Prehistoric loanwords in Armenian: Hurro-Urartian, Kartvelian, and the unclassified substrate

Nielsen, R.T.

Evaluation and outlook 5

The primary purpose of this work is the critical evaluation and delimitation of three loanword corpora in Armenian, each representing distinct linguistic contact events in the prehistory of this language. The secondary purpose is to determine to what extent this data may inform our knowledge about the prehistory of Armenian speaking populations, in particular their movements and the timing of those. The linguistic data demonstrates the relative sequence of contact events. Subsequently, informed guesses can be made with regard to the absolute dating of these events, as well as the geographic location of Armenian at the time.

The youngest of these prehistoric events is the contact between Armenian and Urartian. It took place before the introduction of Iranian loanwords but after most sound changes, including the Armenian sound shift, had taken place. This linguistic observation is consistent with the assumption that the Urartian loanwords are all contemporaneous with the existence of the Urartian Kingdom from ca. 860–590 BCE.

The study of the contact between Armenian and Kartvelian languages presents a complex and multifaceted picture. Contact with the Zan languages stretches up until the historical period but appears to have begun already while these languages were beginning to diverge from their closest predecessor, Georgian-

Zan. A single lexical item, Arm. *cov* 'sea', suggests that contact may have taken place before the Armenian sound shift, but the lack of parallels precludes a firm conclusion. With regard to the relative dating, the Armenian influence upon Kartvelian languages provides stronger evidence (see also Thorsø 2022). The Armenian loanwords into Georgian and Georgian-Zan must have taken place well before the adoption of Urartian loanwords, probably already in the latter half of the second millennium BCE. Unless we assume that Kartvelian languages were, at this time, spoken far from their historically attested location, it suggests that already in the second millennium, Proto-Armenian was spoken in the Southern Caucasus.

This conclusion casts considerable doubt on the traditionally favoured hypothesis of how Armenian was introduced to its historical area. This hypothesis states that Armenian speakers migrated from the Balkans and into Eastern Anatolia only after the collapse of the Hittite Empire around 1200 BCE (Tomaschek 1893: 4, Brandenstein 1961, Diakonoff 1964, Burney & Lang 1971, Mallory 1989: 33–5, Fortson 2010: 382). Fundamentally, the Balkan Hypothesis relies on statements of ancient historians like Herodotus[1] and Eudoxus[2] that Armenians were (closely related to) Phrygians or had come from Phrygia. The Balkan Hypothesis also helps explain why there is no historical record of an Armenian nation or ethnos before the sixth century BCE. On the other hand, if Armenian speakers were present close to Kartvelian speakers already in the second millennium, it appears doubtful that they migrated across Anatolia, since these migrations would have been recorded in Hittite, Luwian, or Assyrian sources. In any case, the material evidence for a migration of Indo-European speaking people from somewhere in the Balkans or Western Anatolia in the Early Iron Age is virtually non-existent. As for the evidence of linguistic phylogeny, the relationship between Armenian and Phrygian appears much more distant than it did to many scholars of the early twentieth century, and there is now broad consensus that the closest relative of Phrygian is not Armenian, but Greek (Obrador-Cursach 2019).

[1] *Histories* 7.73: Ἀρμένιοι δὲ κατά περ Φρύγες ἐσεσάχατο, ἐόντες Φρυγῶν ἄποικοι (The Armenians were armed like the Phrygians, being Phrygian colonists).

[2] Attested only in Stephanus Byzantius, *Ethnica*, s.v. Armenia: Ἀρμένιοι δὲ τὸ γένος ἐκ Φρυγίας καὶ τῇ φωνῇ πολλὰ φρυγίζουσιν (As for their origin, the Armenians are from Phrygia and they speak much like Phrygians).

At the same time, it may be justified to assume that Armenian was spoken north of Urartu, when the latter emerged as a local power. The events of the Southern Caucasus were virtually undocumented at this time, but Urartian and Assyrian sources tell of rival confederations here, such as the Etiuni, with whom they were in frequent conflict. This tribe or 'kingdom' may even have played a decisive role in the eventual downfall of Urartu (see Petrosyan 2018: 158–65 for an overview). More importantly, a movement of plausibly Indo-European-speaking people from the Pontic-Caspian steppe and into this area can be documented already from the Middle Bronze Age, with the emergence of the Trialeti-Vanajor culture (ca. 2100–1700) BCE. This event was a dramatic transition from the sedentary, agricultural, and largely egalitarian Kura-Araxes culture to a nomadic, pastoralist, and socially stratified economy (Sagona 2017: 309–13, Drews 2017: 89–92, Kristiansen 2018: 113–5). This is also the period when the Armenian Highlands see the emergence of the *višapakᶜarer* 'dragon stones' (Barseghian 1968). These curious, zoomorphic stone stelae may be interpreted in the context of a cultic ritual with clear Indo-European elements (Martirosyan 2015). Their connection with the Trialeti seems highly likely but is yet to be established. In any case, the introduction of Trialeti-Vanajor can only be seen as a major social turnover which serves as a plausible staging area for language contact and language shift. From the point of view of ancient DNA, Lazaridis et al. (2022b) are able to demonstrate an admixture of approximately fifteen per cent ancestry associated with the Yamnaya culture of the Pontic-Caspian steppe at this point in time. From around 1500 BCE, the Trialeti-Vanajor culture is gradually replaced by the similar Lčašen-Mecamor culture, whose territory around Lake Sevan plausibly overlaps with that of the aforementioned Etiuni (Diakonoff 1964: 7, Avetisyan et al. 2019). We are thus able to glimpse a more or less direct line from people living at the outskirts of the Urartian empire in the early first millennium BCE to the Yamnaya culture of the third millennium BCE, whose people were most plausibly speaking Indo-European languages (Schrader 1883, Mallory 1989, Anthony 2007). On the basis of linguistic data, this route through the Caucasus around 2000 BCE is the most likely vector for the introduction of Armenian into its historical area.

The third layer of loanwords evaluated in this work clearly represents the oldest. It testifies to a prolonged contact between Armenian and one or more unclassified, non-IE languages. This

contact event predated all or most Armenian sound changes. Crucially, most other Indo-European languages, with the exception of Anatolian and Tocharian, were to some degree in contact with the same stratum. These facts, taken together, suggest that this period of language contact must have begun relatively shortly after the dissolution of the Core Indo-European languages. Therefore, it most likely represents contact between speakers still residing near the Indo-European homeland, and speakers of those languages neighbouring them. It seems clear that Armenian, Greek, and Albanian remained in close contact with the same language(s) for the longest period of time. This is consistent with the data showing that these languages shared innovations on the basis of inherited material as well (Matzinger 2012, Lamberterie 2013, Olsen & Thorsø 2022). At the same time, there is also a considerable overlap between non-Indo-European vocabulary in Armenian and that found in Germanic, Italic and Celtic. Among these loanwords are terms for agricultural crops, like 'barley' and 'some pulse' (> Arm. 'alfalfa'), indicating that Proto-Armenian existed within the core of Indo-European languages whose speakers migrated Westward across the steppe and went through a gradual transition from a completely herding-based economy to a more sedentary culture with elements of agriculture, starting from around 3300 BCE (cf. Kroonen et al. 2022). Nevertheless, Armenian does not share as much foreign agricultural vocabulary with Germanic, Italic, and Celtic as these languages do with one another. Thus, there is reason to believe that its speakers did not take part in those population movements that later gave rise to the Corded Ware and Bell Beaker cultures in Europe. Again, given that population movements around 2000 BCE are a plausible vector for the movement of Proto-Armenian speakers into the Caucasus, it is tempting to preliminarily locate these Proto-Armenian speakers somewhere in the Late Yamnaya and perhaps in the Catacomb culture, which emerges from Yamnaya starting around 2500 BCE. Future studies combining linguistic, archaeological, and genetic evidence will hopefully be able to confirm or reject this hypothesis.

With respect to its origin, the Armenian lexicon is highly variegated and complex, and a large part of it remains obscure. While this work has hopefully advanced the understanding of this lexicon, it should also serve to accentuate the need for much more work within the field of loanwords. Obviously, future studies need not

only follow the same roads that this study has taken. There are many other potential foreign sources of Armenian words. These might include 'substrate words' shared with Anatolian languages and non-IE languages of Western Asia; direct loanwords from Nakh-Daghestanian and Abkhaz-Adyge languages, as well as words of completely unknown origin, which may however still be classified according to formal and semantic criteria. Obviously, many inherited words may still be uncovered as well. With the advancement of these studies, Armenian and Indo-European studies will surely see advancement as a whole.