons after the independence of Albania from the Ottoman Empire (1912).

➤ E.g., Alb. *penxhere* f. 'window' (← Turk. *pencere* 'id.') replaced by *dritare* (← *dritë* 'light').

4. The oldest attestations of Albanian

4.1. Pre-literary attestations, in Old Gheg (15th cent. CE).

4.1.1. **Baptism Formula** (*Formula e pagëzimit*), in a pastoral letter written in Latin by the archbishop of Durrës Paulus Angelus (Pal Engjëlli, <gj> = [j]) in 1462:

<Vnte' paghesont premenit Atit et Birit et Spertit Senit>

= /Un të pagëzonj pr'emënit Atit e t'Birit e t'Shpertit Shenjt/²

"I baptize you in the name of the Father and the Son and the Holy Spirit".

4.1.2. **Arnold von Harff's lexicon**: list of about 40 Alb. words with (Middle Low) German translation, compiled by the Cologne Knight (*Ritter*) Arnold von Harff during his brief stay in Durrës in 1497 (Hetzer 1981; Elsie 1984).

➤ E.g., <vene wijn> = /vënë 'Wein [wine]'/.

4.2. Literary attestations (from 16th cent. CE onward)—see, generally, LHGjSh.

4.2.1. Old Gheg.

4.2.1.1. **Gjon Buzuku: *Missale* / *Mishari* (1555)**—translation of Biblical passages and liturgical texts from Latin (Çabej 1968; Qendro 2013).

4.2.1.2. **Pjetër Budi**: inter alia, *Dottrina Christiana* / *Doktrina e krishterë* (1618)—translation of the *Piccolo catechismo* (*Small Catechism*) by Roberto Bellarmino (Demiraj, Omari & Kapia 2022).

4.2.1.3. **Pjetër Bogdani** (archbishop of Shkodra and Skopje): *Cuneus Prophetarum* (1685)—considered the father of Alb. literary prose, first author of a non-translated work in Alb. (Omari 2015; 2016).

4.2.2. Old Tosk.

4.2.2.1. **Lekë Matrënga: *Dottrina cristiana* / *Mbsuame e krështerë* (1592)**—more on this in Lesson 4.

4.2.2.2. **Jul Varibóba** (from San Giorgio Albanese, Calabria): *Vita della Santa Vergine Maria* / *Gjella e Shën Mëris Virgjër* (1762)—poetic composition (Fortino 1984, with Italian translation).

5. The ethnonym(s) of the Albanians

5.1. Today, the Alb. endoethnonym relies on the adj./adv. *shqip* 'Albanian, in Albanian (language)', cf. *shqip-tar* m. 'Albanian (person)', *Shqip-ëri* (Gh. *Shqip-ni*) f. 'Albania', etc. (<q> = [c]).

5.1.1. *Shqip* '(in) Albanian' probably originated from a semantic specialization of the adj./adv. *shqip* 'clear, comprehensible' (Matzinger 2013).

5.1.2. Cf. NHG *deutsch* 'German' (< OHG *thiut-isk* 'id.') ~ *deut-lich* 'clear, comprehensible'.

5.2. The older endoethnonym of the Albanians relied on a base *AlbVn-* (~ *ArbVn-*).

² To make Old Albanian forms more easily accessible to the reader, it is good practice not to present them in their original orthography (transliteration), but in semiphonological transcription, noted as /.../. See VAA:199-203.

- 5.2.1. Ptolemy (2nd cent. CE) mentions the tribe of the Ἀλβανοί as well as their capital city Ἀλβανόπολις (geographic localization uncertain).
- From the Classical sources, this ethnonym entered most Western European languages, cf. the exoethnonyms PDE *Alban-ian*, NHG *Alban-isch*, It. *Alban-ese*, etc.
- 5.2.2. The base *AlbVn-* (~ *ArbVn-*) is still used as endoethnonym by the diaspora communities of Greece (Ἀρβανίτες) and Southern Italy (*Arbër-esh*).
- *Terminus post quem*: the ethnonymic replacement *AlbVn-* → *shqip* must have happened after the emigration of Tosk speaking people to Southern Italy, i.e., after the 15th cent. CE.
- 5.2.3. When Slavs got in contact with Alb. speakers in today Southern Albania, they borrowed from them the denomination of that region as **Albanía*.
- This underwent South Slavic liquid metathesis and was then reborrowed from the Albanians as **Labanía* > (Tosk rhotacism) **Labaría* > (vowel weakening) *Labëria* (Pic. 8).