George Balabanian

Masters Thesis
University of Ottawa

A Co-phonological Approach to Persian Loanwords within Armenian Phonology

Introduction

Different co-phonological approaches have been fruitful in describing phonotactic patterns in cases of prolonged language contact and heavy borrowing, such as for native Japanese stock vis-à-vis Sino-Japanese vocabulary (Kurisu, 2001), or native Persian stock vis-à-vis Perso-Arabic vocabulary. For the purposes of this paper, the research question to be examined if we can propose a set of nested co-phonologies (a non-derivational morphologically-conditioned phonology developed within OT chiefly by Orgun 1996, 1998 and Inkelas 1998) within Armenian grammar as a way to adequately and economically describe phonological differences between the directly inherited Indo-European lexicon, and the layers of Persian loanwords. Due to the sheer amount of time passed and high degree of nativization of these loanwords, the expected hypothesis here would be that the Persian loanwords have been so fully integrated as to be indistinguishable from the native Armenian lexicon.

Historical Overview

Much of the vocabulary of Armenian comes from Parthian, a Northwestern Middle Iranian language, testimony to the extent to which Armenia was permeated by the political and religious institutions of pre-Islamic Iran. This very large number of loanwords covers over a thousand separate lexical items not counting derivatives or compounded forms (Clackson, 2008). It is often said that the Iranian influence on the Armenian language is comparable to the influence of Norman French and Latin on English, but not only is the larger part of vocabulary of

administration, military life, and religion borrowed from Iranian, but also adjectives and prepositions and a number of adjectival[1] (such as *-agin*, 'like X, endowed with X', Korn and Olsen 2012), adverbial, and nominal suffixes. Even phrasal combinations of noun-and-verb and noun-and-noun are calqued from Persian (such as, respectively, *dandanawan*, from *dantan*, 'tooth', and *banda* 'to tie' 'birdle' and *šahanšah*[2] 'king of kings', from fully reduplicated *šah* 'king').

On the basis of generally accepted historical periods of Armenian history, we can divide these loanwords into three separate periods in which loanwords entered Armenian from Persian: (i) during the Urartean era (pre-6th century BCE., though perhaps many centuries farther back) and the Achaemenid era (6th to 4th century BCE.; mostly from Old Persian; much work has been done to this purpose ever since the publication of Hübschmann's fundamental *Armenische Grammatik*, but the inventory of Persian loans, though considerably furthered, still awaits completion (Godel, 1975), though Meyer 2017 has recently added to this inventory); (ii) during the Parthian period (*c*. 200 BCE to 400 CE), cultural and political contacts between the Armenians and Persians were closest, and there was a large influx of words from Parthian including common terms such as *mah* 'death', *ašxarh* 'land', *šat* 'very', *seaw* 'black' and *spitak* 'white' (Clackson 1994, 2008). (iii) in the later Sasanian period, contact was much less close and loanwords from this period are not well integrated into the Armenian lexicon (Clackson, 2008: 142). Though Clackson does not expound upon this last point, we can *a priori* surmise that the third period's loanwords are contrastive to some degree compared to the two others by its different level of integration.

The bulk of earlier work on Persian loanwords within Armenian predate modern phonology, hence the application of a co-phonological approach to this language is novel. With this approach, we can use the idea of 'Markedness Reversal', where a markedness constraint can be re-ranked in different phonological constructions in the same language to account for apparent irregularities which are lexically derived. Much like Itô & Mester (1999, 2003) propose that

1 The same is true of for English with its numerous Normal French-derived adjectival suffixes and prefixes, such as *-al*, *-ial*, *-orial*, *-ual*, *-ment*, *-ty*, *-ion*, *re-*, *de-*, and others.

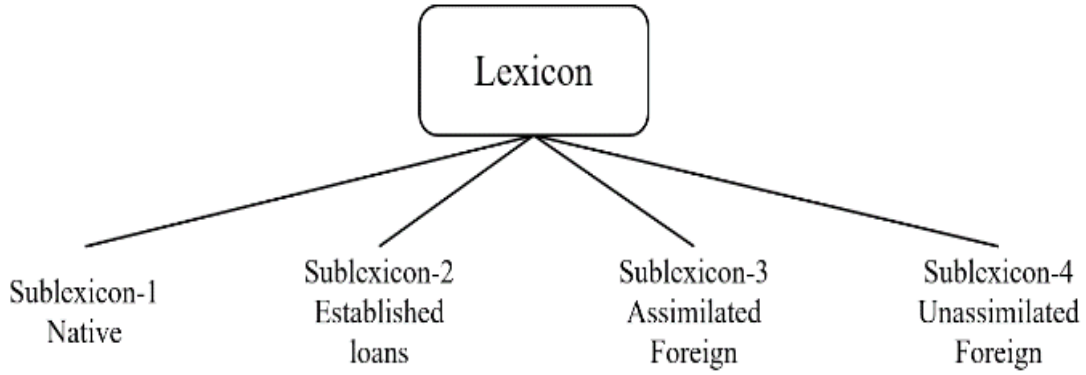
2 The spelling convention used here is the Hübschmann-Meillet (1913) Classical Armenian transliteration, which is the standard way of transliterating Armenian in the historical linguistics literature.

Japanese has strata that have a Core-Periphery relationship such that in the core stratum, all the relevant markedness constraints outrank faithfulness constraints, while in each successive stratum going towards the periphery specific faithfulness constraints are ranked above markedness, permitting more varied phonotactic patterns (Downing, 2008). It should be possible to show a similar relationship in Armenian by briefly focusing on phonotactics, reduplication, prothesis, and metathesis. The purpose of this mémoire is to verify that the Core-Periphery model of phonology and phonotactics applies to the Armenian lexicon, and if so, to what degree it is successful in capturing the differences between each purported lexical layer. Due to the strongly diachronic nature of the problem, a certain degree of framework inter-borrowing is used in this paper, and a brief quantitative discussion will be included.

Theoretical Frameworks

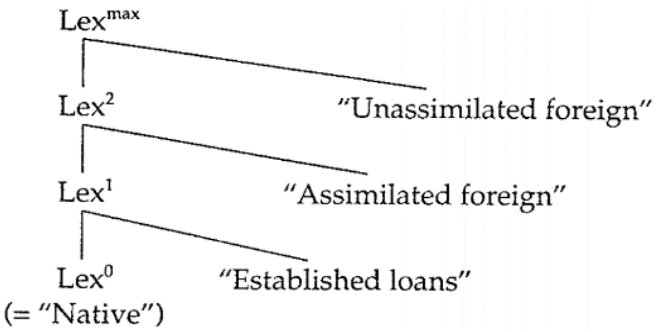
Itô and Mester's Core-Periphery model (1995) is a hierarchical arrangement of lexical strata, where "the relevant structural organization of the lexicon is set inclusion, leading from the innermost lexical core Lex_0 to the most inclusive set Lex_{max} comprising all lexical items" (Itô & Mester 2004: 553), as seen in Figure 1, which represents a non-stratified lexicon, where each sublexicon follows its own independent indexed constraints rather than cophonologies[3], in contradistinction to Figure 2:

(1) Simplified schema for lexical stratification assuming unrelated and separate co-phonologies



3 This important distinction here is that indexed constraints are organized into a single ranking, in other words, a single phonological grammar where certain constraints are indexed to morphemes from some sublexicon. For a co-phonology approach, there are actually several distinct phonological grammars (with potentially very different rankings, as we will see below with reduplication) that are called upon independently.

(2) Vertical representation of assimilatedness of lexical strata



(Source: Itô & Mester 2004: 64-65)

Itô and Mester's 1999 work refers to Kiparsky (1968) who argues that lexical items do not come neatly packaged into groups labeled either [+foreign] or [–foreign], and that we ought to stratify the lexicon of a language based on a hierarchy of foreignness.

The model is built out of an ordering or hierarchy of implicational relations: items that are subject to constraint A are always subject to constraint B, but not all items that are subject to B are subject to A. In this way, A would be a constraint with a more restricted domain than B (A's domain is properly included in B's domain) putatively in terms of phonotactics or morphophonological phenomena. To flesh this out for the Japanese case:

- for the Yamato (core) stratum, "both the occurrence of multiple voiced obstruents within a stem and sequences of a nasal followed by a voiceless obstruent are disallowed – i.e., both OCP-VOICE and *NÇ are enforced" (Hsu & Jesney, 2016); the *COMPLEX constraint indicates that only basic CV-type syllable structure is allowed;
- in the second, less-nativized Sino-Japanese stratum, "OCP-VOICE is still enforced, but sequences of a nasal followed by voiceless obstruent are admitted" (*ibid.*) hence the violation of NO-NT (same as *NÇ in the bullet point above);

- finally, among recent loanwords in the Foreign stratum (though it could be argued that Portuguese and Dutch[4] form an earlier stratum with slightly different rules than the 20th-21st century chiefly English layer), "violations of both OCP-Voice and *NC̥ are allowed [which] gives rise to two asymmetric implicational patterns – if a nasal + voiceless obstruent sequence is repaired in a stem, multiple voiced obstruents within the same stem will also be repaired. Likewise, if multiple voiced obstruents are permitted within a stem, nasal + voiceless obstruent sequences will also be permitted." (*ibid.*); singleton-p (NO-P constraint) is violated for both assimilated and unassimilated loanwords, whereas voiced obstruent geminates (NO-DD) only occur in unassimilated loanwords.

(3) Tableau representing violations in markedness constraints in Japanese co-phonologies

SYLLSTRUCTURE:	basic syllable structure constraints (e.g. *COMPLEX)
NOVOICEDGEMM (NO-DD):	no voiced obstruent geminates (*dd, *gg, etc.)
NOVOICESSLAB (NO-P):	no singleton-p: a constraint against nongeminate [p]
NONASAL VOICELESS (NO-NT):	postnasal obstruents must be voiced (*nt, *mp)[5]

	SYLLSTRUCTURE	NO-DD	NO-P	NO-NT
Yamato (core)	✓	✓	✓	✓
Sino-Japanese	✓	✓	✓	violated
Assimilated foreign words	✓	✓	violated	violated
Unassimilated foreign words (periphery)	✓	violated	violated	violated

(Ariyaee 2019, adapted from Itô & Mester 1999:73)

Japanese lends itself well to such an etymologically-based analysis[6], but Armenian is likely a tougher case because unlike the Sinitic layer in Japanese (Classical Chinese), which is extremely well-studied and comprehensively attested, Old and Middle Persian varieties suffer

4 See, for example, pp. 20, 29, 61, and 76 of Labrune (2012).

5 As Kevin McMullin pointed out, this is not a definition of a markedness constraint, but rather a description of a process (presumably resulting from ID-[voice] being lower ranked).

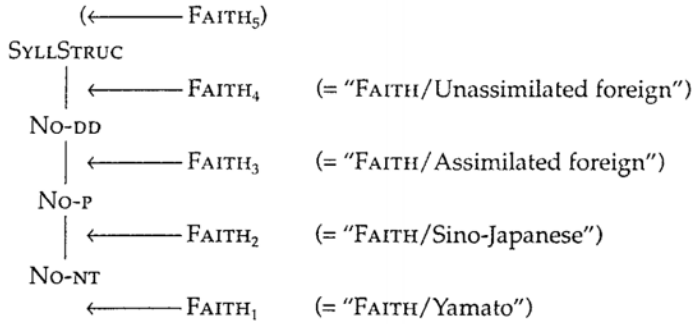
6 Other scholars (Legendre, Miyata & Smolensky 1990, Smolensky & Legendre 2006) have incorporated weighted scalar constraints as in Harmonic Grammar into Itô & Mester's model. "This approach allows implicational patterns to be captured without multiplying the set of Faithfulness and/or Markedness constraints, and without the need to impose metaconditions on the set of possible rankings […] this approach allows the analysis of implicational patterns seen in loanword adaptation to be united with the analysis of implicational patterns elsewhere in phonology." (Hsu & Jesney 2016).

from having smaller surviving corpora, which makes the task of the lexical reconstructionist more difficult. For example, take the Armenian word *nirh* 'dormancy, slumber' – on the basis of its appearance, the word is seen as a loan from an Iranian **niδrā*, but no such word is attested, yet philologists know that it exists in the Indo-Iranian subgroup at large because of Vedic Sanskirt. *nidrā́*- f. 'slumber, sleepiness' (Martirosyan 2013: 105).

In this stratification, the core lexical items (Lex_0) fulfill markedness constraints in its maximal way. And the more we move outward toward the outer layer from the core, we see the more violation of these markedness constraints. The constraint hierarchy (seen in the tableau below in (4)) clarifies the point that this lexical stratification transpires through different faithfulness constraint rankings within this fixed hierarchy of markedness constraints. FAITH_5 would be 19th and 20th century French and English loanwords, which can tolerate, for example, nasalized vowels or close-mid front rounded vowels such as in *Eugénie* [øʒɛni].

For the purpose of this paper, we can explain the implementation of Figure 3 and 4, by means of "moving faithfulness/marknedness constraints", though a full analysis using a scalar harmonic grammar (or another framework, see Itô & Mester, 1998 and Itô & Mester, 2001 for stratal faithfulness) is outside the scope of the research presented here.

(4) Vertical representation of faithfulness constraints



(Source: Itô & Mester 1999:73)

Thus, the stratum for the core Indo-European ("IE")-inherited lexicon of Armenian ranks below all the other strata's markedness constraints and cannot comply with the demands of the two or three (or more) markedness constraints[7]. The constraint INHERITED-IE-PHONOLOGY below is an umbrella term for all regularly (predictably) derived sound changes from Proto-IE to Armenian[8]. Loanwords derived from Persian necessarily follow another set of diachronic rules from IE to Proto-Iranian down to the various Persian languages, which is then modified when borrowed into Armenian.

---

7 Regarding level-ordering and empirical motivation for lexical strata and on the question of whether or not co-phonologies are extrinsically ordered, as is claimed in Lexical Phonology, other frameworks within morphology disagree – in "Sign-Based Morphology, level ordering is not the expected case […] (see also Inkelas and Orgun 1996), though it can be stipulated if necessary in any particular case" (Orgun 1996: 3).

8 For a brief explanation – when one compares Armenian to other Indo-European languages, one immediately notices the great variety of developments from IE voiceless stops (IE voiced stops generally became voiceless stops and IE aspirates become voiced stops, with the exceptions indicated in the chart below), as only **kˆ* appears to consistently yield Armenian *s*. Meillet, in his seminal work *Esquisse d'une grammaire comparée de l'arménien classique*, gives a brief outline of the issue, though according to Winter (1955) "his interpretation of the phonetic stages leading from former voiceless stops to the actually preserved sound is sometimes vague". Meillet held that IE voiceless stops first developed into voiceless aspirates, though no explanation is offered as to how the assumed aspirates came to yield *w* preconsonantally, or voiced stops after liquids and nasals. Even Pedersen (1905) assumes an aspirate stage, but prefers to think that the aspirates were afterwards replaced by spirants.

Winter (1954) was one of the first open opponents of this traditional view, as he directly challenges Meillet's opinion – "it is generally taken for granted that the present-day pronunciation of Arm. *p' t' k'* represents the original status: Meillet (1936, p. 23) simply states this view without discussion: 'Il y avait trois séries d'occlusives ... chacune d'elles existant sous forme de sourde non aspirée, de sourde aspirée (c'est-à-dire où l'explosion était suivie d'un souffle) et de sonore'." By analyzing loanwords from Greek, Syriac, and Persian, and specifically by looking over recordings of Armenian words in Greek script instead of vice-versa, Winter proposes that the IE **p *t *k* must be interpreted as representing allophones of Proto-Armenian /*p t k*/, and that there must have been a lengthy phase wherein these were actually fricatives.

Over the course of writing his trilogy of articles, Winter (1954, 1955, 1962) comes to the tentative conclusion that, much like the intermediary stages of Grimm's Law for the Germanic languages, Proto-Armenian must have gone through a phase wherein the IE **p *t *k* triplet had become [f θ x], at least word-initially and after resonants and semivowels, before becoming the more familiar *p' t' k'* series by the 5th century. IE **p *t *k* remain unchanged after s, yet the *p in this series is exceptional as it is likely to be deleted, such as in *otn* 'foot' < **pod-m̥* or became *h*, as in *howr* 'fire' < **peh₂ur̥* (cp. Greek *πῦρ*).

After a vowel or resonant, the Proto-Armenian allophones of the IE plain plosives are probably voiced spirants, according to Winter (1955). He claims that only this assumption can account for such apparently disparate developments as this series becoming *b d g* after resonants and semivowels and becoming *w y y* between vowels. At a relatively early date, but probably after the spirants had been replaced by occlusives after resonants and semivowels, the sounds [ð] and [ɣ] disappeared from the language altogether (this voiced velar fricative is entirely different from the Medieval Armenian development of a uvular fricative based on Classical Armenian *ł*). Between vowels, [ð] and [ɣ] were replaced by *y*, originally perhaps a palatal fricative [ʝ] but early coalescing with the semivowel *y* (Winter, 1955). Moreover, Proto-Armenian [β], which has two different sources from PIE, **u̯-* and intervocalic **p* became *w* (*ew* 'and' from **epi*, (cp. Greek *ἐπί* < IE **h₁epi*) with an intermediary **eβi* stage). The positing of a voiced bilabial fricative is not problematic insofar as its phonological evolution makes articulatory sense.

Though there has been research for Persian (Shademan 2002 and Perry 2005, which are more heavily corpus-based), such an analysis is trickier for Armenian, given that, due to its larger vowel and consonant inventory from Persian (except the interdental fricative, which disappeared in Modern Persian), we must look at other processes to distinguish the two, such as etymologically-matching lexical concatenation (cf. *roy → royal, royalty; *royly, *royness) and other phonological processes mentioned below.

As an aside, lexical domains (especially in regards to registers) also appear to be unequally represented in these strata – for example, in Japanese, the native Yamato (pre-Chinese contact) stock, though smaller than the *kango* (漢語) or "Han words" layer (it is estimated that approximately 60% of the words contained in a modern Japanese dictionary are *kango*, but they comprise only about 18% of words used in speech (Shibatani, 1990: 142)). The same is true of the Norman-French layer in English or Khmer-Sanskrit layer in Thai; though there are no mass etymological corpus data for Armenian, the same tendency holds true only to a lesser extent, as:

Winter further justifies the existence of the /f θ x/ series by pointing out two things – firstly, that the preservation of the IE cluster **st-*, which remains unchanged throughout the Proto-Armenian period, points to a phonemic contrast between /θ/ and the *-t-* of *st-*, and that secondly, the positing of Proto-Armenian */f θ x/* as reflexes of IE **p *t *k* enables one to interpret successfully the fact that IE **tw-* and **sw-* develop in Armenian the same way, such as *kʻaṙasun* 'forty' (cp. Greek *τεσσαράκοντα* < **kʷetwr̥̄k̂omt*, from earlier **kʷetwr̥̄-dk̂omt* ("four-ten")) and *kʻo 'thou* (gen.)', but *kʻoyr* 'sister' < **swesōr* and *kʻirtn* 'sweat' (cp. Greek *ἱδρώς*).

The PIE labiovelars are far less controversial; given that Armenian is a satem language, the PIE labiovelars had lost their labialization and had become velars, as *elikʻ* 'he left' > **elikwet* (Fortson, 2010: 385), though there are some (Winter, 1962: 258) who maintain that Proto-Armenian must have had an original distinction between the plain velars and the labiovelars, since in certain words, the two have different outcomes in front of front vowels, such as the famous example for 'four', *čʻorkʻ*. Winter asserts that the merger of **k* and **kw* and related pairs, so characteristic of satem languages, took place within Proto-Armenian itself and need not be ascribed to a more remote period in the history of IE, though he concedes that forms such as hing for 'five' cannot be accounted for through his hypotheses.

During the early period of Armenian phonology, we also find a *ruki*-rule, though its effects somewhat differ from what is found in Sanskrit. On the strength of the evidence presented by demonstrating several alternations such as *jaṙ* 'evil, wicked' vs. *garšim* 'to abominate, be disgusted', *mošˇ* 'blackberry' vs. *mor* 'black mulberry', *harsanikʻ* 'wedding' vs. *hašnikʻ* 'wedding (dial.), etc.. Martirosyan (2008) tentatively reformulates the *ruki*-rule in Armenian as follows: PIE *-s- following *k or *r yields -š- in post-apocopic internal pretonic or initial (or, simply, in the non-final) positions. In other words, in these positions, **-rs-* and **(-)ks-* yield *-(r)š-* and *-(k)š-* (and in the initial position, *čʻ-*), respectively, in contrast with *-ṙ-* and *-cʻ-* in the remaining positions.

> "Parthian material is not restricted to any part of the lexicon, or indeed any one grammatical category, but is found across the spectrum in both the basic lexicon (items concerning nature, body parts, abstract vocabulary of everyday life, etc.) and in more specialised segments (e.g. martial and technical vocabulary), in both of which the may occur as nouns, adjectives, adverbs, verbs, even invading closed classes such as prepositions and numbers[9]." (Meyer 2017: 20)

As for classification of the data into separate lexemes – the majority of the work was done through meticulously pouring over various etymological authorities (mainly J̌ahukyan 2010, Ačaṙean 1979, and Awetikʿean, Siwrmēlean & Awgerean 1836–37), with some guesswork necessary for certain lexemes (by deductive reasoning or checking backformations). Afterwards, these lexemes were separated into the three conventionally accepted periods of intense borrowings (as per Clackson 2008), and a systematic qualitative analysis was conducted to tease out the differences (in terms of phonological or morphological behaviour) among the three Persian loanword layers.

Given that this paper uses a modified version of a core-periphery-type cophonological framework, it was necessary to examine in some depth diachronic processes (which only rarely or haphazardly have fossilized or active synchronic effects[10]) which would help us draw conclusions about the organization of the Armenian lexicon based on the various morphophonological processes – four in particular, namely prothesis, phonotactics, metathesis, and reduplication. There are unresolved methodological issues given that the tools chosen herein are synchronic yet our problem is one chiefly of diachrony – in synchronic research, one is only allowed to model a naïve speaker's mental representation, whereas in diachrony we are dealing with an unchanging historical record.

9 For example, Armenian *hazar* '1,000' derives directly from Western Middle Iranian *hz'r* /*hazār*/ and Armenian *biwr* '10,000' from Western Middle Iranian *bywr* / *bēwar*/.

10 Hence the confusion between the Inkelas-style co-phonological frameworks versus the index-constraint frameworks (Itô and Mester-style approaches), as well as between lexical strata (i.e., core-periphery, etc.) vs. strata in a phonological grammar (i.e., levels of a morphophonological derivation in Stratal OT); such confusion, or fusing of various parts of different frameworks, is partly due to the author's insufficient knowledge of this vast field, and partly due to the (nearly exclusively) synchronic tools available in modern phonological theory which are ill-suited to diachronic issues.

Prothesis

We know that prothesis occurs in a particular stem either because of direct attested evidence in Persian (thus showing us the original, unprothesized form) or via phonological reconstruction, the way it has been customarily done in Indo-European studies for Greek, Phrygian, Indo-Aryan, etc. (PIE laryngeals are an especially rich source of prothesis in the daughter languages). In some cases, we can find alternants of a particular stem where prothesis has not happened, or has happened with a different vowel, e.g. [e] and [a] variants "lack a clear phonological distribution, and there are doublets for some words showing both [e] and [a] developments (e.g., *ełbayr* 'brother' beside *ałb′ar* in the Mirakʿ dialect)" (DeLisi, 2015:59).

Prothesis also appears to be differentiated within the two co-phonologies – native words generally add an initial *e*-, like in *erek*‘ 'three', whereas Persian borrowings, especially of Middle Median origin, generally add *a*-, such as in *ašxet* 'reddish, chestnut-coloured' (Périkhanian, 1966). In modern dialects, the Middle Median ([a])/Persian ([i[11]]) prothetic vowels are fossilized and can no longer be used as a repair strategy for onsets which violate Armenian phonology, but we know from documentary evidence that there was a long period where such prothesis was productive. Macak (date unavailable) elaborates that we know that unstressed high vowels were still faithfully realized as genuine high vowels in the Parthian period (ca. 250 BCE – ca. 230 CE), since high vowels in Iranian loanwords are in Armenian reflected with reduced vowels; cf. Manichaen Parthian *nyša'*(= *[ni:ʃa:]) → pre-OA [Old-Armenian] *[ni.ʃɑ] > OA նշան *nšan* [nəʃɑn] 'sign, symbol'. In contrast, loanwords from the later Sassanian period (ca. 230 CE – ca. 650 CE) represent high vowels 'faithfully' i.e., unreduced by the native phonology, cf. Pahlavī puštīkpā → OA փուշպան [*p'owštipan*] 'body-guard' etc. (cf. Ravnæs 1991: 61)."

A chronological difference lies behind the divergent treatment of Persian initial **r*-, which is in part rendered with a prothetic vowel as *ar*- or *er*- as in the inherited vocabulary, but in part appears as *ṙ*- as in the case of borrowings in later times and from other sources (Schmidt & Bailey, 1987).

11 Meyer (2017, FN29): "Prothesis with *i*- is less common than that with *a*- or *e*-; no conditioning factors for the choice between the three options have as yet been discovered (cf. e.g. Greppin 1982)."

There is no native Armenian word that starts with an /a/ + CC; for example, we have the proper name *Axšahrsart*[12] (which interestingly was used as a shibboleth during the reign of King Artašês (189-160 BCE) (*ibid*.)).

Moreover, the literature sometimes distinguishes between earlier loans and later loans based on which of the prothetic vowel(s) they use, as seen in the table below, with the former being supported by words such as *b**oy**ž* 'cure, remedy', *b**u**žel* 'to cure, heal' from Parthian *b**w**j*-, pronounced /*bōž*-/ 'to save, redeem' and *d**ē**mk*' 'face' from Western Middle Iranian *d**y**m,* pronounced /*dēm*/ (this word is so well-integrated that its wiktionary.org entry shows us 258 derived words, including the oft-used adverb *ənddēm* 'against', also used as a preposition to mean 'contrary, opposed, opposite'), and the latter being supported words such as *d**e**n* (no ablaut in the genitive *d**e**ni*) 'religion, faith' from Parthian *d**y**n* /*dēn*/, Armenian *hr**e**štak* 'angel, messenger' from Parthian *fr**y**štg* /*frēštag*/; Arm. ***ṙ**o**t* 'river' from Western Middle Iranian ***rw**d* /*rōd*/, Arm. *t**o**hm* 'family, seed', either from Parthian *t**w**xm* /*t**ō**xm*/ or Middle Persian *t**w**hm* / *t**ō**hm*/ (Meyer 2017:18). Regarding *ṙ* (represents a trilled rhotic instead of a flap), earlier loanwords appended a prothetic initial *e*- (like in *eram*, 'troop, flock', from Western Middle Iranian *ram*, 'flock, Manichaean community'), whereas later loanwords convert the Persian flap to an initial trill (*ṙazm* 'fight, battle' from Western Middle Iranian *razm*).

(5) Prothetic vowel differences in early versus later loanwords

Arm. (early loans)	Arm. (later loans)	Pth.
oý, u	*o*	*w* /o/
é, i	*e*	*y* /e/
er-	*ṙ*-	*r*- /r-/

Table 1.3 – *Stratal differentiation of loans from Parthian*

(Source: Meyer 2017: 19)

Certain authors (Kortland 1980, Meyer 2017) suggest that in terms of the chronology of sound change, the period corresponding to Lex_1 cannot have ended before the rise of secondary

12 Regarding this particular word, Meyer (2015) in FN26 explains: "This inscription is discussed in Périkhanian 1966. It is noteworthy for the [Aramaic] spelling of the name *'ḥštrsrt* /Axšahrsart/, a compound whose first part is cognate with Avestan *xšaθra* 'power, kingdom'; the prothetic vowel *a*- (denoted by aleph), together with other phonological changes, suggests a West Middle Iranian, but non-Parthian origin of the name. Based on this and a few other Armenian lexical items, Perikhanian suggests that the source language may have been (Middle) Median, which is otherwise unattested."

prothetic vowels (Kortlandt 1980:103), which the oldest layer of Persian loanwords exhibit. Secondary prothetic vowels are those which that "did not arise from [PIE] laryngeals as in, e.g., Arm. *anun* 'name', cp. Gk. *ὄνομα*, Lat. *nōmen*, or Arm. *erek* 'evening', cp. Gk. *ἔρεβος* 'darkness', Skt. *rájas* 'id.', ON *røkkr* 'twilight'. Secondary prothetic vowels occur before wordinitial consonant clusters and *r*-, for instance in Arm. *erek*' 'three', cp. Gk. *τρεῖς*, Lat. *trēs*, or Arm. *ełbayr* 'brother', cp. Gk. *φράτηρ*, Lat. *frāter*" (Meyer, 2017:337). Kortlandt thus indicates that certain early Persian words have undergone Armenian-internal sound changes.

Regarding the interaction between prothesis and metathesis – native stock words have a complex (for an overview of the scholarship's internal disagreements on this very point, see Picard 1989) series of phonological changes which involve both processes (the *CrV words shown in the table below exemplify this) which likely require rule-ordering, with prothesis being triggered before metathesis (*CrV → CerV or CarV → erCV or arCV) as shown in the table below.

(6) Armenian words containing both prothesis and metathesis

Armenian word	Indo-European origin	Gloss
aṙu (<*arsuy)	*srudis	'canal'
artawsr	*dráku	'tear' (n.)
erkan	*g^{w}ráwon	'millstone'
ałbewr	*b^{h}rwḗr̥	'spring-well'
ełbajr	*b^{h}rā́ter	'brother'

(Source: Adapted from Picard, 1994:15)

Lastly, there is still some debate as to the explanation behind the choice of various prothetic vowels for each of the layers, as DeLisi (2015:90) explains:

> "Ideally, a full theory of extraprosodicity should be able to integrate the results of the Articulatory Phonology experiments above with the preferences seen in loan incorporation and diachrony. Vaux & Wolfe[13]'s appendix theory cannot account for

13 This refers to the article written by these authors: Vaux, Bert & Andrew Wolfe. 2009. The appendix. In Eric Raimy & Charles Cairns (eds.), *Contemporary Views on Architecture and Representations in Phonology*, 101–

the bias towards initial prothesis associated with these segments. If the sibilant is a mere appendix, why would it preferentially attract a prothetic rather than cluster-internal epenthetic vowel?"

Owing to the fact that diachronic elements cannot be readily captured by cophonological approaches, one must look for synchronic traces as in the few examples above – in this case, we summarize this section by making use of several constraints that compare forms from different lexical layers – for instance, prothetic vowel differences among the different lexical layers. As in Stratal OT, we will need a reranking, suppression, or removal[14] of constraints to give us the correct candidate for each lexical layer. In the case of prothesis, the constraints are the same for both Lex_0 and Lex_1, but different for Lex_2 (and presumably $Lex_{3\setminus}$):

(7) Native (Lex_0) example

/reyek'/, 'three', from PIE *_tréyes_[15]	*[o, e, ɾ̊]PROTHESIS	[u,i,e/a(r)]PROTHESIS	DEP-IO
☞ a. erek'			**
b. ɾ̊ek'	*!		
c. orek	*!		
d. rek		*!	

(8) Earlier loanword (Lex_1) example

/ʾrwst/, 'craft', from Parthian צרכתת)	*[o, e, ɾ̊]PROTHESIS	[u,i,e/a(r)]PROTHESIS	DEP-IO
a. ɾ̊owest	*!		
b. rowest		*!	
☞ c. arowest			***
d. orowest	*!		

143. Cambridge: MIT Press.

14 As Prince (1998) puts in, "Antagonism: [a] constraint and its "anti-constraint" cannot both be active in one hierarchy: the lower ranked of the pair may be simply removed. No such property holds of the rule in serial derivation."

15 Initial PIE /t/ is dropped early on, and "all historical grammars of Armenian since Meillet (1903) have agreed that PIE *s was lost between vowels in the prehistory of the language, with resulting contraction of vowels in hiatus: the classic example is the nominative of 'three', erek" (Kim 2018:100). Just like the PIE word for 'two' (*_dwo_), 'three' received an identical prothetic vowel at some point before the Classical era, through numerous sound changes dw > dg > rg > erg > erk or dw > tw > tk > rk > erk (Winter, 2011:355).

(9) Later loanword (Lex_2) example

/razm/, 'fight, battle', from WMI *razm*)	*[u,i,e/a(r)]PROTHESIS	[o, e, ɾ̓]PROTHESIS	DEP-IO
a. irazm	!*		
b. erazm	!*		
c. razm		*!	
☞ d. ɾ̓azm			*

Rough Sketch

This paper will briefly focus on the fourteen rules[16] described below, with phonotactics given special consideration, and the analysis will concern the lexical layers of the three intense (Lex_2 being the most influential on the language) periods of contact between Armenian speakers and speakers of various Old and Middle Persian variants. As stated by Meyer (2017), regarding these rough timeline approximations, "a more than relative dating of the differences between layers of loanwords is, unfortunately, impossible owing to the lack of continuous evidence [of various Old or Middle Persian varieties] and its imprecise writing system" (p.19).

a. INHERITED-IE-PHONOLOGY — every word, native or borrowed, must adapt to the phonological inventory and rules of the inherited Indo-European layer;

b. [ə]EPENTHESIS — allows for post-lexical phonological rule inserting a schwa to break up illegal clusters in either the word or derivations;

c. NONOSALV — no nasal[17] vowels allowed;

d. METATHESIS — exhibits various patterns[18] of metathesis;

16 Because this paper deals with large swaths of a language's phonological structure, these rules can be considered constraints but with the important caveat that such constraints are defined more loosely than what most authors in the literature use. Some of these rules are umbrella terms, such as "NO-NONNATIVE-CLUSTERS" which would, in reality, capture potentially dozens of constraints in one, and many of the other rules are results of interactions of separate markedness and faithfulness constraints.

17 In terms of articulatory phonetics, we are referring to phonemic nasal vowels proper, not merely nasalization effects due to nearby nasal consonants; however, in phonological terms, given that this is a constraint-based model that follows the basic assumptions of OT, we cannot make this distinction as there are no Morpheme Structure Constraints due to richness of the base. For a new theory of non-derived environment blocking that attributes blocking to an opaque interaction between Morpheme Structure Constraints (which constrain possible underlying forms in the lexicon) and the usual phonological mapping from underlying forms to surface forms, see Rasin (2016).

18 The PIE-derived lexicon has different types of metathesis not seen in the Persian-derived Lex_1 and Lex_2.

e. PARTIAL-REDUPLICATION	allows partial reduplication of morphemes within a word;
f. NO-[ø]	no close-mid front rounded vowels;
g. PRODMORPHOLOGY	allows for productive derivational morphology;
h. [o, e, ṙ]PROTHESIS	allows for prothesis using prothetic [o], [e], or [ṙ[19]];
i. [u,i,e/a(r)]PROTHESIS	allows for prothesis using prothetic [u], [i], or [e/a(r)];
j. CLOSED-CATEGORY	contains words belonging to closed categories such as prepositions, conjunctions, numbers, determiners, and inflectional morphology;
k. FULL-REDUPLICATION	allows full reduplication of morphemes within a word;
l. NO-NONNATIVE-CLUSTERS	no non-native clusters allowed;
m. ABLAUT	nouns undergo ablaut in different case markings; and,
n. CASE-MONOPHTHONGIZATION	monophthongization of diphthongs in case markings.

Below is a rough sketch of the core-periphery model applied to Armenian, with different Itô & Mester-type strata in the Armenian lexicon, ordered such as in Figure 2, each successive level is more marked than the previous one (and allowing the variable positioning of certain faithfulness constraints as explained in the Reduplication section below) and thus has fewer markedness constraints. These layers of diachrony will be accompanied by data showing the relevant morphophonological processes if extant and surface-oriented phonotactic constraints in other cases (since there are many cases of phonological processes which are no longer productive in Modern Armenian and even no longer productive by the Classical (5th century CE) era). Figure 10 below shows us that with each successive layer, there are additional constraints that can be violated.

19 Judging from a few alternating pairs and dialectal variation alive today, IE **rs* tends to become Proto-Armenian **rš*, which later became either a double-*r* sequence (double flap), or a trilled *ṙ* (on this topic, Vaux (1998) and others has often doubted the phonemic status of the two rhotics and others interpret *ṙ* as merely a geminate of *r,* but there exist certain minimal pairs such as *taṙ* 'letter' – *tarr* 'element'), such as *awṙ* 'bottom', cf. Greek *ὄρρος* from **$h_1$orsos,* and *tʿaṙanim*, *tʿaršanim* 'I wither'.

(10) Lexical strata

Lex_{MAX} -
Lex_5 - (="FAITH/NONOSALV")
19th – 20th century
Lex_4 - (="FAITH/NO-[ø]")
12th century – 19th century
Lex_3 - (="FAITH/PRODMORPHOLOGY")
5th century – 11th century
Lex_2 - (="FAITH/METATHESIS/FULL-REDUPLICATION/[o, e, ṙ]PROTHESIS/CLOSED-CATEGORY")
3rd century BCE – 5th century CE
Lex_1 - (="FAITH/PARTIAL-REDUPLICATION/NO-NONNATIVE-CLUSTERS/[u,i,e(r)]PROTHESIS")
3rd millennium BCE – 4th century BCE
Lex_0 - (= "FAITH/INHERITED-IE-PHONOLOGY/[ə]EPENTHESIS/")
> 4th millennium BCE

Phonotactics

As DeLisi 2013 explains, during the earliest stages corresponding to most of the period covered in our Lex_0 and Lex_1, "[d]ue to extended and intensive contact with early East Caucasian languages, the […] lexicon was forced to undergo rather remarkable phonotactic and phonological changes" (p. 476) and that consequently, "Armenian acquired much stricter phonotactic constraints than the Maximum Syllable Template[20] it had inherited from Proto-Indo European" (DeLisis, 2015:47), as the usual maximal syllable in modern East Caucasian languages is CVRC, and "CR clusters were prohibited in Proto-East Caucasian in both initial and medial position" (Kassian & Yakubovich 2002:44). DeLisi (2015) explains that these strict phonotactic constraints were an areal feature, affecting at least Proto-East Caucasian, Armenian, and Ossetic.

Some Armenian phonemes, namely *p* and *č*, but also *š*, *ž*, and *x*, appear only exceptionally in words inherited from Indo-European[21], but commonly in Persian loanwords; characteristic of

20 DeLisi (2015) bases herself on Byrd (2010)'s work, which states that the Maximum Syllable Template "consists of two consonants in the onset and two consonants in the coda. The onset may violate the SSP [Sonority Sequencing Principle]; the coda may not." One extrasyllabic segment is thus allowed at the left edge of the word, and multiple extrasyllabic segments are allowed word-finally.

21 As mentioned in FN26 of Meyer, 2017:16-17: "Arm. *p* can derive from PIE **b* (e.g. Arm. *əmpem* 'to drink' [from] PIE **pi-ph₃-* with analogical nasal infix (cf. Martirosyan 2010:277–8), but the latter sound is rare in Indo-European; some lemmata suggest that PIE **p* may result in Arm. *p* in consonant clusters, e.g. Arm. *aṙaspel*

Persian loanwords are final consonant combinations, in particular *-zd*, *-zm, -xt*, *-nd*, *-nj*, *-šx*, *-šk*, *-št*, *-sp*, *-st*, *-rd*, *-rz*, *-rk*, *-rh*, and *-rt*. (Schmitt & Bailey, 1987).

Codas with nasal + affricate (and to a lesser extent, + plosive) are usually only found in Persian loanwords, such as in *kinč*, 'boar' (compare Persian کنج (*kinj*), کنجر (*kinjar*, 'large elephant')), *hnazand* 'obedient' (from Proto-Iranian **hu-nazand*), and *varung* 'cucumber' (from Middle Persian *vātrang*).

CC-*anel*-type verbs are common in Lex_0 stock (*mtanel*, 'to enter, go, come in, introduce, insinuate or intrude oneself', *gtanel*, 'to know, find out, gain, discover' (cognate with English *wit* and German *weisen*, *Wissen*), *lk'anel* "to desert, forsake'), and some Lex_1 verbs from Persian are actually remolded to fit this pattern, such as *snanel* 'to feed, nurse' when one would have expected **sananel* (Considine 1979) if it had followed the Persian pattern. Lex_0 stock of this type invariably feature a full unreduced vowel for its root noun (CVC which becomes CC-*anel*, with V representing any vowel except a lexical schwa) – this process of vocalic reduction is still synchronically productive, such as the nominative-accusative Western Armenian *tuṙ* to *təṙan* 'door', or *ʃun* 'dog' and *ʃənal* or *ʃənanal* 'to act like a dog, to prostitute oneself'. The Persian pattern would have not reduced a monosyllabic word's vowel to a schwa in derived forms.

As seen in the table below, the majority of vowel borrowings have been unproblematic, though we do see some degree of nativization insofar as vowel length is concerned (vowel quantity status is unclear in Proto-Armenian but was likely noncontrastive (had a low functional load), thus presumed to be lost by the 5th century CE in the first Classical Armenian texts). What is interesting here is that Armenian has a stress-conditioned word-final ablaut rule (usually resulting in the syncopation of vowels in pre-tonic syllables) which applies fully to all the correspondences seen in Figure 11. Later loanwords of Persian (Lex_3) and non-Persian origin (Lex_4 and Lex_5) generally do not follow this syncopation or vocalic reduction rule – for example, high vowels become schwas in derived forms of words like *bžišk* 'doctor, healer' → *bžəškutʿiwn*

'myth, fable', cp. **spel-*, Goth. *spill* 'fable', OE *spell* (cf. Beekes *apud* Kortlandt 2003:197). Arm. *š* occurs in some inherited words such as *šun* 'dog' [from] PIE **ḱu̯ōn*, cp. Gk. *κύων*. These and the other sounds mentioned are, however, only sparsely attested in Indo-European heritage words."

‘medicine’ and *bžəškakan* ‘medical’, yet such processes are blocked in later loanwords like *kʻimia* ‘chemistry’ → *kʻimiarar* ‘chemist’ (not **kʻəmiarar*), and *zibil* ‘trash’ (from Ottoman Turkish, ultimately from Arabic زِبْل (*zibl*)) → *zibilanocʻ* ‘dumpster, trashy place’, and not **zibəlanocʻ*.

(11) Strata of Iranian (Persian) lexical and phonological influence on Armenian

1.3. Strata of Iranian lexical and phonological influence on Armenian

Arm.	Pth.	Examples
a	*a, ā*	Arm. *azat* ‘free, noble’ < WMIr. *ʾzʾd* /āzād/; Arm. *marz* ‘border, province’ < WMIr. *mrz* /marz/
e	*e*	Arm. *pet* ‘chief, head’ < WMIr. *-byd* /-bed/
i	*i, ī*	Arm. *dpir* ‘scribe’ < WMIr. *d(y)byr* /dibīr/; Arm. *Mihr* ‘Mihr, sun god’ < WMIr. *myhr* /mihr/
ea	*ya, yā*	Arm. *seaw* ‘black’ < Pth. *syāw*
u	*u, ū*	Arm. *bun* ‘root, origin’ < WMIr. *bwn* /bun/; Arm. *bazuk* ‘arm’ < WMIr. *bʾzwg* /bāzūg/

Table 1.2 – *Phonological correspondences between Armenian and Parthian vowels*

PIE **s* usually disappeared at the beginning of words, but sometimes it changed into *h* initially, a sound change that has taken place in Armenian, Greek, Iranian, Phrygian, Lycian and also in Brythonic Celtic (Szemerényi 1985; Clackson 1994: 53-54), e.g. *hin* 'old (man)', **seno-* (cp. Latin *senex*). PIE *h_1 disappeared in Armenian, in Persian loanwords the initial *h* is preserved like in *hasb* ‘cavalry’, but *eš* ‘donkey’ in Armenian, both ultimately from **hₓéḱwos*, ‘horse’. Regarding an interesting case from the oldest layer of loanwords, “the well-known case of *partēz* ‘garden’, which is usually treated as a very old Iranian loan reflecting the devoicing shift d > t)” (Martirosyan 2013: 99), we can see that Persian words with voiced plosives tended to become unaspirated voiceless plosives in Armenian.

According to Pedersen (1905: 196), the Proto-Armenian intervocalic *-w-, itself derived from PIE *u̯, “erscheint als arm. *v* wo es auslautend geworden ist, sonst aber als *g*”. Persian loanwords such as *govem* 'I flatter' are misleading here and should be ignored. Pre-Armenian prevocalic **w*- always passes to g: e.g. *gini* 'wine', and *gorc* 'work' from Proto-Indo-European **worǵ*-, cognate with Ancient Greek ἔργον, Avestan *varəzəm*, Persian برز (*barz*, 'agriculture,

seed, sown field'), and Old English *weorc*. Note also a loan from Armenian to Georgian: *agaraki* 'field' (Greppin, 1991).

In contrast with the clear-cut differences between the canonical Sino-Japanese and native Japanese stock as mentioned above, Armenian and Iranian are independent branches of Indo-European and share sometimes parallel phonetic developments which complicate judgments on the status of a lexeme; a frequently cited example is Armenian *naw* 'boat, ship': is it an Iranian loan (cf. Ossetian *naw*/*nawæ* 'boat', Khotanese *no* 'boat', Parthian *nāwāz* 'skipper' > Arm. *nawaz* 'boatman') or an inherited word next to Sanskrit. *náu*- 'boat', Gr. *ναῦς* 'ship', Lat. *nāvis*, 'ship', and Old Irish *nau* 'ship'? (Martirosyan, 2013: 105).

An Armenian word starting with *c*, *čʿ*, or *cʿ* cannot ever have a Persian origin from any dialect of any era (the three words included in the data list have strongly contested etymologies) – they must either be from inherited vocabulary or borrowings from non-Persian sources.

The phonotactic constraints we have seen thus far can be graphically formalized as such for codas:

(12)

Coda	*zd*	*zm*	*xt*	*nd*	*nj*	*šx*	*šk*	*št*	*sp*	*st*	*rd*	*rz*	*rk*	*rh*	*rt*
Lex_0	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Lex_1	*	*	*	✓	✓	✓	*	*	*	✓	✓	*	*	✓	✓
Lex_2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lex_3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

It should be pointed out that Lex_0 is meant to be the most stringent, so Figure 13, especially at first glance, might appear strange since it has checks in Lex_0, but not in Lex_1, Lex_2, and Lex_3 for certain phonemes such as c, cʿ, and čʿ, though as noted above, this is a historical artifact because these are single phonemes that simply did not exist in the inventory being borrowed from (it would have been a different issue than when a cluster that might be repaired in

Lex_0, but that stays faithful in Lex_1, etc.), so this artifact does not actually contradict a core-periphery model.

The phonotactic constraints we have seen thus far can be graphically formalized as such for initial and medial positions:

(13)

	p-	*c-*	*cʿ-*	*čʿ-*	*č-*	*š-*	*ž-*	*x-*	*hr-*	*-hr-*	*sm-*	*-mb-*	*vs-*	*vč-*	*l-*
Lex_0	✓ʳ	✓	✓	✓	✓ʳ	✓ʳ	✓ʳ	✓ʳ	*	*	*	*	*	*	✓
Lex_1	✓	*	*	*	✓	✓	✓	✓	*	*	*	✓	✓	*	*
Lex_2	✓	*	*	*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓ʳ
Lex_3	✓	*	*	*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓ʳ

Note: superscript lowercase R indicates that the sound is rare in that layer.

However, surface phonotactics alone cannot be used to elucidate the problem – irregular declension in the morphology may prove to be more useful, as at least with the morphology we have a more readily available synchronic process to use as a diagnostic. The ABLAUT umbrella rule attempts to capture different morphological constraints in the cophonologies (it could be simplified to an *IRREGULARPARADIGM constraint), with earlier lexical layers showing far more ablaut-type effects in declensional paradigms than later layers, which become increasingly regular, and are lacking any sort of ablauting effect. This area of Armenian works in the opposite way from what Itô & Mester predict, as the constraint applies to newer forms that do not apply to older forms.

Metathesis

Without going into great detail, there appear to have been several metathetic processes between Indo-European and Indo-Aryan, and many more from Indo-European to Proto-Armenian, and in later periods, when Armenian borrowed extensively from different Persian dialects, it appears to have added a few more instances of metathesis, likely as an incomplete repair strategy for the new-coming words to conform to Armenian phonological rules. One such example is *ganj* ‘treasure’ from Parthian *gazn* (*z* later changed to *j*).

An important note here is that from the development of PIE to Proto-Armenian arose numerous instances of regular metatheses, partly from its own innovations and partly because of areal feature acquisition from nearby Caucasian languages – as evidence for this, DeLisi (2013) points out the fact Ossetian, a modern descendant of a Middle Persian variety spoken in the heart of the Caucasus, has "likewise undergone epenthesis and metathesis in words with branching onsets due to contact with East Caucasian languages. Both CR clusters and C[w] clusters metathesize in Ossetic" (p. 479). These non-Persian sources (PIE-derived or Sprachbund-influenced) of metathesis will not be considered here.

In certain situations, it is difficult to determine whether we are dealing with metathesis or deletion. For example, PIE **sCV* often becomes Armenian *sV*-, yet both phenomena can adequately explain *sV*-. If we accept metathesis, we would presumably reconstruct the sound changes as (here *p* is the consonant used based off the alterations for the word for 'army', *sah*, *spah*, *spay* (which is an early Persian loanword) and the dialectal alterations of the anthroponym *Step'anos* from Greek *Στέφανος*, *Tep'an(os)* and *Sep'an* (Ačaŕyan quoted in Martirosyan, 2008)) **sp*- > **ps*- > s- (Lidén, 1933: 50-52); if we accept deletion, we would presumably reconstruct it as merely *spV- > *sV-. Martirosyan (2008) provides evidence for both possibilities.

A chronological difference lies behind the divergent treatment of Persian initial *r*-, which is, in part, rendered with a prothetic vowel as *ar*- or *er*- as in the inherited PIE vocabulary, but also appears as *ṙ*- as in the case of borrowings in later times and from other sources; whereas consonant groups of stop plus *r* (with maintenance of the stop as such) underwent metathesis in inherited words (e.g. Armenian *rt* from Indo-European **dr*), Persian clusters like *dr* or *gr* are kept unchanged (Schmidt & Bailey, 1987). While clusters of occlusive and **r* of Indo-European pedigree regularly undergo metathesis in Armenian (e.g. PIE **bʰréh₂tēr* > CA *ełbajr* 'brother', Western Armenian *jeʁpʰajr*), words of Iranian origin do not undergo this change, thus Armenian *draxt* 'garden, paradise' derives from Western Middle Iranian. *drxt* /draxt/ 'tree', but other clusters regularly undergo metathesis when borrowed into Armenian, such as: *ašxat*, from *ʾxšʾdyh*, 'suffering', cf. Sanskrit क्षत (*kṣata*, "hurt"). Thus, the core stratum in Armenian generally has many more cases of diachronic metathesis than any of the later loanword layers:

Such a comparison strongly suggests, given what we know about the general properties of linguistic change, that instead of metathesis having occurred only word-finally in the majority of Iranian dialects, on one hand, and both initially and finally in Ossetic [an isolated northern Iranian language spoken in Georgia] as part of a single process, on the other hand, it is much more likely that final metathesis emerged first in the common ancestor of all these dialects long before #Cr ever became VrC in Ossetic. In addition to this obvious similarity to what occurred in Ossetic, another reason for viewing initial and final metathesis as separate and unrelated historical processes in Armenian is that sound changes are so seldom found to operate in these two environments simultaneously. For instance, one would be hard put to find a language where vowel deletion or obstruent-devoicing affected word-initial and word-final segments at the same time (Picard 1989:67).

However, the picture becomes more complicated when we compare Persian loanwords in a neighbouring non-Indo-European language; in Georgian, we have *pʿarmani* 'permit, licence' from Middle Persian. *framān* as opposed to Armenian *hraman* 'order'; Georgian *pʿarsaxi* 'parasang' from Parthian. **frasax* (implied by the Syriac loanword *prsḥʾ*) as opposed to Armenian *hrasax*, or, because of the absence of metathesis, Georgian *pʾitʾiaxši* 'governor, viceroy' as opposed to Armenian *bdeašx*.

Another complication here is that we also do not see any expected metathesis from *xr-* to **rx-* as in Armenian *xrat* 'wisdom, reason' from Western Middle Iranian *xrad* and Armenian *xoyr* 'headgear, diadem' from Parthian *xwwd* /xōδ/ 'helmet'.

From the set of syllable-final consonant clusters given above in the Phonotactics section, we also see peculiar metathesis (Armenian *šx* and *rh* from Persian sources *xš* and *hr*) which does not occur elsewhere in the language. The figure below sums up the rule-ordering for metathesis[22]:

22 Though Figure 14 does not directly pertain to the core-periphery modelling seen here, it helps us understand the diachronic relationship amongst these four processes.

(14) Rule-ordering for metathesis-elision-prothesis

	*#CrV	*VCrV(C)#
Vowel Deletion	—	VCr(C)#
Final Metathesis	—	VrC(C)#
Vowel Insertion	#CerV	—
Initial Metathesis	#erCV	—
	#erCV	VrC(C)#

(Source: adapted from Picard, 1989: 68)

Unfortunately, even in the non-native lexical layers where metathesis occurs most often, it is a process too rare for us to reliably use as a determinant for lexical stratum identification. Though it may seem unusual that the periphery would allow for a smaller set of metathetic processes (and for the outer ones, none at all), we can find a few parallels with other processes studied in other languages, such as Kertész (2003:76) who points out that for light versus heavy syllable structures in Hungarian, "the 'possibilities' admitted on the periphery are more restricted than in the native stratum [...] introducing an extra restriction in the peripheral stratum, the language permits more structures (both heavy and light syllables) in the native vocabulary than in foreign words."

Reduplication

Since there has been considerable research on using co-phonological approaches to explain different reduplication patterns in various languages (Jaafar & Raihan 2012 and Downing 2008), reduplication patterns within the two co-phonologies, which exhibit certain differences, will be briefly explored in this section.

Both native (such as *cicatil*, 'to laugh', *ddiel* or *ccel* 'to suck (milk from the breast)', *ačel*, 'to grow, increase', *mrmnǰel*, 'to murmur, mumble') and Persian-derived words are capable of partial reduplication, though the Persian ones, at least for the later loans, tend to have already been borrowed into Armenian with its partially reduplicated forms intact (thus are unlikely to have been a result of a productive process). An illustrative example of what is likely to be an

early loanword is *popoz* ‘sharp, pointy’ (origin disputed but likely from Middle Persian *pōz* /pwc/, meaning ‘horn’, later ‘nose’ as in Modern Persian), which has dialectal variants *pipoz* (Karabakh), *pupuz* (Goris), and even *poploz* (Moks), but no **pozpoz*.

Native words also sometimes exhibit consonantal changes, such as Armenian *kaskac* 'doubt, fear' (found in Classical Armenian, the Bible, and several dialects; in the Łarabał and Ararat dialects we find a more archaic version: *kackac*) derives from **kac-kac*, a reduplication of **kac-*, probably found in *karcem* 'to assume, to doubt, to opine' (Hübschmann, 1897: 533-534). The phonetic change *-ck- > -sk* indicates a consistent shift in Proto-Armenian and can help to reinterpret and understand some formations and etymologies according to Martirosyan (2008). It also helps explain *kas-karmir* 'entirely red' (Ačaŕyan 1913: 553), which is treated by Vaux (1998: 242-244) as a fixed coda reduplication but Martirosyan (2008: 550) proposes to treat *kas-karmir* as a compound of two words: *ka(y)c* 'spark' + *karmir* 'red' = Proto-Armenian **kac-karmir* > *kas-karmir*. Rarely, it is possible to find native words with both valid fully reduplicated and partially reduplicated forms, such as *parap* ‘empty’, *parap-parap* ‘idly, wastefully’, and *pas-parap* ‘completely empty, thoroughly hollow’.

Jaafar & Raihan (2012) show that, for Perak Malay, one of the Malay dialects spoken within the subnational state of Perak, it is possible to have a coherent system whereby the faithfulness constraint, MAX-BR[23] (requires that every element in the base to have a correspondent in the reduplicative morpheme) and the markedness constraint NOCODA switch ordering from Lex_0 to Lex_n.

Persian loanwords typically cannot have full reduplication (a notable exception is *zanazan*, ‘different, various’, from Parthian *zanag*): from Proto-Armenian **mar-mar-* we have *mar-m(a)r-il* ‘to shimmer, flicker, glimmer, extinguish gradually’, which is precisely matched in

23 Kiparsky (2010:127): “In Stratal OT, there are no reduplication-specific correspondence constraints, i.e. no B/R or I/R constraints, and no O/O constraints either. The shape of a reduplicated or truncated element — the REDUPLICANT or TRUNCATUM — is determined by the interaction of normal Input/Output (I/O) faithfulness constraints with markedness constraints in a morphologically selected constraint ranking (a co-phonology, along the lines of Inkelas & Zoll 2005).”

Greek *μαρμαίρω* ‘to flash, sparkle, gleam’, nor ablaut reduplication – like in the native word *sarsur* ‘extremely cold’ (from *sar*, ‘ice’).

Persian loanwords, at least for earlier ones, can be fully reduplication but without a linking morpheme, such as the *-a-* infix[24]; examples are *azg azg*, ‘various, different, manifold’ (from a secondary meaning of the word, which typically means ‘nation, people, tribe’), *apuš apuš* ‘very stupidly’ (from Persian prefix *ap-* ‘without’ and *ʻuš* ‘ear’, compare Middle Persian (ʾp̄y- /abē-/) and Modern Persian بی (bē-), both negative prefixes, and Middle Persian ʾwš (ōš) respectively). This shows us that the full reduplication process can target different strata but through different means – one with and one without infixing.

Thus, basing ourselves on Jaafar & Raihan’s analysis of heavy versus light reduplication in different strata in Perak Malay, by applying co-phonology to account for Armenian partial and full reduplication, the full reduplicative morpheme is explained by the tendency of prosodic constituents to be of maximal size in Lex_0, while the partial reduplicative morpheme seen in many Persian loanwords is explained “by the opposing tendency of some prosodic morphemes to have unmarked structure and be distinctly ranked” (*ibid*., 99).

Moreover, there is a vocalic reduction process[25] that affects the reduplicant that is specific to lexical strata Lex_2 and earlier, such as *caxel* ‘to sell’ and *caxc(ə)xel* ‘to sell out’ or ‘to sell all of one's possessions quickly and cheaply’, and a small number of these words have other reduplicated forms using a different vowel from the first segment, such as *caxcux* ‘trade, commerce’.

24 Though very rare, certain later loanwords from Classical Persian (Lex_3) allow for infixes other than *a*, such as *kuzekuz*. ‘hunchback’.

25 Though this can be the topic of a detailed paper, partial reduplication in Armenian also often involves the shifting of the initial base consonant within the reduplicant. For a similar process in other languages, a dissimilatory phenomenon has been proposed: “in those cases the difference between the two copies is a consequence of this ranking, but it is not uncommon to find cases of reduplication where it appears that the reduplicant is actually mandated by the grammar to be non-identical to the base along some dimension. Consider, for example, the case of melodic overwriting […] [i]n Abkhaz (NW Caucasian), reduplicated nominal constructions meaning “X etc.”, *m-* replaces the onset of the second copy (or supplies an onset in case of vowel-initial stems) (Vaux 1998, Bruening 1997). When the base itself is *m*-initial, however, the reduplicant begins with *č*“” (Inkelas & Zoll, 2000:28).

The full set of constraints below is based on McCarthy & Prince (1993: 16, 122-124), adapted from Zukoff, 2002:10-11); where "along each of these correspondence dimensions, the family of correspondence constraints evaluates the faithfulness of the relationships between segments" (*ibid*.):

MAXIMALITY (MAX) – Every element of S_1 has a correspondent in S_2

= No Deletion

DEPENDENCE (DEP) – Every element of S_2 has a correspondent in S_1

= No Insertion

IDENTITY(F) (IDENT(F)) – Corresponding segments have identical values for feature F

= No Feature Changing

CONTIGUITY (CONTIG) – (a) The portion of S_1 standing in correspondence forms a contiguous string (= No Skipping); (b) The portion of S_2 standing in correspondence forms a contiguous string (= No Intrusion)

ANCHOR – Any element at the designated periphery (i.e. left-edge or right-edge) of S_1 has a correspondent at the designated periphery of S_2

= No Insertion or Deletion at edges

LINEARITY – S_1 is consistent with the precedence structure of S_2, and vice versa

= No Metathesis

UNIFORMITY (UNIF) – No element of S_2 has multiple correspondents in S_1

= No Coalescence

INTEGRITY (INTEG) – No element of S_1 has multiple correspondents in S_2

= No Breaking/Splitting

Since such a maximally large set of constraints is too detailed for our purposes, we can simplify them by removing the latter four (since they mention phonological features partially dealt with elsewhere) and moving the ranking of MAX to account for partial reduplication (moving the ranking of CONTIG to account for the infixing morpheme if we were to fine-tune it further). Whichever type of reduplication we describe, all segments on the base have correspondents in the reduplicant, but in the case of full reduplication, the segment is repeatedly

wholly – the first and second string within a reduplicated word are instantiated by two free-standing, morphologically related outputs (this is output-output-correspondence (Kager 1999:263)), but the situation is more complicated if we wished to entirely account for partial reduplication as there is some variation within Armenian as to which elements get repeated (and whether or not the vowel changes[26]), but such details can be safely ignored here. We can further simplify matters by combining these constraints to come up with a global hierarchy of constraints (co-phonology theory here is implemented as a type of Stratal OT), which for the native lexicon would be:

PARTIAL-REDUPLICATION >> FULL-REDUPLICATION >> FAITH (MAX-BR)

Thus, in a fully reduplicated form, neither of these constraints are violated, whereas in a partially-reduplicated form the FULL-REDUPLICATION constraint is violated. Partial reduplication can thus be said to be more unmarked or less marked. The ranking in Lex_1 – Lex_2 would be:

PARTIAL-REDUPLICATION >> FAITH (MAX-BR) >> FULL-REDUPLICATION

And finally, the ranking in Lex_3, which very rarely engages in reduplication of any kind, would be:

FAITH (MAX-BR) >> PARTIAL-REDUPLICATION >> FULL-REDUPLICATION

If we were blind to any etymological information, we would have better than chance odds to correctly guess that a word that allows full reduplication would belong to the native stratum, and that a word that allows partial reduplication would likely belong to either Lex_0 or Lex_{1-2}, though we would need to exercise caution as reduplication of either kind is not very common. To sum up our stratified lexicon analysis, we can rely on a reranking of constraints specific to each layer within OT for us to find out the correct candidate, as in the following nonstandard tableaux:

26 The reduplicant will be reduced vis-à-vis the base since it bears fewer contrasts and almost always has simplified codas and vowel contrasts, such as in *kokord* (from *kord* which derives from PIE *$*g^{w}erh_3$).

(15) Native (Lex_0) example

/corak/ ‘nape’	PARTIAL-REDUPLICATION	FULL-REDUPLICATION	FAITH (MAX-BR)
a. corak			✓!
☞ b. cocorak	✓		
c. corakcorak		✓!	

(16) (Lex_1) example (also applicable to Lex_2)

/pōz/ ‘’sharp, pointy’	PARTIAL-REDUPLICATION	FAITH (MAX-BR)	FULL-REDUPLICATION
☞ a. popoz	✓		
b. pozpoz			✓!
c. poz		✓!	

(17) (Lex_3) example

/kʿičʿ/ ‘few’	FAITH (MAX-BR)	PARTIAL-REDUPLICATION	FULL-REDUPLICATION
☞ a. kʿičʿ	✓		
b. kʿkʿičʿ		✓!	
c. kʿičʿkʿičʿ			✓!

Residual Issues: Derivability and Etymological Difficulties

Judging from the large number of attested derived forms of lexical items in Lex_1 and especially Lex_2, and comparing it to the different behaviour of most words in Lex_3, we can contend that Clackson (2008) was correct in stating off-hand that loanwords from the Sassanid period were not as well-integrated into the lexicon. For example, the word *axt* ‘illness, disease, indisposition, vice’, borrowed from Middle Persian, is extremely well integrated into the rest of the lexicon by measure of its great number of compounds and derivations[27]. Most of the data

27 By no means an exhaustive list (bare forms only, all attested): *axtabar, axtabek, axtaber, axtabic, axtaboyž, axtaboyc, axtaborbokʿ, axtagorc, axtažet, axtažetim, axtažetutʿiwn, axtažēt, axtažētutʿiwn, axtali, axtalicʿ, axtaxonawacin, axtackʿ, axtakan, axtaker, axtakic, axtakir, axtakicʿ, axtakrakan, axtakrem, axtakrim, axtakrutʿiwn, axtakcʿabar, axtakcʿagoyn, axtakcʿem, axtakcʿim, axtakcʿutʿiwn, axtahalac, axtahawak, axtahawakʿ, axtamart, axtamol, axtamoli, axtamolutʿiwn, axtayin, axtanam, axtankeal, axtašarž, axtašaržutʿiwn, axtapašar, axtasēr, axtasirem, axtasirutʿiwn, axtaspan, axtarar, axtarcarc, axtacʿucʿanem, axtawor, axtaworabar, axtaworagoyn, axtaworakan, axtaworim, axtaworutʿiwn, axtakʿał, axtunak, amēnaxtalicʿ, anaxt, anaxtabar, anaxtagoyn, anaxtacin, anaxtakan, anaxtakanutʿiwn, anaxtakir, anaxtakicʿ, anaxtakcʿutʿiwn, anaxtanali, anaxtanam, anaxtapēs, anaxtaworutʿiwn, anaxtutʿiwn, anxaxtakan, anxaxteli, bazmaxtean, bazmaxtutʿiwn, žantaxt, žantaxtakan, canraxtutʿiwn, heštaxtasēr, heštaxtim, heštaxtutʿiwn,*

from Lex_3 in this paper have either zero or a few derived words, such as *xišt* ‘spear, lance’, which has *xištel* ‘to spear, to skewer’ and *xištik* ‘short spear’ (most nouns imported during this period have no verbal correspondences).

The majority of the etymological information was obtained from J̌ahukyan 2010, Ačaŕean 1979, and Awetikʿean, Siwrmēlean & Awgerean 1836–37. Where there were numerous sources that located the source word (from say, a specific Persian variety), its lexical layer class was secured; where there were significant disagreements, a question mark was used after its suspected lexical layer.

Knowing beyond reasonable doubt that a lexical item belongs in Lex_1 presents us with a hard task, since it is “difficult to work out the details because of the scanty evidence available for the older Iranian dialects” (Schmitt & Bailey, 1987).

The methodology of dealing with such borrowings has been developed and applied by Kuiper (1995), Beekes (1998; 2000; 2003), Schrijver (1997), and Lubotsky (2001). It has been pointed out that an etymon is likely to be a loanword if it is characterized by some of the following features: 1) limited geographical distribution; 2) phonological or morphophonological irregularity; 3) unusual phonology; 4) unusual word formation; 5) specific semantics (see Schrijver 1997: 293-297; Beekes 2000: 22-23; L ubotsky 2001: 301-302). These are useful guidelines when having to deal with multiple layers of loanwords, though they often fail to elucidate us in regards to "areal" words, which we will briefly deal with later. However, for the purposes of reconstructive phonology, the mere 450 or so (Godel, 1975:67 places the number at 438, but some have been withdrawn or added since then) inherited IE root words Armenian has remains a solid diagnostic.

This may be evidence of later loans not participating in earlier phonological processes which swept earlier lexical strata of Armenian. The earlier loans regularly undergo differentiation between tonic and pre-tonic positions, and regularly undergo ablaut in different case markings (*dēmk‘* (nom. sing.), ‘face’, ***di**mac‘* (gen. sing.) and *ēš* ‘donkey (nom. sing.)’ to *išoy* (gen. sing.)),

naxaxtutʿiwn, čʿaraxtakan, čʿaraxtakcʿutʿiwn, čʿaraxtavat, čʿaraxtavar, and others.

whereas the later ones do not participate in these processes, suggesting that ablaut alternation was no longer productive in Armenian by the later Sasanian period.

In terms of nominal morphology, virtually all words from Lex_3 (and Lex_4 and Lex_5) have entirely predictable case marking systems and have no ablauting phenomena, whereas words from Lex_2 and earlier often undergo ablaut (particularly front vowels, such as *gês* (nom. sing.) 'long hanging hair' to *gisoy* (gen. sing., whose *i* form also shows up in the instrumental case)) The most irregular alternations occur in Lex_0, perhaps unsurprisingly, as this layer also exhibits ablauting with other vowels, monophthong-diphthong alternations within the declensional system, such as *keankʿ* 'life' (nom. sing.) to *kenacʿ* (dat. sing.), *nžoyg* 'excellent quality horse' (nom. sing.) to *nžugi* (gen. sing.) and even rhotic consonant shifting, such as in *leaṙn* 'mountain (nom. sing.)' to *lerinkʿ* (nom. pl.), *lerins* (acc. pl.), with attested variation within the ablative, *leṙnē* and *leaṙnē*. (this word for 'mountain' actually has both of the aforementioned phenomena, as *ea* is a diphthong). This thus supports the proposition that one may separate lexical strata by diachronic processes, in the sense that a much larger percentage of older loanwords exhibit ablauting for noun case and virtually all Lex_3 or later loanwords exhibit the predictable, regular declensional paradigms.

Data and Brief Quantitative Analysis

Lex_1

For the earliest layer of Persian loanwords, going by phonotactic pattern alone makes for a poor predictor of stratum identification – there are 94 lexemes in Appendix A, 69 of which cannot be predicted by surface phonotactics, and out of 25 tokens where our phonotactic tableau above (Figures 12 and 13) would predict affiliation with the Lex_1 stratum, 24 are correctly identified as belonging to Lex_1, along with one faulty prediction. However, due to the very nature of the co-phonological approach that later layers are supposed to accept a greater range of phonotactic patterns (which necessarily means that any phonotactic pattern seen in an earlier

layer will also be valid, thus accepted, for a later layer), 24 out of 24 of these correct predictions here can also be erroneously fitted in Lex_2 or Lex_3.

Derivability provides interesting support for a high degree of integration of these words, but they are not as high as one would expect, especially when comparing them to Lex_2 – 35 are highly derivable, 14 have medium derivability, and 45 have no or little derivability.

Lex_2

Like for Lex_1, Lex_2 lexemes at Appendix B show a high degree of morphological integration – out of 630 tokens, 280 have high derivability, 91 have medium derivability, and 259 have low derivability.

Phonotactics here provide some predictive power – 137 tokens for which the phonological tableau correctly predicts but overspecifies group identification (indicated by "1 2 3" under the column "lexical layer by phonotactics"), 54 tokens for which the phonological tableau correctly predicts group identification, and two tokens (*c'anc'* and *loramarg*) with incorrect predictions, where, in both cases, judging by their phonotactics alone, they ought to belong to the Lex_0 layer.

Lex_3

This stratum (Appendix C) suffers from a paucity of lexemes – a *prima facia* fact that may indicate a lesser degree of integration. Out of 8 tokens, only one can be correctly categorized by surface phonotactics, and even then, we overpredict as a word like *xišt* ('lance' or 'spear' directly borrowed from pre-Modern Persian) could also belong to the Lex_2 layer, and for the rest of the tokens it is not immediately apparent that they should belong to this layer. However, if we consider that none of these lexemes show any kind of process related to prothesis, metathesis, and reduplication (contrary to what we saw for Lex_1 and Lex_2), we can be surer of their more recent introduction into the language.

Unlike Lex_1 and Lex_2, Lex_3 lexemes show a lower degree of morphological integration – out of 8 tokens, 2 have high derivability, one has medium derivability, and 5 have low derivability.

For all layers, prothesis, metathesis, and reduplication in our data do not represent a sufficiently large sample for us to make solid statistical inferences – though the findings we do have, which are summarized in Figure 19, are interesting nonetheless. For instance, none of the later loanwords participate in reduplication processes, whether partial or full, and very few Lex_1 lexemes take later prothetic vowels (as defined by the literature).

Lexemes of externally uncertain origin

For lexemes where etymological authorities either strongly disagree with each other or simply cannot determine the time period in which the borrowings occurred, phonotactics can help us determine the possible stratum for 11 out of these 43 uncertain lexemes at Appendix D. From these 11, 2 make likely false predictions (*aspar* and *dašt*). 19 of these tokens have high derivability, 6 have medium derivability, and 18 have low derivability, which is in line with what we see in Lex_2.

It is interesting to see that compared to native words, a large percentage of these lexemes in this table below violate the previously mentioned NO-NONNATIVE-CLUSTERS constraint, which at least would indicate that they do not derive directly from PIE to Proto-Armenian, though that constraint alone cannot further elucidate which of the three Persian layers these words belong to.

Piecing all co-phonologies in Armenian together

We can combine the above observations and short analyses on phonotactics, prothesis, metathesis, the two types of reduplication, morphological productiveness, and other phenomena in a tableau (below at Figure 19) which uses the abovementioned core-periphery stratification

based on violations of markedness and faithfulness constraints that exist in the (morpho-)phonology of Armenian. This tableau also includes non-phonological elements (features of the grammar, essentially) such as the acceptance of new lexical items into closed grammatical categories, and includes Lex_4 (Ottoman Turkish and its influences from the 12th century to the early 20th) and Lex_5 (modern French and English loans).

We can identify bundles of properties and alternations that are exclusive to a particular stratum – for the Persian layers, it is harder to distinguish between the first and second stratum, but fairly easy to discern the third from the earlier ones. In general, earlier layers exhibit quirkier behaviour that is reminiscent of Armenian's innermost lexical core, and later layers exhibit more predictability and regularity; on the other hand, the phonotactics of the inner layers is more limited (stricter) and the outer layers, as the model predicts, allow for greater configurations of clusters and onsets not typically allowed by earlier layers.

Figure 18 below gives us four example of words that each belong to a different stratum – due to the vast range of phenomena captured by our fourteen rules, there are no words that can be affected by all fourteen.

(18) Tableau for a sample lexemes belonging to each stratum

	INHERITED-IE-PHONOLOGY	[ə]EPENTHESIS	NONOSALV	METATHESIS	PARTIAL-REDUPLICATION	NO-[ø]	PRODMORPHOLOGY	[o, e, ɛ]PROTHESIS	[u,i,e/a(r)]PROTHESIS	CLOSED-CATEGORY	FULL-REDUPLICATION	NO-NONNATIVE-CLUSTERS	ABLAUT	CASE-MONOPHTHONGIZATION
Lex_0 - hur	✓	✓	✓	-	-	-	✓	-	-	-	-	✓	✓	✓
Lex_1 - seaw	✓	-[28]	✓	-	✓	-	✓	-	-	-	-	✓	✓	✓
Lex_2 - ašxoyž	✓	-	✓	✓	-	-	✓	-	✓	-	-	*	✓	✓
Lex_3 - tʿarxan	✓	-	✓	-	-	-	*	-	-	-	-	*	-	-

28 "-" means that this particular lexeme does not participate in this rule.

(19) Tableau with a restatement of constraints and features separated by lexical strata

	INHERITED-IE-PHONOLOGY	[ə]EPENTHESIS	NONOSALV	METATHESIS	PARTIAL-REDUPLICATION	NO-[ø]	PRODMORPHOLOGY	[o, e, ɾ]PROTHESIS	[u,i,e/a(r)]PROTHESIS	CLOSED-CATEGORY	FULL-REDUPLICATION	NO-NONNATIVE-CLUSTERS	ABLAUT	CASE-MONOPHTHONGIZATION
Lex_0 (native IE)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lex_1 (earliest period of Persian Persian)	✓	✓	✓	✓	✓	✓	✓	N/A	✓	✓	✓	*	✓	✓
Lex_2 (second period of Persian)	✓	✓	✓	✓	✓	✓	✓	✓	N/A	✓	✓[r]	*	✓	✓
Lex_3 (third period of Persian)	✓	✓	✓	N/A[29]	N/A	✓	*	✓	N/A	N/A	N/A	*	*	*
Lex_4 (Turkish)	✓	✓	✓	N/A	N/A	*	*	N/A	N/A	N/A	✓[r30]	*	*	*
Lex_5 (French and English)	✓	✓	*	N/A	N/A	*	*	N/A	N/A	N/A	N/A	*	*	*

Conclusion

In general, this paper confirms Kiparsky (1968) who argues that there are different degrees of nativization and conventionalization among foreign words, but there are some issues. If lexical items are borrowed, they either need to be adapted to the phoneme inventory of the replica language, or that inventory needs to be appended. Armenian has done both. However, co-phonologies within a language may help disambiguate certain phonological processes which may be treated differently for each lexical stratum, even if they eventually become unproductive and opaque.

29 N/A means that the lexical stratum does not play a role in that process.

30 There are some Turkish-derived expressions or adjectives in modern Western Armenian (like *yavaş-yavaş* 'very slowly, methodically', *zaman-zaman* 'from time to time') which exhibit full reduplication, but it is unclear if these were loaned directly in their fully reduplicated forms.

"Competing analyses in theoretical linguistics are typically evaluated on their empirical coverage and theoretical parsimony" (Inkelas & Orgun 2003), which opens up a problem for us given that there are no competing analyses for our topic – we can thus tentatively say that the account of Armenian lexical layers developed in this mémoire using the Core-Periphery approach has decent, but not perfect, empirical coverage, captures some generalizations and is theoretically more streamlined than what other possible alternatives can offer. Ranking various strata of the lexicon according to their phonotactic permissiveness constitutes a piece of evidence for the ancientness or the degree of assimilation of particular lexical items – how well such a ranking matches the independently attested external (etymological) facts is a good measuring stick for how much predictive power such an analysis has. In our case here, relying on phonotactics alone proves to be a somewhat mediocre predictor of lexical layer identification, but if it is used as one factor along with an array of morphophonological processes, it can complement our search in teasing out these loanword layers.

Phenomenologically if a native speaker, assuming no knowledge of linguistics, asks themselves how they can know or intuitively feel that a particular word is Persian or native Armenian without knowing etymologies, they would only be able to suspect a non-Armenian origin for Lex_3 words – Lex_1 and Lex_2 act too similarly to native stock for the speaker to discern the difference. However, once a speaker is made aware that certain clusters are in fact borrowed from various Old and Middle Persian varieties, it becomes likely that they would be able to discern a higher percentage of Persian-derived words successfully.

In terms of morphological case ending paradigms, the findings in this paper appear to violate Itô & Mester's conclusions regarding concentric permissiveness – in our case, the newer layers appear to be more morphologically restrictive (thus showing more morphological regularity), but they are indeed more phonotactically permissive as the theory would predict.

In terms of our initial hypothesis – we can cautiously say that there is enough diachronic evidence (with important caveats mentioned throughout the paper) to propose that the layers of Persian loanwords are acting not only as self-contained Lex_n layers, but that they have not been

nativized to the point of being assimilated to Lex_0, though they are very close in most respects. The sheer amount of time that has passed (16-30+ centuries, depending on which Persian loanword layer we are concerned about) and the presence of initially counter-intuitive and murky sound changes make Armenian a particularly hard case for positing clearly defined co-phonologies compared to most other examples in the literature. Further research is required to flesh out a full co-phonological account of Persian loanword layers within Armenian – one that would analyze the entire lexicon (not just purported loanwords) using computational models.

Bibliography

Ačaṙean, H. (1979). *Hayerēn armatakan baṙaran* [Dictionary of Armenian Root Words] (in Armenian), volume IV, 2nd edition, reprint of the original 1926–1935 seven-volume edition, Yerevan: University Press.

Ariyaee, K. (2019). Loanword Adaptation in Persian: a Core-Periphery Model Approach, *Phonetics-Phonology MOT 2019 conference*.

Awetikʿean, G.; Siwrmēlean, X.; Awgerean, M. (1836–1837). *Nor baṙgirkʿ haykazean lezui* [New Dictionary of the Armenian Language] (in Classical Armenian), Venice: S. Lazarus Armenian Academy.

Beekes, R. S. P. (1998). Hades and Elysion, in Jasanoff, J., Melchert, H. C. & Oliver, L. (eds.), *Mir curad: Studies in Honor of Calvert Watkins Innsbruck*: Institut für Sprachwissenschaft der Universität Innsbruck, 17-28.

Beekes, R. S. P. (2000). European substratum words in Greek, in Ofitsch, M. & Zinko, C. (eds.), *125 Jahre Indogermanistik in Graz Leykam*, 21-31.

Beekes, R. S. P. (2003). Indo-European or substrate? φάτην and κῆρυξ, in Bammesberger, A., Vennemann, T. *Languages in Prehistoric Europe*, Heidelberg: Universitätsverlag Winter, 109-115.

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.

Byrd, A. M. (2010). *Reconstructing Indo-European Syllabification*. Los Angeles: UCLA dissertation.

Clackson, J. (1994) *The Linguistic Relationship between Armenian and Greek*. Publications of the Philological Society (30), Blackwell.

Clackson, J. (2008) Classical Armenian, ch. 11 in Woodard, R. D., *The Ancient Languages of Asia Minor*, Cambridge: Cambridge University Press, 124-144.

Considine, P. (1979). A Semantic Approach to the Identification of Iranian Loanwords in Armenian. *Studies in Diachronic, Synchronic, and Typological Linguistics: Festschrift for Oswald Szemerényi on the Occasion of His 65th Birthday*, 213.

Davidson, L., & Noyer, R. (1997). Loan phonology in Huave: nativization and the ranking of faithfulness constraints. In *Proceedings of WCCFL15* (65-80). CSLI Publications.

DeLisi, J. L. (2015) *Epenthesis and Prosodic Structure in Armenian: A Diachronic Account*, (Doctoral dissertation).

Downing, L. J. (2008). Co-phonologies and morphological exponence in OT. In *3rd Exponence Network meeting and Workshop on Theoretical Morphology* (Vol. 4).

Godel, R. (1975). *An Introduction to the Study of Classical Armenian*. L. Reichert.

Hsu, B. & Jesney, K. (2016) Loanword Adaptation in Québec French: Evidence for Weighted Scalar Constraints, *Proceedings of the 34th Meeting of the West Coast Conference on Formal Linguistics*, University of Utah.

Hübschmann, H. (1897) *Armenische Grammatik. 1. Theil: Armenische Etymologie* (in German), Leipzig: Breitkopf & Härtel.

Hübschmann, H., & Meillet, A. (1981). *Altarmenisches Elementarbuch*. Heidelberg. reprint, New York.

Inkelas, S. (1998). The theoretical status of morphologically conditioned phonology: a case study of dominance effects. In *Yearbook of Morphology 1997* (121-155). Springer, Dordrecht.

Inkelas, S., & Zoll, C. (2000). *Reduplication as Morphological Doubling*. Manuscript, University of California, Berkeley and Massachusetts Institute of Technology.

Inkelas, S., & Orgun, C. O. (2003). Turkish stress: a review. *Phonology*, 20(1), 139-161.

Itô, J & Mester A. (1995). The Core-Periphery Structure of the Lexicon and Constraints on Reranking. *Papers in Optimality Theory,* (18), 181–209.

Itô, J & Mester A. (1998). *Markedness and Word Structure: OCP Effects in Japanese*, unpublished paper, University of California, Santa Cruz.

Itô, J & Mester A. (1999). The phonological lexicon. In Natsuko Tsujimura (ed.), *The Handbook of Japanese Linguistics*. Oxford: Blackwell, 62-100.

Itô, J & Mester A. (2001). Covert Generalizations in Optimality Theory: The Role of Stratal Faithfulness Constraints. *Studies in Phonetics, Phonology, and Morphology*, 7(2), 273-299.

Itô, J & Mester A. (2003). *Japanese Morphophonemics: Markedness and Word Structure*. Cambridge, Mass.: MIT Press.

Itô, J & Mester A. (2004). *The Phonological Lexicon*, 552–564, in McCarthy, John J (ed), *Optimality Theory in Phonology: A Reader.* Wiley-Blackwell.

Jaafar, S., & Raihan, S. (2012). Co-phonology vs. Indexed constraint theory: a case study of Perak dialect partial reduplication. *Journal of the Southeast Asian Linguistics Society*, (5), 97.

J̌ahukyan, G. (2010) *Hayeren stugabanakan baŕaran* [Armenian Etymological Dictionary] (in Armenian), Yerevan: Asoghik.

Kager, R. (1999). *Optimality theory*. Cambridge University Press.

Kassian, A. S. & Yakubovich, I. S. (2002). The reflexes of IE initial clusters in Hittite. In Shevoroshkin V. & Sidwell P. (eds.), *Anatolian Languages*, 10–48. Canberra: Association for the History of Language.

Kertész, Z. (2003). Vowel Harmony and the Stratified Lexicon of Hungarian. *The Odd Yearbook*, 7, 62-77.

Kim, R. I. (2018) The Prehistory of the Classical Armenian Weak Aorist. In *Acta Linguistica Petropolitana, Transactions of the Institute for Linguistic Studies*, 24(1), 86-136.

Kiparsky, P. (1968). Linguistic universals and linguistic change, in Emmon B. & Harms, R. (eds.), *Universals in Linguistic Theory*. Holt, Rinehart, and Winston. 170-202.

Kiparsky, P. (2010). Reduplication in Stratal OT. *Reality Exploration and Discovery: Pattern Interaction, Language & Life*, 125-142.

Korn, A. (2012). Final troubles: Armenian stem classes and the word-end in Late Old Persian. In Lur'e P. & Tokhtasev S. (eds.), *Commentationes Iranicae*, 74-92.

Korn, A. & Olsen, B.A. (2012) On Armenian -*agin*: additional evidence for a third West Middle Iranian dialect?, *Münchener Studien zur Sprachwissenschaft*, 66 (2), 201–220.

Kubozono, H., Itô, J., & Mester, A. (2008). Consonant gemination in Japanese loanword phonology. In *18th International Congress of Linguistics*, Seoul, 953-973.

Kuiper, F. B. J. (1995) Gothic *bagms* and Old Icelandic *ylgr*. *North-Western European Language Evolution*, vol. 25, 63-88.

Kurisu, K. (2011). *Cophonology in Sino-Japanese Vowel Harmony*. NELS39, 521-532.

Labrune, L. (2012). *The Phonology of Japanese*. Oxford University Press.

Lidén, E. (1933) Armeniaca. *Göteborgs Högskolas Årsskrift*. Göteborg: Wettergren & Kerber (Elander), vol. 39, 39-56.

Lubotsky, A. M. (2001). The Indo-Iranian Substratum. In Carpelan, C., Parpola, A. & Koskikallio, P. (eds.), *Early contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Papers presented at an international symposium held at the *Tvärminne Research Station of the University of Helsinki* (8-10 January, 1999). Helsinki: Suomalais-ugrilainen seura, 301-317.

Macak, M. (n.d.). *On the Affricate Dissimilation in Old Armenian Aorist Subjunctive*, unpublished paper.

Martirosyan, H. (2008) *Studies in Armenian Etymology with Special Emphasis on Dialects and Culture, Indo-European Heritage*, doctoral thesis, Universiteit Leiden.

McCarthy, J. J. & Prince A. (1993). Prosodic Morphology I: Constraint Interaction and Satisfaction. *Linguistics Department Faculty Publication Series,* Paper 14.

Meyer, R. (2015). Morphosyntactic Alignment, Pattern Replication, and the Classical Armenian Periphrastic Perfect. In Jamison, S. W., Melchert, H. C. & Vine, B. (eds.) (2015). *Proceedings of the 26th Annual UCLA Indo-European Conference*. Bremen: Hempen.

Meyer, R. (2017). *Iranian-Armenian language contact in and before the 5th century CE: An investigation into pattern replication and societal multilingualism* (Doctoral Dissertation).

Nercissians, E. (1987). The development of standard literary Armenian and Persian: a comparative study. *Journal of Pragmatics*, 11(5), 623-639.

Olsen, B. A. (1999) The Noun in Biblical Armenian: Origin and Word-Formation: With Special Emphasis on the Indo-European Heritage, *Trends in Linguistics. Studies and Monographs 119*, Berlin, New York: Mouton de Gruyter.

Orgun, C. O. (1996). *Sign-based morphology and phonology with special attention to Optimality Theory*, University of California at Berkeley (Doctoral dissertation).

Orgun, C. O. (1998). Cyclic and noncyclic phonological effects in a declarative grammar. In *Yearbook of Morphology 1997* (179-218). Springer, Dordrecht.

Pedersen, H. (1905) Zur armenischen Sprachgeschichte. *Zeitschrift für vergleichende Sprachforschung*, (38) 194-240.

Périkhanian, A. (1966) "Une inscription araméenne du roi Artašēs trouvée à Zanguézour (Siwnik')," *Revue des Études Arméniennes* 3, 17–29.

Perry, J. R. (2005). Lexical areas and semantic fields of Arabic loanwords in Persian and beyond. Linguistic Convergence and Areal Diffusion: *Case Studies from Iranian, Semitic, and Turkic*, 97-110.

Picard, M. (1989). A Reanalysis of Armenian Prothesis and Metathesis. *Folia Linguistica Historica*, 23(*Historica* vol. 10, 1-2), 61-70.

Picard, M. (1994). Principles and Methods in Historical Phonology: *From Proto-Algonkian to Arapaho*. McGill-Queen's Press-MQUP.

Pisowicz, A. (1995). How did New Persian and Arabic Words Penetrate the Middle Armenian Vocabulary? Remarks on the Material in Kostandin Erznkac 'i's Poetry'. *New Approaches to Medieval Armenian Language and Literature*, 3, 95.

Prince, A. (1998). Two lectures on Optimality Theory. In *Handout from Phonology Forum*.

Rasin, E. (2016). Morpheme Structure Constraints and Blocking in Nonderived Environments, *MIT Papers*.

Schmitt, L., & Bailey, H. W. (1987). Armenia and Iran IV. *Encyclopaedia Iranica II*, 450.

Schrijver, p. C. H. (1991) The reflexes of the Proto-Indo-European laryngeals in Latin. *Leiden Studies in Indo-European*, Amsterdam: Rodopi.

Shademan, S. (2002). *Epenthetic Vowel Harmony in Farsi*. University of California at Los Angeles (MA thesis).

Shibatani, M. (1990). *The languages of Japan*. Cambridge University Press.

Smolensky, P & Legendre, G. (2006). *The Harmonic Mind: From Neural Computation to Optimality Theoretic Grammar*. Cambridge, MA: MIT Press.

Szemerényi, O. (1985). Armenian between Iran and Greece, in Pieper, U. & Stickel, G. (eds.), *Studia linguistica, diachronica et synchronica Werner Winter sexagerio anno MCMLXXXIII*. Berlin: Mouton de Gruyter, 783-799.

Vaux, B. (1998). *The Phonology of Armenian*. Oxford: Oxford University Press.

Winter, W. (1954). Problems of Armenian Phonology I. *Language*, 30(2), 197-201.

Winter, W. (1955). Problems of Armenian Phonology II. *Language*, 31(1), 4-8.

Winter, W. (1962). Problems of Armenian phonology III. *Language*, 38(3), 254-262.

Winter, W. (2011) Armenian. In Gvozdanovic, J. (ed.). (2011). *Indo-European Numerals* (Vol. 57), Walter de Gruyter, 347-360.

Zukoff, S. (2012). *The Phonology of Verbal Reduplication in Ancient Greek: An Optimality Theory Approach* (Masters dissertation, University of Georgia).

Appendices

Legend for the word lists:

Lexical Layer by etymological authority:

1 – corresponds to Lex_1
2 – corresponds to Lex_2
3 – corresponds to Lex_3

Lexical layer by phonotactic pattern:

1 – corresponds to Lex_1
2 – corresponds to Lex_2
3 – corresponds to Lex_3
X – lexical layer cannot be predicted based on surface-oriented phonotactics

Morphological Derivability

(L)ow – save for a few rare instances like nonce words, further deriving the word is impossible
(M)edium – some derivations possible (up to ten)
(H)igh – ten up to many dozens or hundreds of derived words possible, showing a very high degree of integration

Prothesis:

(E)arly – oy, u (ow), ē, i, er
(L)ater – o, e, ṙ
(M)odern – ə
N)ot applicable

Metathesis:

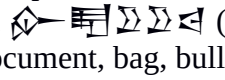
(Y)es – the word contains metathesis
(N)o – the word does not contain metathesis

Reduplication:

(F)ull – fully reduplicated forms possible
(P)artial – partially reduplicated forms possible
(N)one – no reduplicated forms possible

Appendix A: Lex$_1$

Transliteration	Armenian script	Lex layer by etym. auth.	Lex layer by phonotactics	Derivability	Prothesis	Metathesis	Reduplication	Notes
-ական	-akan	1	X	H	N	N	N	
-ան	-an	1	X	L	N	N	N	instrument suffix (19 words containing this suffix)
-անակ	-anak	1	X	H	N	N	N	noun-forming suffix; used in 6 words
-արան	-aran	1	X	H	N	N	N	place- or container-forming noun
-աւանդ	-awand	1	1 2 3	M	N	N	N	
-աւանդակ	-awandak	1	X	L	N	N	N	
-աւէտ	-awêt	1	X	M	N	N	N	
-եակ	-eak	1	X	H	N	N	N	
-եղէն	-ełên	1	X	H	N	N	N	
-էճ	-êč	1	X	L	N	N	N	diminutive
-իկ	-ik	1	X	H	N	N	N	(not the homophonous PIE -ik diminutive suffix)
-կար	-kar	1	X	H	N	N	N	
-կէն	-kên	1	X	M	N	N	N	
-ճան	-čan	1	X	H	N	N	N	
-ոյկ	-oyk	1	X	H	N	N	N	Etymology ‘probable’ but not certain
-ուհի	-owhi	1	X	H	N	N	N	Feminizing noun suffix
-ստան	-stan	1	X	H	N	N	N	Same ‘-stan’ suffix as in Kazakhstan.
ազբն	azbn	1	X	L	N	N	N	May be inherited from PIE
ազգ	azg	1	X	H	N	N	N	
աճառ	ačar̄	1	X	M	N	N	N	
անդրուար	androwar	1	X	M	N	N	N	
ապահարզան	apaharzan	1	X	L	N	N	N	
ապաւէն	apawên	1	X	M	N	N	N	
ասպական ի	aspakani	1	X	L	N	N	N	
ասպետ	aspet	1	X	H	N	N	N	
Ասպուրակ	Aspowrak	1	X	L	N	N	N	From Old Ossetic?

Էս	ê							
ավեր	aver	1	X	M	N	N	N	
Արամ	Aram	1	X	L	N	N	N	
Արամայիս	Aramayis	1	X	L	N	N	N	
Արամանեակ	Aramanea k	1	X	L	N	N	N	
Արիստակէս	Aristakês	1	X	L	N	N	N	
աւար	awar	1	X	M	N	N	N	
բամբիշն	bambišn	1	1 2 3	H	N	N	N	
բդեաշխ	bdeašx	1	1 2 3	L	N	Y	N	
գաւիթ	gawitʻ	1	X	L	N	N	N	
գոյն	goyn	1	X	H	N	N	N	
գուշակ	gowšak	1	X	L	N	N	N	
դանակ	danak	1	X	H	N	N	N	
դէտ	dêt	1	X	L	N	N	N	
դժ-	dž-	1	X	M	N	N	N	
դրոշմ	drošm	1	X	H	N	N	N	
դրուատ	drowat	1	X	L	N	N	N	
երախ	erax	1	X	L	E	N	N	
երախան	eraxan	1	X	L	E	N	N	
Երուազ	Erowaz	1	X	L	E	N	N	
զատիկ	zatik	1	X	H	N	N	N	
ընկոյզ	ənkoyz	1	X	M	M	N	N	
թագ	tʻag	1	X	H	N	N	N	
թագաւոր	tʻagawor	1	X	H	N	N	N	
թագուհի	tʻagowhi	1	X	H	N	N	N	
ժիր	žir	1	1 2 3	M	N	N	N	
խոնաստան	xonastan	1	1 2 3	L	N	N	N	
խորհ	xorh	1	1 2 3	H	N	Y	N	
ծիրան	ciran	1	X	H	N	N	N	Contested
կահոյր	kahoyr	1	X	L	N	N	N	
կապուտակ	kapowtak	1	X	L	N	N	N	
կնիք	knikʻ	1	X	H	N	N	N	Strongly disputed – could also be from Akkadian  (/kanīku/, ‘sealed object: document, bag, bulla’) or  (kunukku, “seal-cylinder”), perhaps via Hurrian (Hübschmann 1897:307)
հազարապետ	hazarapet	1	X	L	N	N	N	
հանապազ	hanapaz	1	X	L	N	N	N	Adverb

հասբ	hasb	1	X	L	N	N	N	
ճիւ	čiw	1	1 2 3	H	N	N	N	
մահիկ	mahik	1	X	L	N	N	N	
մատակարար	matakarar	1	X	L	N	N	N	
նժոյգ	nžoyg	1	X	L	N	N	N	
շարաւանդ	šarawand	1	1 2 3	L	N	N	N	
շճեմ	ščem	1	1 2 3	M	N	N	N	
պահ	pah	1	1 2 3	H	N	N	N	
պաշար	pašar	1	1 2 3	H	N	N	N	
պաշտպան	paštpan	1	1 2 3	H	N	N	N	
պատնէշ	patnêš	1	1 2 3	L	N	N	N	
պատշգամբ	patšgamb	1	1 2 3	L	N	N	N	
պատուար	patowar	1	1 2 3	L	N	N	N	
պարեգաւտք	paregawtkʻ	1	1 2 3	L	N	N	N	
պարիսպ	parisp	1	2 3	M	N	N	N	
պարկէն	parkên	1	1 2 3	L	N	N	N	
պարտ	part	1	1 2 3	H	N	N	N	
Պարտաւ	Partaw	1	1 2 3	L	N	N	N	
պարտէզ	partêz	1	1 2 3	H	N	N	N	
պղինձ	płinj	1	1 2 3	H	N	N	N	From Old Median
Սաթենիկ	Satʻenik	1	X	L	N	N	N	From Scythian
սաղաւարտ	saławart	1	1 2 3	L	N	N	N	
Սանատրուկ	Sanatrowk	1	X	L	N	N	N	
սեաւ	seaw	1	X	H	N	N	N	
սեպուհ	sepowh	1	X	H	N	N	N	
ստարան	staran	1	X	L	N	N	N	
Վահէ	Vahê	1	X	L	N	N	N	
վասն	vasn	1	X	L	N	N	N	Preposition
վարազ	varaz	1	X	H	N	N	N	
տատրակ	tatrak	1	X	L	N	N	N	
տար-	tar-	1	X	M	N	N	N	Prefix meaning ‘beyond’.
տէգ	têg	1	X	H	N	N	N	
Տիգրան	Tigran	1	X	L	N	N	N	
Փայտակարան	Pʻaytakaran	1	X	L	N	N	N	
օրիորդ	ôriord	1	1 2 3	L	L	N	N	

Appendix B: Lex₂

Transliteration	Armenian script	Lex layer by etym. auth.	Lex layer by phonotactics	Derivability	Prothesis	Metathesis	Reduplication	Notes
-ակ	-ak	2	X	M	N	N	N	
-ան	-an	2	X	H	N	N	N	collective suffix (3 words containing this suffix)
-անի	-ani	2	X	H	N	N	N	
-անօր	-anôr	2	X	M	N	N	N	Adverbial suffix indicating location. J̌ahukyan considers the origin uncertain.
-գին	-gin –	2	X	H	N	N	N	
-երէն	-erên	2	X	H	N	N	N	
-կերտ	-kert	2	1 2 3	H	N	N	N	Found also in wholly-borrowed terms
-պան	-pan	2	1 2 3	H	N	N	N	
-պանակ	-panak	2	1 2 3	M	N	N	N	
ազատ	azat	2	X	H	N	N	N	
ազդ	azd	2	2 3	H	N	N	N	
ազն	azn	2	X	H	N	N	N	
Աժդահակ	Aždahak	2	X	L	N	N	N	Mythological figure
ախոռ	axorˉ	2	X	H	N	N	N	
ախորժ	axorž	2	X	H	N	N	N	
ախտ	axt	2	2 3	H	N	N	N	
ակահ	akah	2	X	L	N	N	N	
ահոկ	ahok	2	X	L	N	N	N	
ամանակ	amanak	2	X	L	N	N	N	
ամբաստան	ambastan	2	1 2 3	L	N	N	N	
ամբար	ambar	2	1 2 3	M	N	N	N	
ամբարտակ	ambartak	2	1 2 3	L	N	N	N	
ամբոխ	ambox	2	1 2 3	H	N	N	N	
ամիճ	amič	2	X	H	N	N	N	
Անակ	Anak	2	X	L	N	N	N	Male given name derived from Middle Iranian *anāk ‘evil, bad’

Անահիտ	Anahit	2	X	L	N	N	N	Mythological figure (Armenian equivalent of Artemis or Venus)
անապատ	anapat	2	X	H	N	N	N	
անարի	anari	2	X	L	N	N	N	In the sense of ‘non-Aryan’, not ‘weak’ (negative ADJ an + ari)
անդամ	andam	2	X	H	N	N	N	
անդոհական	andohakan	2	X	L	N	N	N	
անդրավարտիք	andravartikʿ	2	X	L	N	N	N	
անձուկ	anjowk	2	X	L	N	N	N	Disputed – sometimes analyzed as from PIE *$h_2$éng^{h}us
անոյշ	anoyš	2	X	H	N	N	N	
անուշակ	anowšak	2	X	H	N	N	N	
աշակերտ	ašakert	2	1 2 3	H	N	N	N	
աշխատ	ašxat	2		H	N	N	N	
աշխար	ašxar	2		L	N	Y	N	From Manichaean Middle Persian ʾxšʾd *axšād ‘troubled, distressed; distress’
աշխարհ	ašxarh	2	1 2 3	H	N	Y	N	
աշխարհակալ	ašxarhakal	2	X	M	N	Y	N	From ašxarh + -a- + kalum. The whole formation is a calque of Middle Persian štr'dʾl /šahr-yār/, ‘lord, sovereign, ruler’, literally ‘holding the world’
աշխոյժ	ašxoyž	2	X	M	N	Y	N	
աշտանակ	aštanak	2	X	L	N	N	N	
ապ-	ap-	2	X	M	N	N	N	
ապագովեմ	apagovem	2	X	L	N	N	N	
ապակի	apak	2	X	H	N	N	N	
ապաշխարեմ	apašxarem	2	X	H	N	Y	N	
ապաստան	apastan	2	X	L	N	N	N	
ապատ	apat	2	X	M	N	N	N	
ապարանջան	aparanǰan	2	X	L	N	N	N	
ապարանք	aparankʿ	2	X	L	N	N	N	
ապարօշ	aparôš	2	X	L	N	N	N	
ապիրատ	apirat	2	X	M	N	N	N	
ապուռ	apowr̄	2	X	M	N	N	N	
ապսպարեմ	apsparem	2	X	H	N	N	N	
ապստամբ	apstamb	2	1 2 3	H	N	N	N	
առապար	arāpar	2	X	L	N	N	N	
առաւիր	arāwir	2	X	L	N	N	N	
Ասորեստան	Asorestan	2	X	L	N	N	N	
ասորի	asori	2	X	L	N	N	N	

ասպազԷն	aspazên	2	X	L	N	N	N	
ասպակ	aspak	2	2 3	L	N	N	N	
ասպահանիկ	aspahanik	2	2 3	L	N	N	N	
ասպահապետ	aspahapet	2	2 3	L	N	N	N	
ասպաստան	aspastan	2	2 3	L	N	N	N	
ասպատակ	aspatak	2	2 3	L	N	N	N	
ասպարապետ	asparapet	2	2 3	M	N	N	N	
ասպարԷզ	asparêz	2	2 3	H	N	N	N	
ասպնջական	aspnǰakan	2	2 3	L	N	N	N	
աստուճ	astowč	2	X	L	N	N	N	
ատրագոյն	atragoyn	2	X	L	N	N	N	
ատրճան	atrčan	2	X	L	N	N	N	
ատրճանակ	atrčanak	2	X	L	N	N	N	
արահետ	arahet	2	X	M	N	N	N	
Արամազդ	Aramazd	2	2 3	L	N	N	N	
արեւելք	arewelkʻ	2	X	H	N	N	N	
արժան	aržan	2	X	H	N	N	N	
արծաթ	arcatʻ	2	X	H	N	N	N	
արկանեմ	arkanem	2	X	M	N	N	N	
արհամարհ	arhamarh	2	1 2 3	M	N	Y	N	
Արհմն	Arhmn	2	X	L	N	Y	N	
արձակ	arjak	2	X	H	N	N	N	
արմաւ	armaw	2	X	M	N	N	N	
արշալուրշ	aršalowrš	2	X	L	N	N	N	
Արշակ	Aršak	2	X	L	N	N	N	
արուեստ	arowest	2	X	H	N	N	N	
Արուսեակ	Arowseak	2	X	L	N	N	N	
արջասպ	arǰasp	2	2 3	M	N	N	N	
արջն	arǰn	2	X	H	N	N	N	
Արտաշէս	Artašês	2	X	L	N	N	N	
Արտաշիր	Artašir	2	X	L	N	N	N	
Արտաւան	Artawan	2	X	L	N	N	N	
աւազ	awag	2	X	M	N	N	N	
աւազ	awaz	2	X	H	N	N	N	
աւազակ	awazak	2	X	H	N	N	N	
աւան	awan	2	X	M	N	N	N	
աւատ	awat	2	X	M	N	N	N	
աւերանք	awerankʻ	2	X	L	N	N	N	
Բագարատ	Bagarat	2	X	L	N	N	N	
բադ	bad	2	X	H	N	N	N	
բազմիմ	bazmim	2	X	H	N	N	N	
բազում	bazowm	2	X	H	N	N	N	

բազրիք	bazrikʻ	2	X	L	N	N	N	
բաժ	baž	2	X	H	N	N	N	
բաժակ	bažak	2	X	H	N	N	N	
բախտ	baxt	2	2 3	H	N	N	N	
բահ	bah	2	X	M	N	N	N	
բամբակ	bambak	2	1 2 3	H	N	N	N	
բանակ	banak	2	X	H	N	N	N	
բանտ	bant	2	X	H	N	N	N	
բաշխ	bašx	2	1 2 3	H	N	Y	N	
բարակ	barak	2	X	H	N	N	N	
բարգաւաճ	bargawač	2	X	L	N	N	N	
բարձ	barj	2	X	H	N	N	N	
բժիշկ	bžišk	2	2 3	H	N	N	N	
բնակ	bnak	2	X	H	N	N	N	
բոյր	boyr	2	X	H	N	N	N	
բովանդակ	bovandak	2	X	M	N	N	N	
բոր	bor	2	X	M	N	N	N	
բորենի	boreni	2	X	L	N	N	N	
բուծին	bowcin	2	X	L	N	N	N	
բուն	bown	2	X	H	N	N	N	
բրինձ	brinj	2	1 2 3	H	N	N	N	
գազան	gazan	2	X	H	N	N	N	
գահ	gah	2	X	H	N	N	N	
գաղափար	gałapʻar	2	X	H	N	N	N	
գանձ	ganj	2	1 2 3	H	N	N	N	
գանձաւոր	ganjawor	2	X	L	N	N	N	
գարիմ	garim	2	X	L	N	N	N	
գարհայեմ	garhayem	2	X	L	N	Y	N	
գարշապար	garšapar	2	X	L	N	N	N	
գարշիմ	garšim	2	X	L	N	N	N	
գերդաստան	gerdastan	2	X	H	N	N	N	
գերեզման	gerezman	2	X	H	N	N	N	
գԷս	gês	2	X	H	N	N	N	
գմբեթ	gmbetʻ	2	1 2 3	H	N	N	N	
գոհար	gohar	2	X	L	N	N	N	
գոմԷշ	gomêš	2	X	L	N	N	N	
գոռ	gorʻ	2	X	L	N	N	N	
գունդ	gownd	2	1 2 3	M	N	N	N	
գուսան	gowsan	2	X	L	N	N	N	
գրապան	grapan	2	X	H	N	N	N	
գրաւ	graw	2	X	H	N	N	N	
գրիւ	griw	2	X	H	N	N	N	

գրտանակ	grtanak	2	X	H	N	N	N	
դաժան	dažan	2	X	H	N	N	N	
դահիճ	dahič	2	X	H	N	N	N	
դահլիճ	dahlič	2	X	L	N	N	N	
դայեակ	dayeak	2	X	L	N	N	N	
դան	dan	2	X	H	N	N	N	
դանգ	dang	2	X	L	N	N	N	
դանդանաւան դ	dandanawand	2	1 2 3	L	N	N	N	
դաշն	dašn	2	X	L	N	N	N	
դաշնակ	dašnak	2	X	H	N	N	N	
դաշոյն	dašoyn	2	X	M	N	N	N	
դաշտան	daštan	2	X	L	N	N	N	
դաս	das	2	X	H	N	N	N	
դաստակ	dastak	2	X	M	N	N	N	
դաստակերտ	dastakert	2	1 2 3	L	N	N	N	
դաստիարակ	dastiarak	2	X	L	N	N	N	
դատ	dat	2	X	H	N	N	N	
դատախազ	dataxaz	2	X	L	N	N	N	
դատաւոր	datawor	2	X	H	N	N	N	
Դատոյ	Datoy	2	X	L	N	N	N	
Դատոյեան	Datoyean	2	X	L	N	N	N	
դարգ	darg	2	X	L	N	N	N	
Դարեհ	Dareh	2	X	L	N	N	N	
դարիճենիկ	daričenik	2	X	L	N	N	N	
դաւաճան	dawačan	2	X	H	N	N	N	
դեհ	deh	2	X	M	N	N	N	
դեն	den	2	X	M	N	N	N	
դեսպան	despan	2	X	M	N	N	N	
դերձակ	derjak	2	X	L	N	N	N	
դեւ	dew	2	X	H	N	N	N	
դԷզ	dêz	2	X	M	N	N	N	
դԷմ	dêm	2	X	H	N	N	N	
դԷպք	dêpkʻ	2	X	H	N	N	N	
դԷսպակ	dêspak	2	X	L	N	N	N	
դժոխ	džox	2	X	H	N	N	N	
դժուար	džowar	2	X	H	N	N	N	
դիպակ	dipak	2	X	L	N	N	N	
դիպահ	dipah	2	X	L	N	N	N	
դիտակ	ditak	2	X	H	N	N	N	
դիւան	diwan	2	X	L	N	N	N	
դղեակ	dłeak	2	X	L	N	N	N	

դոյլ	doyl	2	X	H	N	N	N	Alternatively from Classical Syriac ܕܠܐ (dawlā’, “pail, bucket; Aquarius”).
դպիր	dpir	2	X	H	N	N	N	
դսրով	dsrov	2	X	L	N	N	N	
դրախտ	draxt	2	2 3	H	N	N	N	
դրամ	dram	2	X	H	N	N	N	
Դրասխանակերտ	Drasxanakert	2	1 2 3	L	N	N	N	
դրուագ	drowag	2	X	L	N	N	N	
դրօշ	drôš	2	X	M	N	N	N	
երագ	erag	2	X	L	E	N	N	
երազ	eraz	2	X	H	E	N	N	
երախտիք	eraxtikʻ	2	2 3	L	E	N	N	
երակ	erak	2	X	M	E	N	N	
երամ	eram	2	X	H	E	N	N	
երամակ	eramak	2	X	L	E	N	N	
երանգ	erang	2	X	M	E	N	N	
երաշխ	erašx	2	1 2 3	M	E	Y	N	
երաշխիք	erašxikʻ	2	X	L	E	Y	N	
երաշտ	erašt	2	2 3	L	E	N	N	
երասան	erasan	2	X	M	E	N	N	
Երասխ	Erasx	2	X	L	E	N	N	
Երաստ	Erast	2	X	L	E	N	N	
երիվար	erivar	2	X	L	E	N	N	
երիտասարդ	eritasard	2	1 2 3	H	E	N	N	
երկ	erk	2	2 3	H	E	N	N	
զամբիւղ	zambiwł	2	1 2 3	L	N	N	N	
զանազան	zanazan	2	X	L	N	N	F	
զանգապան	zangapan	2	X	L	N	N	N	
զառամ	zarām	2	X	L	N	N	N	
Զարուհի	Zarowhi	2	X	L	N	N	N	
զարտագոյն	zartagoyn	2	X	H	N	N	N	
Զաւէն	Zawên	2	X	L	N	N	N	
զեան	zean	2	X	L	N	N	N	
զենում	zenowm	2	X	L	N	N	N	
զէն	zên	2	X	H	N	N	N	
զմրուխտ	zmrowxt	2	2 3	L	N	N	N	
զնդան	zndan	2	X	L	N	N	N	
զոհ	zoh	2	X	H	N	N	N	
զուարակ	zowarak	2	X	L	N	N	N	
զուր	zowr	2	X	M	N	N	N	

զրահ	zrah	2	X	H	N	N	N	
զօր	zôr	2	X	H	N	N	N	
թակոյկ	tʿakoyk	2	X	M	N	N	N	
թաշկինակ	tʿaškinak	2	X	M	N	N	N	
թմբուկ	tʿmbowk	2	1 2 3	M	N	N	N	
թշնամի	tʿšnami	2	X	H	N	N	N	
թշուառ	tʿšowarə̄	2	X	M	N	N	N	
թոշակ	tʿošak	2	X	M	N	N	N	
թութ	tʿowtʿ	2	X	L	N	N	N	
ժահր	žahr	2	1 2 3	L	N	N	N	
ժաման	žaman	2	1 2 3	M	N	N	N	
ժամանակ	žamanak	2	1 2 3	H	N	N	N	
ժանգ	žang	2	1 2 3	H	N	N	N	
ժառանգ	žarə̄ang	2	1 2 3	M	N	N	N	
ժպիրհ	žpirh	2	1 2 3	M	N	Y	N	
իշխան	išxan	2	X	H	N	Y	N	
իր	ir	2	X	H	N	N	N	Disputed - because the Persian is etymologically unexplained, Olsen (1999) allows the possibility that for once the loan went in the opposite direction, i.e. that the Iranian word is borrowed from Armenian
լորամարգ	loramarg	2	N/A	L	N	N	N	Only -marg is Persian
խազ	xaz	2	1 2 3	M	N	N	N	
խարբուզ	xarbowz	2	1 2 3	L	N	N	N	
խաւար	xawar	2	1 2 3	H	N	N	N	
խոհ	xoh	2	1 2 3	H	N	N	N	
խոյր	xoyr	2	1 2 3	M	N	N	N	
խոնարհ	xonarh	2	1 2 3	H	N	Y	N	
խոշտանգեմ	xoštangem	2	1 2 3	L	N	N	N	
խոստանամ	xostanam	2	1 2 3	H	N	N	N	
Խոսրով	Xosrov	2	1 2 3	L	N	N	N	
խորակ	xorak	2	1 2 3	L	N	N	N	
խորան	xoran	2	1 2 3	H	N	N	N	
խորտակեմ	xortakem	2	1 2 3	M	N	N	N	
խորտիկ	xortik	2	1 2 3	L	N	N	N	
խուճապ	xowčap	2	1 2 3	L	N	N	N	
խունկ	xownk	2	1 2 3	M	N	N	N	
խրատ	xrat	2	1 2 3	H	N	N	N	
խոզ	xôz	2	1 2 3	H	N	N	N	
խօս	xôs	2	1 2 3	H	N	N	N	
կախ	kax	2	X	H	N	N	N	

կախարդ	kaxard	2	1 2 3	H	N	N	N	
կաղամբ	kałamb	2	1 2 3	H	N	N	N	
կաղապար	kałapar	2	X	L	N	N	N	Perhaps from Ancient Greek καλοπόδιον
կամ	kam	2	X	L	N	N	N	In the sense of ‘threshing sledge’
կանեփ	kanepʻ	2	X	H	N	N	N	
կաչաղակ	kačʻałak	2	X	L	N	N	N	
կապանք	kapankʻ	2	X	M	N	N	N	
կապարճ	kaparč	2	X	M	N	N	N	
կապիկ	kapik	2	X	H	N	N	N	Middle Persian from Sanskrit
կապիճ	kapič	2	X	L	N	N	N	
կապոյտ	kapoyt	2	X	H	N	N	N	
կասկ	kask	2	X	H	N	N	N	
կատակ	katak	2	X	H	N	N	N	
կատակագուսան	katakagowsan	2	X	L	N	N	N	
կար	kar	2	X	H	N	N	N	
կարաս	karas	2	X	L	N	N	N	
կարաւան	karawan	2	X	M	N	N	N	
կարաւանդք	karawandkʻ	2	1 2 3	L	N	N	N	Plural form only
Կարէն	Karên	2	X	L	N	N	N	
կարիճ	karič	2	X	L	N	N	N	
կարճ	karč	2	X	H	N	N	N	
կարմիր	karmir	2	X	H	N	N	N	
կարշն	karšn	2	X	M	N	N	N	
կերպ	kerp	2	X	H	N	N	N	
կէս	kês	2	X	H	N	N	N	
կէտ	kêt	2	X	H	N	N	N	
կինճ	kinč	2	X	M	N	N	N	
կիրթ	kirtʻ	2	X	H	N	N	N	
կիրճ	kirč	2	X	L	N	N	N	
կնգմենի	kngmeni	2	X	L	N	N	N	
կնդրուկ	kndrowk	2	X	L	N	N	N	
կոյս	koys	2	X	H	N	N	N	
կոյտ	koyt	2	X	H	N	N	N	
կոյր	koyr	2	X	H	N	N	N	
կոռտիկ	korṙtik	2	X	L	N	N	N	
կոտակ	kotak	2	X	L	N	N	N	
կուժ	kowž	2	X	M	N	N	N	
կրակ	krak	2	X	H	N	N	N	
կրպակ	krpak	2	X	M	N	N	N	
կօշիկ	kôšik	2	X	M	N	N	N	

հազար	hazar	2	X	H	N	N	N	
հաճ	hač	2	X	H	N	N	N	
համ-	ham-	2	X	H	N	N	N	
համար	hamar	2	X	H	N	N	N	
համբար	hambar	2	1 2 3	H	N	N	N	
համբոյր	hamboyr	2	1 2 3	H	N	N	N	
համհարզ	hamharz	2	2, 3	L	N	N	N	
համոզեմ	hamozem	2	X	L	N	N	N	
հանգամանք	hangamankʻ	2	X	L	N	N	N	
հանդերձապետ	handerjapet	2	X	L	N	N	N	Borrowed in its derived form already from Middle Persian hndlcpt' /handarzbed/, ‘chancellor’
հանդէս	handês	2	X	H	N	N	N	
հանճար	hančar	2	X	M	N	N	N	
հաշտ	hašt	2	2 3	H	N	N	N	
հասարակ	hasarak	2	X	H	N	N	N	
հարազատ	harazat	2	X	M	N	N	N	
հարկ	hark	2	2 3	H	N	N	N	
հաւան	hawan	2	X	M	N	N	N	
հաւասար	hawasar	2	X	H	N	N	N	
հաւաստ	hawast	2	X	M	N	N	N	
հաւատ	hawat	2	X	H	N	N	N	
հեշտ	hešt	2	2 3	L	N	N	N	
հզօր	hzôr	2	X	H	N	N	N	
հէն	hên	2	X	H	N	N	N	
հմայք	hmaykʻ	2	X	H	N	N	N	
հնազանդ	hnazand	2	1 2 3	H	N	N	N	
հնար	hnar	2	X	H	N	N	N	
Հնարակերտ	Hnarakert	2	1 2 3	L	N	N	N	
Հրազդան	Hrazdan	2	2 3	L	N	N	N	
հրամայեմ	hramayem	2	2 3	H	N	N	N	
հրաման	hraman	2	2 3	H	N	N	N	
հրամատար	hramatar	2	2 3	L	N	N	N	
հրաշ	hraš	2	2 3	H	N	N	N	
հրաշակերտ	hrašakert	2	1 2 3	M	N	N	N	
հրապար	hrapar	2	2 3	M	N	N	N	
հրասախ	hrasax	2	2 3	L	N	N	N	
Հրասեակ	Hraseak	2	2 3	L	N	N	N	
հրեշտակ	hreštak	2	2 3	H	N	N	N	
հրովարտակ	hrovartak	2	2 3	L	N	N	N	
հրուանդան	hrowandan	2	2 3	L	N	N	N	
Հրուդէն	Hrowdên	2	2 3	L	N	N	N	

ձագ	jag	2	X	H	N	N	N	
ձերբակալ	jerbakal	2	X	L	N	N	N	Literally “caught by the hand”, from *ձերբ (*jerb, an old instrumental of jeṙn) + -ա- (-a-) + կալում (kalum), a calque of Middle Persian dstglwb' (*dast-graw, ‘captivity’, literally ‘caught by the hand’)
ճակատ	čakat	2	X	H	N	N	N	
ճակնդեղ	čakndeł	2	X	L	N	N	N	
ճահուկ	čahowk	2	X	L	N	N	N	
ճաղատ	čałat	2	X	L	N	N	N	
ճամբրուկ	čambrowk	2	1 2 3	L	N	N	N	
ճանապարհ	čanaparh	2	1 2 3	H	N	Y	N	
ճանդան	čandan	2	1 2 3	L	N	N	N	
ճանդարի	čandari	2	1 2 3	L	N	N	N	
ճանկ	čank	2	1 2 3	M	N	N	N	
ճաշ	čaš	2	1 2 3	H	N	N	N	
ճաշակ	čašak	2	1 2 3	H	N	N	N	
ճատրակ	čatrak	2	1 2 3	L	N	N	N	Middle Persian from Sanskrit
ճարակ	čarak	2	1 2 3	L	N	N	N	
ճարմանդ	čarmand	2	1 2 3	L	N	N	N	
ճարպ	čarp	2	1 2 3	H	N	N	N	
ճարտուկ	čartowk	2	1 2 3	L	N	N	N	
ճերմակ	čermak	2	1 2 3	H	N	N	N	
ճիշդ	čišd	2	1 2 3	H	N	N	N	
ճշմարիտ	čšmarit	2	1 2 3	H	N	N	N	
ճրագ	črag	2	1 2 3	H	N	N	N	
մազդեզն	mazdezn	2	X	L	N	N	N	
մակոյկ	makoyk	2	X	L	N	N	N	
մահակ	mahak	2	X	L	N	N	N	
մանուշակ	manowšak	2	X	L	N	N	N	
մատակ	matak	2	X	L	N	N	N	
մատեան	matean	2	X	H	N	N	N	
մար	mar	2	X	H	N	N	N	
մարախ	marax	2	X	L	N	N	N	
մարգ	marg	2	X	L	N	N	N	
մարգարէ	margarê	2	X	H	N	N	N	
մարգարիտ	margarit	2	X	H	N	N	N	
մարզ	marz	2	2 3	H	N	N	N	
մարզպան	marzpan	2	X	L	N	N	N	
մէգ	mêg	2	X	H	N	N	N	Possibly from PIE
Միհրան	Mihran	2	2 3	L	N	N	N	

միշտ	mišt	2	2 3	H	N	N	N	
միրգ	mirg	2	X	L	N	N	N	
մշակ	mšak	2	X	H	N	N	N	
մշտիկ	mštik	2	X	H	N	N	N	
մոգ	mog	2	X	M	N	N	N	
մոգպետ	mogpet	2	X	L	N	N	N	
մոզանամ	mozanam	2	X	L	N	N	N	
մոմ	mom	2	X	H	N	N	N	
մոյկ	moyk	2	X	L	N	N	N	
մովպետ	movpet	2	X	L	N	N	N	
մուճակ	mowčak	2	X	M	N	N	N	
մուրհակ	mowrhak	2	X	L	N	Y	N	
յազեմ	yazem	2	X	L	N	N	N	
յակունդ	yakownd	2	1 2 3	L	N	N	N	
յանդիման	yandiman	2	X	H	N	N	N	
յատակ	yatak	2	X	L	N	N	N	
յաւէժ	yawêž	2	X	L	N	N	N	
յաւէտ	yawêt	2	X	L	N	N	N	
յոյզ	yoyz	2	X	H	N	N	N	
յովազ	yovaz	2	X	L	N	N	N	
յօսեմ	yôsem	2	X	L	N	N	N	
նազիմ	nazim	2	X	H	N	N	N	
նաժիշտ	nažišt	2	2 3	L	N	N	N	
նախ	nax	2	X	H	N	N	N	
նախարար	naxarar	2	X	H	N	N	N	
նախճիր	naxčir	2	X	L	N	N	N	
նամակ	namak	2	X	H	N	N	N	
նապաստակ	napastak	2	X	L	N	Y	N	
նաւ	naw	2	X	H	N	N	N	Disputed.
նաւազ	nawaz	2	X	H	N	N	N	
նաւակատիք	nawakatikʻ	2	X	L	N	N	N	
նաւաստի	nawasti	2	X	L	N	N	N	
նաւթ	nawtʻ	2	X	M	N	N	N	
Ներսեհ	Nerseh	2	X	L	N	N	N	
ներքինի	nerkʻini	2	X	M	N	N	N	
նժդեհ	nždeh	2	X	M	N	N	N	
նիզակ	nizak	2	X	L	N	N	N	
նիհար	nihar	2	X	H	N	N	N	
նիշ	niš	2	X	H	N	N	N	
նիրհ	nirh	2	1 2 3	H	N	Y	N	
նկար	nkar	2	X	H	N	N	N	
նկուն	nkown	2	X	L	N	N	N	

նկրտեմ	nkrtem	2	X	L	N	N	N	
նհանգ	nhang	2	X	L	N	N	N	
նման	nman	2	X	H	N	N	N	
նշան	nšan	2	X	H	N	N	N	
նշխար	nšxar	2	X	H	N	Y	N	
նոխազ	noxaz	2	X	M	N	N	N	
նոճ	noč	2	X	L	N	N	N	
նուագ	nowag	2	X	H	N	N	N	
նուարտան	nowartan	2	X	L	N	N	N	
նուէր	nowêr	2	X	H	N	N	N	
նուիրակ	nowirak	2	X	L	N	N	N	
նպատակ	npatak	2	X	H	N	N	N	
նօթճեմ	nôtʿčem	2	X	L	N	N	N	
շահանշահ	šahanšah	2	1 2 3	L	N	Y	N	
շահապ	šahap	2	1 2 3	L	N	Y	N	
շահաստան	šahastan	2	1 2 3	L	N	Y	N	From Old Persian 𐎧𐏁𐏂𐎱𐎠𐎺𐎠 (xšaçapāvā)
շահդանակ	šahdanak	2	1 2 3	L	N	Y	N	
Շահրապղական	Šahrapłakan	2	1 2 3	L	N	Y	N	
շամղիտակ	šamłitak	2	1 2 3	L	N	N	N	
շապիկ	šapik	2	1 2 3	L	N	N	N	
Շապուհ	Šapowh	2	1 2 3	L	N	N	N	
Շաւասպ	Šawasp	2	2 3	L	N	N	N	
շէն	šên	2	1 2 3	H	N	N	N	
շիճուկ	šičowk	2	1 2 3	M	N	N	N	
շիշ	šiš	2	1 2 3	L	N	N	N	
շնորհ	šnorh	2	1 2 3	H	N	Y	N	
շուք	šowkʿ	2	1 2 3	H	N	N	N	
շպետ	špet	2	1 2 3	M	N	N	N	
շտեմարան	štemaran	2	1 2 3	L	N	N	N	
շտրպաղանգ	štrpałang	2	1 2 3	L	N	N	N	
ոյժ	oyž	2	1 2 3	H	E	N	N	
ոստան	ostan	2	1 2 3	L	L	N	N	
ոստիկան	ostikan	2	X	H	L	N	N	
ուխտ	owxt	2	2 3	H	E	N	N	
ունակ	ownak	2	X	M	E	N	N	
ուշ	owš	2	X	H	E	N	N	
ուրախ	owrax	2	X	H	E	N	N	
ուրուկ	owrowk	2	X	L	E	N	N	
պահապան	pahapan	2	1 2 3	H	N	N	N	
պաճէն	pačên	2	1 2 3	L	N	N	N	

պայիկ	payik	2	1 2 3	L	N	N	N	
պայման	payman	2	1 2 3	H	N	N	N	
պայուսակ	payowsak	2	1 2 3	H	N	N	N	
պայքար	paykʻar	2	1 2 3	H	N	N	N	
պանիր	panir	2	1 2 3	H	N	N	N	
պաշտեմ	paštem	2	2 3	H	N	N	N	
պաշտօն	paštôn	2	1 2 3	H	N	N	N	
պապ	pap	2	1 2 3	L	N	N	N	In the sense of ‘grandfather’, not ‘pope’, which is from Byzantine Greek παπάς
պառաւ	parāw	2	1 2 3	L	N	N	N	
պատանդ	patand	2	1 2 3	L	N	N	N	
պատառ	patarʀ̄	2	1 2 3	H	N	N	N	
պատասխանի	patasxani	2	1 2 3	H	N	N	N	
պատգամ	patgam	2	1 2 3	H	N	N	N	
պատերազմ	paterazm	2	2 3	H	N	N	N	
պատիժ	patiž	2	1 2 3	H	N	N	N	
պատիւ	patiw	2	1 2 3	H	N	N	N	
պատկեր	patker	2	1 2 3	H	N	N	N	
պատճառ	patčarʀ̄	2	1 2 3	H	N	N	N	
պատճէն	patčên	2	1 2 3	H	N	N	N	
պատմուճակ	patmowčak	2	1 2 3	L	N	N	N	
պատշաճ	patšač	2	1 2 3	L	N	N	N	
պատուանդան	patowandan	2	1 2 3	L	N	N	N	
պատուհան	patowhan	2	1 2 3	M	N	N	N	
պատուհաս	patowhas	2	1 2 3	M	N	N	N	
պատրաստ	patrast	2	1 2 3	H	N	N	N	
պատրաստեմ	patrastem	2	1 2 3	H	N	N	N	
պատրուակ	patrowak	2	1 2 3	L	N	N	N	
պարաւանդ	parawand	2	1 2 3	L	N	N	N	
պարգեւ	pargew	2	1 2 3	H	N	N	N	
պարզ	parz	2	2 3	H	N	N	N	
պարիկ	parik	2	1 2 3	L	N	N	N	
պարկեշտ	parkešt	2	2 3	L	N	N	N	
պարսիկ	parsik	2	1 2 3	M	N	N	N	
պետ	pet	2	1 2 3	H	N	N	N	
պէս	pês	2	1 2 3	H	N	N	N	
պէտ	pêt	2	1 2 3	H	N	N	N	
պինդ	pind	2	1 2 3	H	N	N	N	
պիսակ	pisak	2	1 2 3	L	N	N	N	
պիստակ	pistak	2	1 2 3	M	N	N	N	

պողովատ	połovat	2	1 2 3	L	N	N	N	
պուրակ	powrak	2	1 2 3	L	N	N	N	From Middle Persian from Sanskrit
պսակ	psak	2	1 2 3	H	N	N	N	
ջատագով	ǰatagov	2	X	L	N	N	N	
ջոկ	ǰok	2	X	H	N	N	FP	
ռազմ	rāzm	2	2 3	H	L	N	N	
ռահ	rāh	2	X	M	L	N	N	
ռոճիկ	rōčik	2	X	L	L	N	N	
սագճիկ	sagčik	2	X	L	N	N	N	
սահման	sahman	2	X	H	N	N	N	
սաղար	sałar	2	X	M	N	N	N	
սամոյր	samoyr	2	X	L	N	N	N	
սանդ	sand	2	1 2 3	L	N	N	N	
սանդարամետ	sandaramet	2	X	L	N	N	N	
սար-	sar-	2	X	H	N	N	N	
սարեակ	sareak	2	X	L	N	N	N	
սափոր	sapʿor	2	X	L	N	N	N	
սեփական	sepʿakan	2	X	M	N	N	N	
սիրամարգ	siramarg	2	X	L	N	N	N	
սկայ	skay	2	X	H	N	N	N	
սմբակ	smbak	2	2 3	H	N	N	N	
սմպատակ	smpatak	2	2 3	L	N	N	N	
սնգոյր	sngoyr	2	X	M	N	N	N	
սոխ	sox	2	X	H	N	N	N	
սոխակ	soxak	2	X	L	N	N	N	
սով	sov	2	X	H	N	N	N	
սուգ	sowg	2	X	H	N	N	N	
սուսեր	sowser	2	X	L	N	N	N	
սուրբ	sowrb	2	X	H	N	N	N	
սպահ	spah	2	X	H	N	N	N	
սպայ	spay	2	X	M	N	N	N	
սպայապետ	spayapet	2	X	L	N	N	N	
սպառ	sparˉ	2	X	L	N	N	N	
սպաս	spas	2	X	H	N	N	N	
սպարապետ	sparapet	2	X	L	N	N	N	
սպիտակ	spitak	2	X	H	N	N	N	
սպուժեմ	spowžem	2	X	L	N	N	N	
ստամբակ	stambak	2	1 2 3	L	N	N	N	
ստուար	stowar	2	X	L	N	N	N	
սրահ	srah	2	X	H	N	N	N	
սրահակ	srahak	2	X	L	N	N	N	
վագր	vagr	2	X	L	N	N	N	Middle Persian from Sanskrit?

վազ	vaz	2	X	H	N	N	N	
Վահագն	Vahagn	2	X	L	N	N	N	
վահան	vahan	2	X	H	N	N	N	
Վահրամ	Vahram	2	2 3	L	N	N	N	
վաճառ	vačarə̄	2	X	H	N	N	N	
վաճառական	vačarə̄kan	2	X	L	N	N	N	
վայր	vayr	2	X	H	N	N	N	
վանգ	vang	2	X	H	N	N	N	
վանդակ	vandak	2	1 2 3	H	N	N	N	-ak is a suffix borrowed directly from Middle Persian
վանեմ	vanem	2	X	M	N	N	N	
վատ	vat	2	X	H	N	N	N	
վատթար	vattʿar	2	X	L	N	N	N	Comparative form of adjective calqued from Middle Persian
վարագոյր	varagoyr	2	X	M	N	N	N	
վարդ	vard	2	1 2 3	H	N	N	N	
Վարդան	Vaan	2	X	L	N	N	N	
վարեմ	varem	2	X	H	N	N	N	
վարզ	varz	2	2 3	L	N	N	N	
վարժ	varž	2	X	H	N	N	N	
վարձ	varj	2	X	H	N	N	N	
վարձակ	varjak	2	X	L	N	N	N	
վարշամակ	varšamak	2	X	L	N	N	N	
վարուժան	varowžan	2	X	M	N	N	N	
վարունգ	varowng	2	X	L	N	N	N	
վարս	vars	2	X	H	N	N	N	
վարտիք	vartikʿ	2	X	L	N	N	N	
վաւեր	vawer	2	X	L	N	N	N	
վեհ	veh	2	X	H	N	N	N	
վերմակ	vermak	2	X	L	N	N	N	
վզուրկ	vzowrk	2	2 3	L	N	N	N	
վէգ	vêg	2	X	L	N	N	N	
վէմ	vêm	2	X	H	N	N	N	
վէպ	vêp	2	X	H	N	N	N	
վիզ	viz	2	X	H	N	N	N	
վիժակ	vižak	2	X	L	N	N	N	
վիժեմ	vižem	2	X	H	N	N	N	
վին	vin	2	X	H	N	N	N	
վիշապ	višap	2	X	H	N	N	N	
վիշտ	višt	2	2 3	H	N	N	N	
վիրամ	viram	2	X	M	N	N	N	
վկայ	vkay	2	X	H	N	N	N	

վկայեմ	vkayem	2	X	H	N	N	N	
վճար	včar	2	2 3	H	N	N	N	
վճարեմ	včarem	2	2 3	H	N	N	N	
վճիռ	včiṙ	2	2 3	H	N	N	N	
վճիտ	včit	2	2 3	L	N	N	N	
վնաս	vnas	2	X	H	N	N	N	
վտանգ	vtang	2	X	H	N	N	N	
վրէժ	vrêž	2	X	H	N	N	N	
վրէպ	vrêp	2	X	H	N	N	N	
տագնապ	tagnap	2	X	H	N	N	N	
տախտ	taxt	2	2 3	H	N	N	N	
տախտակ	taxtak	2	X	H	N	N	N	
տակառ	takaṙ	2	X	L	N	N	N	
տաղաւար	taławar	2	X	L	N	N	N	
Տաճատ	Tačat	2	X	L	N	N	N	
տաճար	tačar	2	X	H	N	N	N	
տաճիկ	tačik	2	X	L	N	N	N	
տանջ	tanǰ	2	1 2 3	H	N	N	N	
տաշեմ	tašem	2	X	H	N	N	N	
տապ	tap	2	X	M	N	N	N	
տապակ	tapak	2	X	H	N	N	N	
տարափ	tarapʻ	2	X	H	N	N	N	
տաւիղ	tawił	2	X	L	N	N	N	
տզրուկ	tzrowk	2	X	L	N	N	N	
տոհմ	tohm	2	X	L	N	N	N	
տոյժ	toyž	2	X	L	N	N	N	
Տրդատ	Tat	2	X	L	N	N	N	
տօթ	tôtʻ	2	X	H	N	N	N	
ցանց	cʻancʻ	2	N/A	H	N	N	N	From Alanic?
փառք	pʻarǩʻ	2	X	H	N	N	N	
փարթամ	pʻartʻam	2	X	H	N	N	N	
փիղ	pʻił	2	X	H	N	N	N	
փլատակ	pʻlatak	2	X	L	N	N	N	
փուշտիպան	pʻowštipan	2	X	L	N	N	N	
փուտ	pʻowt	2	X	H	N	N	N	
քաղաք	kʻałakʻ	2	X	H	N	N	N	
քանդակ	kʻandak	2	X	H	N	N	N	
քանոն	kʻanon	2	X	M	N	N	N	
քարշ	kʻarš	2	X	H	N	N	N	
քարտակատեմ	kʻartakatem	2	X	M	N	N	N	
քէն	kʻên	2	X	H	N	N	N	Possibly from PIE *kʷoynéh₂

քուռակ	kʻowrāk	2	X	L	N	N	N	
քուրձ	kʻowrj	2	X	M	N	N	N	
քսակ	kʻsak	2	X	M	N	N	N	
քրքում	kʻrkʻowm	2	X	M	N	N	P	
օգն	ôgn	2	X	H	L	N	N	
օգուտ	ôgowt	2	X	H	L	N	N	
օժանդակ	ôžandak	2	X	M	L	N	N	
օրէն	ôrên	2	X	H	L	N	N	
օրինակ	ôrinak	2	X	H	L	N	N	
օրհնեմ	ôrhnem	2	X	H	L	Y	N	

Appendix C: Lex₃

Transliteration	Armenian script	Lex layer by etym. auth.	Lex layer by phonotactics	Derivability	Prothesis	Metathesis	Reduplication	Notes
ամպ	amp	3	X	L	N	N	N	Not ‘cloud’ (homophonous and homographous word but derived from PIE) but ‘sponge’ from pre-Modern Persian
գօռ	gôr̄	3	X	L	N	N	N	
թափուր	tʿapʿowr	3	X	L	N	N	N	
խիշտ	xišt	3	2 3	L	N	N	N	
հասբ	hasb	3	X	L	N	N	N	Maybe Northern Kuish?
փող	pʿoł	3	X	H	N	N	N	Uncertain
քիմիա	kʿimia	3	X	M	N	N	N	
քիչ	kʿičʿ	3	X	H	N	N	N	

Appendix D: Lexemes of externally uncertain origin

Transliteration	Armenian script	Lex layer by etym. auth.	Lex layer by phonotactics	Derivability	Prothesis	Metathesis	Reduplication	Notes
Աշխէն	Ašxên	1?	X	L	N	Y	N	
աշխէտ	ašxêt	1?	X	L	N	Y	N	
աշտեայ	ašteay	1?	X	L	N	N	N	
ասպար	aspar	1?	2 3	H	N	N	N	
ատակ	atak	1?	X	L	N	N	N	
ատրուշան	atrowšan	1?	X	L	N	N	N	
արագ	arag	1?	X	H	N	N	N	
արոյր	aroyr	1?	X	H	N	N	N	
արտախոյր	artaxoyr	1?	X	L	N	N	N	
բահուանդ	bahowand	1?	1 2 3	M	N	N	N	
բիւր	biwr	1?	X	H	N	N	N	
բուրւառ	bowrwarə̄	1?	X	L	N	N	N	
գաւազան	gawazan	1?	X	M	N	N	N	
գովեմ	govem	1?	X	H	N	N	N	
դաշտ	dašt	1?	2 3	H	N	N	N	
երագազ	eragaz	1?	X	L	E	N	N	
Երուանդ	Erowand	1?	1 2 3	L	E	N	N	
զառիկ	zarīk	1?	X	L	N	N	N	
խոնաւ	xonaw	1?	1 2 3	H	N	N	N	
Ծղուկ	Cłowk	1?	X	L	N	N	N	Strongly contested, perhaps from Georgian or Scythian
հ-	h-	1?	X	M	N	N	N	Intensifying prefix, might be from PIE *h_1su-, either via Proto-Iranian *hu- (“good”) or “contaminated with it” (translated from Ačaŕean, 1926: 1483).
հռչակ	hrčʿak	1?	X	H	N	N	N	
նկարէն	nkarên	1?	X	L	N	N	N	
նրան	nran	1?	X	H	N	N	N	

վան	van	1?	X	H	N	N	N	
վարդապետ	vaapet	1?	X	H	N	N	N	From Western Iranian
վստահ	vstah	1?	1 2 3	H	N	N	N	
վրան	vran	1?	X	M	N	N	N	
առասան	arāsan	2?	X	L	N	N	N	
Ասպահան	Aspahan	2?	2 3	L	N	N	N	
Արաստ	Arast	2?	X	L	N	N	N	
բազմակ	bazmak	2?	X	H	N	N	N	
բազուկ	bazowk	2?	X	H	N	N	N	
գիժ	giž	2?	X	L	N	N	N	From Medo-Parthian?
երաժիշտ	eražišt	2?	2 3	H	E	N	N	
հիւանդ	hiwand	2?	1 2 3	H	N	N	N	
մահ	mah	2?	X	H	N	N	N	
նպաստ	npast	2?	X	M	N	N	N	
վիպասան	vipasan	2?	X	L	N	N	N	
տաշտ	tašt	2?	2 3	H	N	N	N	
օտար	ôtar	2?	X	H	L	N	N	
թարխան	tʿarxan	3?	X	L	N	N	N	
շաքար	šakʿar	3?	1 2 3	M	N	N	N	