

PIE *e in Albanian

1. There seem to be at least six different reflexes of Proto-Indo-European *e in Albanian, viz. *i*, *e*, *je*, *ie*, *a* and *ja*. Most modern treatments of the historical phonology of Albanian assume the following basic distribution¹:

1. PIE *e is reflected as *e* mainly in front of NC and after Cr- and Cl-.
2. PIE *e was diphthongized to *je* in the remaining environments.
3. The reflex *je* has further developed to *ja* in closed syllables; in front of word-final resonants, *je* yields *ie*.
4. The reflex *ja* yields *a* after some of the palatal consonants.
5. The reflex *i* is the result of raising of *e caused by a high front vowel in the next syllable (*i*-mutation).

Rule number one is safely established on the basis of etymologies such as *pesë* 'five' < PIE *penk^we, *dhemb* 'to hurt' < PIE *ǵemb^h-, *dredh* 'to turn' < *d^hreǵ^h-. Some studies give more detailed rules as to the environments in which *e is retained, but we must be cautious here: there are indications that *je* was analogically replaced by *e* in part of our evidence. For instance, the verb *mbledh* 'to collect' does not show *je* in any of its conjugational forms today, but it appears as *ënbëljedh-* in the participle *enbelieḡune / ënbëljedhunë/* 'collected' in the 16th-century text of Buzuku². Another example has been adduced by Matzinger (1998: 234): the present of *heq* 'to pull' is attested as *hieq- / hieq-* in Buzuku, which confirms the evidence of dialects which show *hjek* instead of *heq*. Similar replacements may have affected other words, for which we do not have sufficient evidence in the older Albanian texts.

Rule number five has resulted in an alternation between *e* and *i* which mainly occurs in verb forms: in the 2pl. present indicative (only attested as subjunctive and future after the particle *të*, e.g. Buz. *teh enbeliḡnih/ të ënbëlidhni/* to *mbledh*, *teh vdisni / të vdisni/* to *vdes* 'to die'), in the 2sg. and pl. imperative (Bogdani sg. *mbëliḡ / mbëlidh/*, Buz. pl. *enbeliḡni / ënbëlidhni/* to *mbledh*), in the imperfect (3sg. Buz. *enbeliḡ / ënbëlidh/* to *mbledh*, *endiḡ / ëndiq/* to *ndjek* 'to follow'), in the present middle (Buz.

¹ Compare the surveys and discussions in Beekes 1995: 265, S. Demiraj 1996: 105ff., B. Demiraj 1997: 41f., Topalli 1998: 75-114, Orel 2000: 3ff.

² Words from Old Albanian authors (16th and 17th century) are given in their original form, followed by a (semi-)phonological interpretation, which follows the present-day orthography of Albanian.

23sg. *diḡe, diḡete / digje, digjetë/* to *djeg* 'to burn') and in the 3sg. *ti-aor.* (Buz. *engriti / ëngriti/* to *ngre* 'to raise'). Whereas Hamp (1971:224) has argued that *e changed to *i* in front of an earlier consonant cluster, Klingenschmitt (2000: 8) and Orel (2000: 145) explain *i* < *e from *i*-mutation. The latter solution seems the more likely one, even if it is impossible to prove the presence of **i* or **i* in the following syllable for all the verb forms in question. The main reasons for preferring *i*-mutation are: the 2pl.pres. originally ended in *-i* rather than *-ni* (cf. Klingenschmitt 1994: 226); the original ending has disappeared altogether in the 3sg. aorist; there is palatalization of stem-final velars in the imperfect and the present middle. Among the inherited nouns, *i* for *e is found in the words *mish* 'meat' < PIE *mēmso-, *nip* 'grandson; nephew' < *nepōt-, *viç* 'calf' < *uete/ōs-(i)o- and Tosk *vit* (Gheg *vjet*) 'year' < *uetos-. In my opinion, the change of *e to *i* in *mish* and *viç* can be due to the influence of the following palatal sibilant (for *mish* see de Vaan 1997: 62). The *i* in *vit* can be ascribed to umlaut in the oblique forms of *vjet*, e.g. gen.sg. **veti* (thus B. Demiraj 1997: 420). The word *nip* can also be explained from umlaut, viz. if we adopt Klingenschmitt's suggestion (2000: 11) that the nom.sg. PAIb. **nepō* < PIE **nepōts* led to the incorporation of this noun in the *n*-stem inflection, as shown by its plural *nipër* < **nepenes*. Since the plural suffix *-ër* regularly causes *i*-mutation of **a* to *e* (*shtrat* 'bed', pl. *shtrëtër*, *vëlla* 'brother', pl. *vëllezër*), it seems possible that the original paradigm of 'grandson' was sg. **nep*, pl. **nipër*, with *i*-mutation of *e in the plural. In order to explain *i*-mutation in front of the suffix *-ër*, Orel 2000: 230 derives it from PIE *-ino-, but since the core of the nouns in pl. *-ër* is formed by kinship terms, it is preferable to assume the suffix *-en-. Possibly, its vowel was raised to **i* in unstressed position in front of **n*. As argued by Klingenschmitt 2000: 8, the sg. *nip* may have acquired its *i* analogically from the plural, just like the sg. *dreq* 'dragon' < Lat. *dracō*, *-nis* must have adopted *e* from the plural, where it arose phonetically.

Rules number two, three and four as given above contain the standard explanations for *ie*, *je* and *ja*. A superficial glance at the evidence reveals their problematic nature:

- The reflex *ja* is not restricted to closed syllables.
- The reflex *je* is found in a considerable amount of closed syllables.
- The reflex *ja* mainly appears in inherited words and hardly in Latin loanwords, whereas *je* is found with the same frequency in inherited words and in loanwords.
- A phonetic change of *je* to *ie* would be surprising; usually, the direction is the reverse.
- The reflex *a* after palatals mainly appears in Latin loanwords.

These objections lead me to reconsider the evidence for diphthongization of PIE *e and of e in loanwords from Ancient Greek and Latin.

2. I start with the discussion of Albanian /ie/. Most scholars of Albanian agree on the view that *ie* has developed out of *je*, by a stress shift from the second to the first part of the diphthong. The reason why this view was proposed is the restriction of *ie* to a limited number of phonetic surroundings, and its alternation with *je* in verbal stems in *-r* and *-ll*, see e.g. Pekmezi 1908: 54, S. Demiraj 1996: 111. It seems to me, however, that the change to *ie* and the change to *je* represent two independent phenomena in the history of the Albanian language.

The distribution of *ie* is as follows: it represents *e in a stressed word-final syllable in front of one of the consonants *-ll* and **-r*. These conditions coincide with the conditions for the rise of the diphthongs *ua* from *ō and *ye* from *ō̄. These two long vowels partly continue long vowels from PIE or Latin, but also short vowels which were lengthened in stressed position in front of *-ll*, **-n* and *-r*. *duar* ‘hands’ < *dōr < *dēr < PIE *ǵ^hesr-es; *kapua* ‘rooster’ < Lat. *capōnem*; *muaj* ‘month’ < *muo+j < *mōn < *mēn < PIE *meh₁n-; *shuall* ‘sole of the foot’ < *shōll < Lat. *solum*; *arsye*³ ‘reason’ < *arsōn < Lat. (*ad*)*ratiōnem*⁴; *dyer* ‘doors’ < *dōr < *dōri < PIE *d^huōr-es.

The diphthongs *ua* (the older form of which was *uo*, as found in Buzuku) and *ye* have the stress on their first element, as the result of a diphthongization by which the first part of the original vowel was raised,

³ The Old Gheg auhtors consistently have a threesyllabic form: Buzuku *aresue* /ar(r)ēsye/, Bardhi *areshe*, *arreshe*, Bogdani *arrēstie* /arrēsye/. Buzuku’s plural form *aresegne* /ar(r)ēsenjē/ confirms the origin of *arsye* from a form in a nasal. Threesyllabic *arrēsye* might reflect Vulgar Latin **ad-ratiōnem*, although no such form is continued in other Romance languages. Jokl (1916: 138) has connected the verb *arsen* ‘to instruct, educate; to give a lesson in how to behave, scold, reprove’, which is attested in dialects as 1sg. *arrēsenj* ‘I reprove, scold, chase away’. Jokl argues that the verb can be cognate with Old French *araisnier* ‘to address’ < **arratiōnāre*, and that *arrēsye* is probably derived from the verb in Albanian. However, there is no hint of the meanings ‘teaching’ or ‘reproval’ in *arrēsye*, and the verb *arrēsenj* cannot be the phonetic reflex of **arratiōnāre*. Therefore, it is probably the noun which was primary in Albanian.

⁴ The difference between the reflexes of *capōnem* and *ratiōnem* is caused by the presence of *i* in front of *ō* in the second word: *ō* was fronted to **ō̄*. The same condition underlies the difference between the vocalic verbs ending in 1sg. *-oj* and those in 1sg. *-ej*, e.g. *bekoj* ‘I bless’ (Latin *benedicō*), *ndëgjoj* ‘I hear’ (*intellegō*) versus *pëlqej* ‘I please’ (*placeō*), *këmbej* ‘I change’ (*cambiō*); cf. Klingenschmitt 1981: 104f.

as e.g. in Italian (*nuovo*, *Pietro*): **ō* > *uo* > *ua*⁵, and **ō̄* > **yō* > *ye*. Since **e* yields *ie* in the same environment, and since a diphthongization of **ē* to *ie* would be the exact mirror image on the front vowel side of the change **ō* > *uo*, we may conclude that the changes **ō* > *uo*, **ō̄* > *ye* and **ē* > *ie* are three members of the same process of diphthongization of long mid vowels to falling diphthongs. This yields the following relative chronology: **-ell* > **-ēll* > *-iell* and **-er* > **-ēr* > **-ier* > *-ie*.

In front of *-ll*, we find three nouns: *diell* ‘sun’ (no certain etymology), *miell* ‘flour’ < PIE **melh₁u-*, *qiell* ‘heaven’ < Lat. *caelum*. The modern standard language contains several verbs in which both variants are allowed: *sjell* beside *siell* ‘to bring’ < PIE **k^wel-*, *mbiell* ‘to sow’ beside *pjell* ‘to beget’, dial. *piell* < **pel-n-*, *shtjell* ‘to throw’ beside *shtiell* dial. < **stel-n-*. As has been recognized by several scholars (e.g. Pekmezi 1908: 54, Cimochowski 1951: 163, Çabej 1975: 214), many peripheric dialects show the originally complementary distribution between *ie* in front of word-final *ll* and *je* in front of word-internal *ll*. Compare, for instance, the Gheg presents *piell* ‘brings forth’, *viell* ‘vomits’ but participles *pjellë*, *vjellë*. Therefore, a modern present form such as *vjel* ‘to reap’ must be due to the analogical transfer of *je* from the participle *vjelë*⁶.

In part of the verbs in which **e* stood in front of **-r*, the vibrant is retained, whereas in others it has been lost. It is clear that the loss of **-r* after *ie* is original, whereas verbs ending in *-ier* in the present have restored *-r* on the basis of the participle and maybe other forms of the paradigm; cf. Topalli 1998: 99f. Thus, we find *bie* ‘to fall; to bring’ < PIE **b^her-* (part. *rënë* ‘fallen’, *prurë* ‘brought’), *shpie* ‘to take’ < **per-* (part. *shpënë*), *shtie* ‘to put into’ < **sterh₃-* (part. *shtëně*), on the one hand, but *bier* ‘to lose’ < **b^her-* (part. *bjerrë*), *tier* ‘to spin’ < **ter(H)-* (part. *tjerrë*), *vier* ‘to hang’ < **h₂uer-* (part. *vjerrë*), on the other. In fact, the standard language now prefers *je* in the present of these verbs: *bjerr*, *tjerr*, *vjerr*. As argued by To-

⁵ There is no need to posit an intermediate stage **ou*, as Topalli 1998: 118ff. does; this suggestion goes back to Jokl 1931: 276ff. The alleged metathesis of **ou* to *uo* is very unlikely.

⁶ The co-occurrence of *-iell* and *-jell* has been one of the main arguments for supposing that *ie* developed out of *je*. However, it is cross-linguistically very rare for *je* to change into *ie*; the reverse is commonplace in languages, and has occurred recently in some Albanian dialects, especially northern Tosk (according to Topalli 1998: 95; more details in Desnickaja 1968: 269f.), where e.g. *miell-i* ‘the honey’ has become *mjelli*.

palli, the original form of the verb *ziej* 'to boil' was *zie*, as still found in dialects; it reflects the PIE root **g^{wh}er-* 'to become hot'⁷.

In *fier* 'fern', *-r* has been restored on the basis of the definite nom.sg. **feri*. This noun was explained as a borrowing from Ancient Greek *ptéris* (πτέρις) or its Modern Greek descendant *ftéri* (φτέρη) by Thumb (1909: 13).

In front of *-m* and *-n*, we find *ie* in a few 1pl. and 3pl. pres.ind. forms of verbs in **-er*: *biem*, *bien* 'we bring, they bring', similarly *shpiem*, *shpien* and *shtiem*, *shtien*. Since the nasal consonant has regularly disappeared after *ua* and *ye* from **-ōn* and **-ōn̄*, it might be expected that the nasals would have disappeared after *ie* too; therefore, these verb forms are suspect. Furthermore, the endings *-m* and *-n* go back to *-mē* and *-(ē)nē*, which can still be found in many dialects, and in older texts such as Buzuku; but in original **be(r)mē*, **be(r)nē*, the condition for diphthongization to *bie-* was absent. Therefore, it will be safer to assume that verb forms such as *biem* and *bien* have analogically introduced *ie* from the singular.

The regular reflex of word-final **-en* is not a diphthong, but a nasalized monophthong *-ē* in Gheg, and *-e* or *-ë* in Tosk; cf. Topalli 1996: 56ff. for a discussion. A Latin loanword showing this development is *fre*, *-ri* 'bridle' < Lat. *frēnum*. Possibly inherited words showing this sequence are *pe*, *-ri* 'thread' (< **pen-* to PIE **(s)penh₂-* 'to stretch' according to Ölberg 1972: 29), and *re*, Gheg *rēja* 'cloud' (compare the Greek gloss *ρίβόν* acc.sg. 'fog' attributed to the Illyrians). Thus, the reflex of **-en* is parallel to that of other stressed vowels before word-final nasal: **-an* > Gheg *-ā* / Tosk *-ë*, **-in* > *-ī/-i*, **-un* > *-ū/-u*, **-ün* > *-ÿ/-y*. Thus, whereas **e* patterns with **ō* and **ō̄* in front of **-r* and *-ll*, it does not in front of **-n*.

In front of *l*, no diphthongization of **e* has taken place, as shown by Topalli 1998: 120ff. The forms which do show *ie* in front of word-final *l*, such as *e diel* 'Sunday', have analogically acquired *ie* from related forms in *-iell*, in this case from *diell* 'sun'. In front of *rr*, there was also no phonetic (lengthening and) diphthongization into *ie*. In front of **nj*, Topalli also argues for the original absence of (lengthening and) diphthongization for the vowel **o*; we may assume the same for **e*. This accords well with tendencies which can still be seen at work in Albanian dialects: lengthening of vowels frequently occurs in front of word-final *r* and *ll*, but not in front of *rr* and *l*.

⁷ A similar explanation of *ziej* was already given by Klingenschmitt 1981: 109, 128f.

3. Albanian *je* is found as a reflex of PIE **e* in front of stops, fricatives, *r*, *rr*, *ll*, and in front of the clusters *rdh* and *rg*. Examples according to B. Demiraj 1997 are: *rjep* 'to peel' < PIE **h₁rep-*; *dhjetë* 'ten' < **dekm̄t*; *vjedh* 'to steal' < PIE **ued^h-*; *dhjes* 'to shit' < **g^hed-je/o-*; *mjekër* 'chin, beard' < **smekur-*; *njeri* 'human being', pl. *njerëz* < **h₂ner^s-*; *djeg* 'to burn' < **d^heg^{wh}-*; *vjehërr* 'father-in-law' < **suekuros*; *vjerrë* 'hanged' < **h₂uer-n-*; *pjerdh* 'to fart' < **perd-*; *djeg* 'to be lying ill' < **d^herHg^(h)-*.

The diphthongization to *je* also affects Latin loanwords with Classical Latin *e* and *ae*: *kështjellë* 'castle' < *castellum*, *mjek* 'doctor' < *medicus*, *pjeshkë* 'peach' < **pesca* < *persica*, *vjetër* 'old' < *vetere*, etc. As in the inherited vocabulary, there is no diphthongization of Latin *e* in front of nasal plus obstruent (*kuvend* < *conventus*, *mend* < *mentem*), nor, apparently, after *Cr-*: *grek* 'Greek' < *graecus*, *pre* 'prey, booty' < *praeda*. Loanwords with Latin *ē* do not undergo diphthongization: *vrer* 'gall, bile' < *venēnum*, *qetë* 'quiet' < *quiētus*.

As was seen by Romanists dealing with Albanian, the distribution of *e* and *je* in Latin loanwords exactly matches the Vulgar Latin distinction between a closed mid front vowel /e/ and an open vowel /ɛ/, the latter of which underwent diphthongization to *ie*, *je* to a greater or lesser degree in most of the Romance languages (cf. Meyer 1904-06: 1043). The similar conditions for the development of **e* in inherited and in borrowed words suggest that, like Vulgar Latin, Proto-Albanian distinguished between two phonemes /e/ and /ɛ/ (thus Ölberg 1972: 145ff., and, from a Romanist's point of view, Bonnet 2000: 73f.). The difference between these two vowels may already have arisen before the Latin period, by means of the monophthongization of the PIE diphthongs **oi*, **h₂eɪ* and **eɪ* to **e*. This new monophthong probably was a higher vowel than the reflex of PIE **e*. The tendency to diphthongize /ɛ/ which was present in the Vulgar Latin adstrate of Proto-Albanian may have triggered this change within the inherited vocabulary. Put differently, the diphthongization in inherited words and in Latin loanwords represents one and the same process. We can date this diphthongization relatively well, viz. before the influx of

⁸ Since the Old Gheg authors consistently spell /*njeri*/ with initial <ni(j)e->, as opposed to initial <gn-> in *një* 'one' and *njef* 'knows', some scholars have assumed that <ni(j)eri> contains the diphthong /ie/, and that this was the original reflex of PALb. **nēr* < **ner* < PIE **h₂ner-*; thus e.g. Ölberg 1972: 65, S. Demiraj 1996: 166, Ashta 2000: 390f. However, as Stefan Schumacher points out to me, there is no way in which initial /nje/ and /nie/ could have been graphically disambiguated by the Old Gheg authors. Since we always find the word *njegullë* 'fog' spelled with initial <nie->, it is more likely that <gn-> indicates a palatalized phoneme /ń/, whereas <ni-> is used for biphonemic /nj/.

Slavic loanwords into the language, that is, before the 7th century (Svane 1992: 290).

Since we have already concluded that the change of *e* to *ie* was caused by lengthening to **ē* in front of a resonant, we can now be certain that the changes *e* > *je* and *e* > *ie* must have been two separate, independent developments. In theory, they might have occurred simultaneously; but more probably, **e* > *ie* post-dates **e* > *je*, since lengthening in front of resonants can still be observed in many dialects today. As it is unlikely that **-eR* > *-ieR* would have passed through a stage **-jeR*, the change *e* > *je* probably did not apply in front of word-final *-r*, *-ll* and *-n* at all. The reason may have been that **e* yielded a closed vowel /*e*/ in this position.

4. The distribution of *ja* is rather different from that of *je*: nearly all certain instances of *ja* concern words which are inherited from PIE. The PIE evidence, according to Demiraj 1997, consists of the following forms: *djathë* ‘cheese’ < PIE **d^hed^hh₁-*, *djathë* ‘right’ < **deks(i)-*, *jashtë* ‘outside’ < **h₁ek^s-t-*, *mjalhtë* ‘honey’ < **meli-t-*, *pjalm* ‘pollen’ < **pel(H)-mo-*, *zjarm* ‘fire’ < **g^hermo-*. The evidence of *jam* ‘I am’ and *janë* ‘they are’ < **h₁esmi* and **h₁senti* is regarded as inconclusive by many scholars (e.g. S. Demiraj 1986: 743), who suspect that *jam* and *janë* have acquired a *u* under the influence of *kam*, *kanë* ‘I have, they have’; the original forms would have been **jem* and **jenë*. Three words in *gja-* are also adduced as evidence for **e* > *ja*, viz. *gjalpë* ‘butter’ < PIE **selpos*, *gjarpër* ‘snake’ < **serpono-* and *gjashtë* ‘six’ < **s(u)ek^s*. However, the sequence *gja-* is ambiguous since PIE **s-* yields *gj-* in front of all accented vowels (cf. Kortlandt 1987: 219), as in *gjallë* ‘alive’ < **soluo-*, *gjak* ‘blood’ < **sok^o-* and *gjumë* ‘sleep’ < **súpnos*. Hence, the *a* in *gjalpë*, *gjarpër* and *gjashtë* could be due to a change **e* > *a* which was conditioned by the preceding *gj-* (or its predecessor). We must therefore leave these forms out of consideration when trying to explain the change of **e* > *ja*; afterwards, we will return to them.

Unlike *je*, *ja* is seldom found in Latin loanwords. Latin loanwords with *a* for **e* after a palatal consonant are sometimes adduced as evidence for original **ja*, under the assumption that **j* was absorbed by the preceding palatal: *qartoj* ‘to quarrel’ < *certāre*, *qarr* ‘oak’ < *cerrus*, *shalë* ‘saddle’ < *sella*, *sharrë* ‘saw’ < *serra*, *shartoj* ‘to graft’ < **sertāre*. However, the change of *e* to *a* in the position between *q* or *sh* to the left and *l* or *r* to the right seems to be regular, so that we may just as well assume that *e* was directly lowered to *a* in this position. Thus, these forms cannot be used as evidence for a stage **ja*.

The only remaining candidates for a Latin origin are Alb. *fjalë* ‘word’

and *javë* ‘week’. The explanation of *fjalë* is disputed. According to some scholars, it reflects Latin *fabella*, but the loss of Latin intervocalic *b* (which in itself is a regular development, cf. B. Demiraj 2001: 61f.) would have us expect a long vowel from the contraction of *a* and *e*. The etymology might be saved by assuming (near-)haplology of **fabēlla* to **fēlla*, whence **fēlē*. *Fjalë* would then be the only Latin borrowing for which the change **e* > *ja* is required. However, it is not certain that a word **fabella* existed in eastern Romance. According to Meyer-Lübke (1935: 271), the word is attested in Italian (*favella* ‘speech’) and Portuguese (*favelas* ‘horse teeth’), and must be derived from the verb **fabellāre*. *Fabella* is thus a relatively recent creation of Vulgar Latin, and it may not have spread to eastern Romance anymore; or it may have arisen independently in the dialects where it is attested. Hence, a PIE etymology for *fjalë* must still be reckoned with. Scholars who take the latter view (e.g. Kortlandt 1987: 220, Beekes 1995: 263, Orel 2000: 95) connect *fjalë* with the PIE root **spelH-* ‘to speak in public’ (LIV²: 576), as in Gothic *spill* ‘story’. In any case, many scholars assume that *f-* can reflect PIE **sp-*, mainly because of the comparison between Alb. *farë* ‘seed, sperm’ and Greek *sporá* ‘seed’ (cf. also Demiraj 1997: 56).

Albanian *javë* is usually regarded as a borrowing of Greek *hebdomás* ‘week’ (acc.sg. *hebdómada*) via Latin, but this is also problematic. Meyer-Lübke (1935: 343) lists two sets of forms: 1. Vegliote *yedma*, Italian *edima*, Old Bolognese *edema* etc. < *hebdoma*, 2. Old Italian *domada*, Old French *domée*, Old Galician *domea* < *hebdomada*. Thus, neither of these two Romance reflexes has preserved a trace of Greek **b*; it would be surprising if Albanian had. Rather, Albanian *-v-* could directly reflect the Greek (post-Classical) pronunciation [evðomas]. In general, intervocalic *-v-* is rare in Albanian, and some of the best examples are found in Greek loanwords, e.g. *qeveris* ‘to govern’ < **kivérnisa* < aorist κιβέρνισα. Thus, *javë* may be an Ancient Greek loanword from before the Roman era. In support of this, I note that *hebdomás* is already attested with the meaning ‘seven days, week’ in Hippocrates (5th century BC), and in the Septuagint. The main problem with the Greek etymology is the unexplained loss of the syllable containing *-ma-*. Kortlandt (personal communication) suggests to me that **evðomada* may have been abbreviated to **evðo*.

The restriction of *ja* to the pre-Roman vocabulary is sufficient to dismiss the generally accepted view that *ja* has developed out of **je*. If the change of *e* to *je* represents a single development in the history of Albanian, which took place after the beginning of the Roman era and before the 7th century, we must conclude that it post-dates the change of **e* to *ja*, which does not affect Latin borrowings. It now remains to be determined, under which

phonetic conditions *ja* arose.

Most of the words containing *ja* from **e* end in *-ë*: the inherited vocabulary *djathë*, *djath(t)ë*, *jashtë*, *mjaltë*; the disputed words *fjalë* and *javë*; and most of the remaining words (of uncertain etymology) which have *ja*: *djalë* 'boy', *ngjalë* 'eel', *dhjamë* 'animal fat, tallow'. The two main exceptions are *pjalm* and *zjarm*. Hence, it is conceivable that the change of **e* to *ja* was caused by the vowel in the following syllable. The noun inflexion of Albanian is characterized by the opposition between masculine forms with a nominative singular in zero, and feminines with a nom.sg. ending in *-ë*. Since Latin loans in *-us* have usually become Albanian masculines, and Latin loans in *-a* have become Albanian feminines, this distinction must have been present in Proto-Albanian at the time when the Latin loanwords came in. Accordingly, most scholars assume that this distinction continues the difference between PIE *o*- and *eh₂*-stems, respectively. The actual endings of Proto-Albanian are reconstructed in different ways by different scholars: Orel 2000 assumes masc. **-a* and fem. **-ā*, whereas Klingenschmitt (1994: 223f.) reconstructs **-oh* and **-ā*, respectively. One might even assume masc. zero and fem. **-a*; the main point being that the feminine ended in a low vowel. The change of **e* to *ja* in the feminines can then be interpreted as a case of *a*-mutation under the influence of **-a* in the following syllable. The change which I propose is comparable to the so-called 'fracture' in Old Norse, whereby Proto-Germanic **e* changed into **ea* to yield *ja*: e.g. ON *gjalda* 'to pay' < PGM. **geldanan*, *jafn* 'even' < **ebnaz*. Closer to Albania, this development is reminiscent of the Rumanian change of *e* to *ea* if the next syllable contains *-ă*, *-a* or *-e*: *negru* 'black' (m.) but *neagră* 'black' (f.), *seară* 'evening' but *seri* 'evenings'. The only scholar who – to my knowledge – has previously considered a similar explanation for *ja* is Ölberg (1972: 72), who suggests that "geschlossene Silbe oder nachfolgender dunkler Vokal (*a* > *ë*) die Öffnung des *je* zu *ja* verusacht habe". Once again, I reject an intermediate stage **je*.

The phonetic origin from *a*-mutation may also explain the fact that, as noted e.g. by Topalli 1998: 133, *ja* is found as a reflex of **e* in nouns, adjectives, adverbs and numerals, but not in verbs. Topalli ascribes this to systemic pressure within the verbal system, which aimed at preserving the ablaut difference between the present (*e*, *je*) and the aorist (*o*); but even if this were true, it would suggest that, from the outset, there were few verb forms in which the change of **e* to *ja* was phonetically motivated. Probably, the condition for *a*-mutation was not given at all in the verbal system: the verbal endings which we can reconstruct for Proto-Albanian contain the vowels **e*, **ē*, **o* (> **a*), **ō* or **oi* (> **ē*), but not **-ā*; see Klingenschmitt 1994: 255ff.

Only part of the words in *-ja-* which have final *-ë* are indeed feminine nouns, viz. *djathë* 'right hand' (*diaḡa* /*djatha*/ in Buzuku), *fjalë*, *javë* and *ngjalë*. The distribution of the remaining forms is as follows: masculine are *djathë* 'cheese' and *mjaltë*; masculine or neuter is *dhjamë*; and *jashtë* is an adverb.

We will start with the masculine or neuter nouns in *-ë*. In fact, Albanian has several other masculine nouns ending in nom.sg. *-ë* instead of zero: *brumë* 'dough', *burrë* 'man', *gjumë* 'sleep', *lumë* 'river', etc. Orel (2000: 233) reconstructs an end-stressed nom.sg. form in *-ós* > *-ë*, suggesting that the zero ending in the majority of masculines reflects the unstressed ending **-os*. Yet there is no independent proof that the masculines in *-ë* had final stress and the others did not; moreover, *gjumë* 'sleep' must have had initial stress (**súpnos*) in order to yield the reflex *gj-* (cf. Kortlandt 1987). We must rather explain these masculines as original neuters, in which the ending *-ë* reflects the original ending of PIE neuters.

As a matter of fact, the noun *dhjamë* is still acknowledged as a neuter noun for the standard language by Buchholz-Fiedler 1987: 209f., who also point out that *djathë* 'cheese' and *mjaltë* are still sporadically attested as neuters in texts after 1944. The further back we go in time, the more masculines appear to have been neuters: a list of 33 nouns is given by Buchholz-Fiedler loc.cit., and many of these take final *-ë*. The noun *djalë* is found as a neuter with the definite article *-të* in Buzuku: *dialete* /*djalëtë*/ 'the son'.

It has already been claimed by Klingenschmitt (2000: 5) that the PIE neuter nom.acc.sg. ending **-om* regularly yielded *-ë* in Albanian, on the strength of examples such as Old Gheg *emënë*, definite *emënitë* 'name'⁹. In support of this, note that the ending **-om* is probably also reflected as *-ë* in several pronominal forms: acc.sg.m. of the definite article *-në* (Old Gheg, dialects) < **-ntom* < **-m* + *tom*, acc.sg.m. *atë* 'him/her', *këtë* 'this', acc.sg. *kë* 'whom?'. Since **-om* must have yielded **-am* in the first place, it is possible that this ending caused the same *a*-mutation of **e* as final **-a* of the feminines did.

⁹ Of course, the assertion that the Albanian neuter really continues the PIE neuter gender has been challenged, but the arguments in favour of continuity seem more compelling than the arguments against it; see the discussion in S. Demiraj 2002: 103-112.

Finally, original **-om* can explain the ending of *jashtë* ‘outside’, which probably contains PIE **h₁eǵs* ‘out’. Final *-të* can go back to the PIE suffix **-tos* conveying an ablative meaning, as in Skt. *itās* ‘from here’, Greek *ἐκτός* ‘(from) outside’, Latin *intus* ‘(from) inside’. As Stefan Schumacher suggests to me, it is conceivable that the ending **-tos* was remade to **-tom* on the analogy with other adverbs which were derived from an acc.sg.

Above we have excepted from the evidence three forms in *gja-*, which are ambiguous because PIE **s-* yields *gj-* in front of stressed vowels anyway: *gjalpë* ‘butter’, *gjarpër* ‘snake’ and *gjashtë* ‘six’. As it turns out, all three forms can also be interpreted in accordance with the rules for *a*-mutation of **e*. Firstly, the PIE neuter *s*-stem **sélpos* which must be reconstructed for *gjalpë* may have been changed to a thematic stem in nom.acc.sg. **-om* because of its neuter gender. Secondly, the numeral *gjashtë* from PIE **s(ǵ)eḱs*¹⁰ has added the suffix *-të* which we find in the Albanian numerals ‘six’ to ‘ten’. As Hamp has argued (1992: 913), this may well go back to PIE **-tā-*, which would again provide the necessary condition for *a*-mutation of **e*. Thirdly, *gjarpër* has been reconstructed either as **serpeno-* or as **serpono-*. Klingenschmitt’s argument (1981: 129) that the vowel *a* in the Old Gheg plural forms proves PIE **-ono-*, is conclusive. Compare the oldest attestations: nom.sg.indef. Bardhi, Bogd. *giarpene* / *gjarpenë*/, acc.sg.def. Buz. *ḱarpenineh* / *gjarpeninë*/, acc.pl.def. Buz. *ḱerpagneteh* / *gjarpanjë*/, Bogd. *giarpagnetë* / *gjarpanjë*/ . Of course, it remains uncertain whether **a* in the second syllable of a threesyllabic word would regularly cause *a*-mutation at all.

We can now discuss *zjarm* and *pjalm*, the two nouns in *-ja-* without final *-ë*. Their meanings ‘fire’ and ‘pollen’ would fit in very nicely with the other neuter nouns, which encompass many mass nouns. However, the absence of *-ë* after *-lm-* and *-rm-* is unlikely to be due to phonetic loss, since the reverse dictionary of Albanian shows quite a number of words in *-lmë* and *-rmë*. In the modern standard language, the usual word for ‘fire’ is the masculine *zjarr*. *Zjarm* has disappeared, but a form *zjarm-* occurs as the basis for a few derivatives such as *zjarmí* ‘fever; passion’. In dialects, *zjarm* survives, and the plural form is sometimes *zjerm*, unfortunately, I

¹⁰ Pedersen (1900: 286-289) has argued that the phonetic outcome of PIE **sǵ-* in front of a stressed vowel was *d-*. If this is correct, it would imply that - like other Indo-European languages - Albanian had changed **sǵeḱs* to **seḱs*, probably on the analogy of **septm* ‘seven’.

do not have enough data to provide a reliable survey of the dialect geography. In the Old Gheg texts we find *zjarm* and *zjarrë* without any difference in meaning. The attested forms are:

- Buzuku: nom.sg.indef. *ɛiarm* / *zjarm*/, acc.indef. *ɛiarm*,
gen.sg.indef. *ɛiarmi*, nom.def. *ɛiarmi*, gen.sg.def. *ɛiarmit*
acc.sg.(in)def. *ɛjare* / *zjarrë*/
- Budi: gen.sg.def. *zjermit*¹¹
acc.sg.indef. *zjarrë*
- Bardhi: nom.sg.indef. *ɛiarm* / *zjarm*/, acc.sg.indef. *ɛiarne* / *zjarmë*/,
nom.sg.def. *ɛiarmi*, gen.abl.sg.def. *ɛiarmit*
acc.sg.def. *ɛiarre* / *zjarrë*/
- Bogdani: nom.sg.indef. *ɛjarm*, *ɛiarm* / *zjarm*/, acc.sg. indef. *ɛjarm*,
abl.sg.indef. *ɛjarmi*, nom.sg.def. *ɛjarmi*, acc.sg.def. *ɛjarminë*
/ *zjarminë*/, abl.sg.def. *ɛjarmit*
acc.sg.def. *ɛiarre*, *ɛiarre* / *zjarrë*/

In the text of Buzuku, we only find one occurrence of *zjarrë*, viz. in *e 8 stið ɛiare ʎuteteuet tune / e u shtiu zjarë qytetevet tyne/* ‘and it set fire to their cities’ (Mt. 22:7), which might represent a definite acc.sg. (but an indefinite is not excluded). All other 37 occurrences of the word ‘fire’ take the stem *zjarm*, of which all sg. case forms are attested except for the definite acc.sg. In Budi’s ‘Rituale’, *zjarrë* would normally have to be interpreted as the acc.sg. of the indefinite: Rituale 43 *ta bajë me zjarë e ta timonjë fontënë* ‘that he make fire and make the fountain smoke.’ Bardhi’s ‘Dictionarium Latino-Epiroticum’ also has *zjarrë* only once, as a definite acc.sg.: *cus dro ɛiarre, n timit ichen / kush dro zjarrë, n timit ikën/* ‘Qui timet ignem fumum fugiat’. In Bogdani’s *Cuneus Prophetarum*, there are only 4 occurrences of the form *zjarrë*, against 82 of the stem *zjarm*. This is the first text in which we find the definite acc.sg. of the stem *zjarm*, viz. twice *zjarminë*. The form *zjarrë* seems to function both as an indefinite and a definite acc.sg.: *Hini pra S. Pietri ndë ketë vend e persë isc ftofetë patnë nde ɛunë ɛjarrë, tue ù nzeem ajo gind e sctepijssë* ‘Saint Peter thus went in there and, because it was cold, those servants had lit a fire, warming themselves’; *Mbij Ajrit crijoj ɛotüne ɛiarre, issili ansctë fort i nzetë, e i ɛḱḱaatë* ‘After the air our Lord created the fire, which is very hot and dry.’

¹¹ In principle, *ja* from **e* cannot be subject to *i*-mutation, because the conditions of both mutations are mutually exclusive: the endings **-a*, **-am* cannot cause *i*-mutation, and the ending(s) which bring about *i*-mutation cannot cause *a*-mutation. Since the other Old Gheg authors all show *zjarmit*, Budi’s *zjermit* is probably due to analogical extension of *i*-mutation from words containing PAlb. **a*, such as *ati* ‘pater’, *tet* ‘patris’ in Bardhi.

The fact that *zjarminë* does not occur before Bogdani might lend support to the suggestion by B. Demiraj (1997: 429) that *zjarrë* arose phonetically from the acc.sg. **zjarm-në*. Yet although this development seems possible from the phonetic point of view¹², it would be very surprising if the definite acc.sg. (which, as we have seen, forms only a minority of the attestations of ‘fire’ in Buzuku and Bogdani) had provided the basis for the present-day preponderance of *zjarr* as the word for ‘fire’.

I think that we are dealing with two originally different words, viz. *zjarrë* and *zjarm-* (or **zjerm-*, see below). They may have denoted a difference in meaning, probably ‘fire’ as a generic denomination versus a concrete ‘fire’. This opposition is found in Latin, viz. Classical Latin *ignis* ‘fire’ versus *focus* ‘a fire-place’; the latter has become the only word for ‘fire’ in VLat. For the formal side of the opposition *zjarrë*: *zjarm*, compare the situation with the word *djalë* ‘boy’ and its plural *djem* < *djelm*, which is described in detail by B. Demiraj 1997: 134f. He concludes that there have been two ‘konkurrierende Stammbildungen’, viz. *djal-* (with the pl. *djel-*) and *djelm-* (with its sg. *djal-*), going back to **dela(-)* and **delmo-*, respectively. They subsequently merged into one paradigm, one form being used for the singular, the other one for the plural. In analogy to *djalë*: *djem*, we might assume that there were two words for ‘fire’ which were built on the PIE root **g^{wh}er-*. One was *zjarrë* < neuter **g^{wh}ernom*¹³, in agreement with the rules for the rise of *ja* as established above. The other was *zjarm-* or, maybe, *zjerm-*: the sg. *zjarm* might be an analogical singularization of the pl. *zjerm*. Since PIE **g^{wh}ermos* ‘warm’ was an adjective, one might prefer a PIE neuter noun **g^{wh}ermom* ‘heat’, whence Albanian ‘fire’. But it is also conceivable that *zjarm* does not go back to PIE at all, but is rather a secondary form built with **-mo-* on the basis of *zjarrë*, just like *djelm* is secondary to *djalë*.

The noun *pjalm* is not attested in the 16th and 17th-century texts, and I do not have sufficient dialect data to say whether *-ë* might have been apocoped or not. If the earlier form was **pjalmë*, the noun can regularly go back to a neuter **pel(H)om*, from the same root **pel-* ‘to procreate’

¹² There is no exact parallel of this assimilation, but the acc.sg. ending **-në* does assimilate to a preceding resonant in Buzuku, cf. Ashta 2000: 203: *-lnë* > *-llë*, *-llnë* > *-llë*, *-rnnë* > *-rë*.

¹³ B. Demiraj (1997: 429) rightly objects that we would expect zero grade of the root in view of Skt. *ghṛnā-* m. ‘glow (of the sun)’, OCS *grünŭ* ‘kettle’; the Albanian noun may have adopted the full grade from the adj. PIE **g^{wh}ermos*. If Alb. *-rr-* can also derive from single intervocalic **-r-*, as Demiraj argues, *zjarrë* could go back to a PIE *s*-stem **g^{wh}er-os*, cognate with Skt. *hāras-* ‘flame’ and Greek *θερός* ‘warmth, summer’.

which, according to B. Demiraj 1997: 323, underlies the verb *pjell* ‘to give birth to; bear fruit’.

The proposed explanation for *ja* would be even more convincing if there were no inherited words which contain PIE **e* > *e* or *je* in front of an ending *-ë*. In fact, I find only four possible counterexamples in the evidence collected in Demiraj 1997, viz. *dhjetë* ‘ten’, *pendë* ‘yoke of oxen’, *pesë* ‘five’ and *shtjerrë* ‘lamb’. Since nasals plus obstruent block the change of **e* to *je* at a later stage, they may also have blocked **e* > *ja*; this concerns *pendë* and *pesë*. For *shtjerrë*, Demiraj 1997: 377 considers a derivation from **ster-en-* to PIE **ster-* ‘sterile’¹⁴; the nasal suffix is needed to explain intervocalic *-rr-*. If this is correct, the absence of *ja* would be regular.

For *dhjetë* < PIE **dekm̥t*, we may reconstruct **dekm̥t* > **deθat*, or, with early loss of **-t*, **deθa*; compare *shta-të* ‘seven’ < **septm̥* + **-tā*. The following stages of the development are uncertain in many respects: it is uncertain whether **-tā* was added early enough to cause *a*-mutation (recall that *gjashtë* ‘six’ is ambiguous); it is uncertain whether **a* from the syllabic nasal caused *a*-mutation, and if it did, it might have been syncopated at an early stage from the preform **déθatā*; initial *dh-* represents lenited, intervocalic **d*, which was probably adopted from composite numerals such as *njëmbëdhjetë* ‘11’ or *tridhjetë* ‘30’ (cf. Hamp 1992: 916). I conclude that *dhjetë* is unsuited to serve as a counterexample to *a*-mutation.

5. The investigation has yielded the following three main results:

1. Albanian *ie* is the result of diphthongization of **ē*, which in its turn reflects lengthening of PIE and Latin **e* in front of word-final *-ll* and *-r*. This diphthongization is relatively recent, and probably contemporary with the changes of **ō* > *uo* and of **ō* > *ye* in front of the same resonants.
2. Albanian *je* is the result of diphthongization of the open mid front vowel **e*, which reflected both PIE short **e* and Latin *e/ae* in part of the Latin loanwords. This change must have been completed before the influx of Slavic loanwords.
3. Albanian *ja* is the result of *a*-mutation of PIE **e*, which took place under the influence of a vowel **-a* or **-a(m)* in the next syllable. This change can be dated before the influx of Latin loanwords.

The analysis of the evidence allows us to establish the following relative chronology of developments:

¹⁴ The assumed semantic development goes from ‘sterile’ via ‘which has not yeaned yet’ to ‘young animal’, ‘lamb’. Compare Vedic Sanskrit *starf-* ‘sterile; who has not given birth yet, young woman’.

1. PIE **-eh₂* > **-ā* > **-a*, PIE **-om* > **-am*
2. PIE **e* > *ja* in front of **-a* and **-a(m)* in the following syllable.
3. PALb. acquires two different phonemes /ɛ/ and /e/, PIE **oi*, **h₂ei* and **eu* being the source of /e/.
PIE **e* and **e* in borrowings join /e/ when stressed in front of *-ll*, *-n*, *-r*.
4. PALb. **ɛ* > *je*.
Diphthongization of stressed **ē* > *ie*, **ō* > *uo*, **ō̄* > *ye* in front of word-final *-ll*, *-r*.
Diphthongization of stressed **ō* > *uo*, **ō̄* > *ye* in front of word-final **-n*.
5. *i*-mutation of **a* to *e*, unrounding **ō̄* > *ē*.

Comparative Linguistics (VTW)
PO Box 9515
NL-2300 RA Leiden
M.A.C.de.Vaan@let.leidenuniv.nl

Michiel de Vaan

Bibliography

- Ashta, K. 2000: *Leksiku historik i gjuhës shqipe*, volume I, Shkodër.
Bonnet, G. 2000: *Les mots latins de l'albanais*, Paris – Montréal.
Beekes, R.S.P. 1995: *Comparative Indo-European Linguistics*, Amsterdam – Philadelphia.
Buchholz, O., Fiedler, W. 1987: *Albanische Grammatik*, Leipzig.
Çabej, E. 1975: 1. *Hyrje në historinë e gjuhës shqipe*, 2. *Fonetika historike e shqipes*, Tiranë.
Cimochowski, W. 1951: *Le dialecte de Dushmani*, Poznań.
Demiraj, B. 1997: *Albanische Etymologien*, Amsterdam – Atlanta.
– 2001: Das Meyersche Gesetz über den Schwund der intervokalischen Media im Albanischen, *MSS* 61, 57-92.
Demiraj, S. 1996: *Fonologjia historike e gjuhës shqipe*, Tiranë. [with a summary in English]
– 2002: *Gramatikë historike e gjuhës shqipe (botim i përmbledhur)*, Tiranë. [with a short summary in English]
Desnickaja, A.V. 1968: *Albanskij jazyk i ego dialekty*, Leningrad.
Hamp, E. 1971: 'Fils' et 'fille' en italique: nouvelle contribution, *BSL* 66, 213-227.
– 1992: Albanian, in: *Indo-European Numerals*, ed. J. Gvozdanović, Berlin – New York, p. 837-921.
Jokl, N. 1916: Beiträge zur albanesischen Grammatik, *Indogermanische Forschungen* 36, 98-164.

- 1931: Zur Geschichte des alb. Diphthongs -ua-, -ue-, *Indogermanische Forschungen* 49, 274-300.
Klingenschmitt, G. 1981: Albanisch und Urindogermanisch, *MSS* 40, 93-131.
– 1994: Das Albanische als Glied der indogermanischen Sprachfamilie, *In honorem Holger Pedersen, Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*, ed. J.E. Rasmussen unter Mitwirkung von Benedicte Nielsen, Wiesbaden, 221-233.
– 2000: 'Albanisch und seine Lehnbeziehungen in früherer Zeit', Handout at the XI. Fachtagung der indogermanischen Gesellschaft, Halle an der Saale, September 2000.
Kortlandt, F. 1987: PIE. **s* in Albanian, *Studies in Slavic and General Linguistics* 10, 219-226.
Matzinger, J. 1998: Review of: V. Orel, 'Albanian Etymological Dictionary', Leiden - Boston - Köln 1998, *Die Sprache* 40, 229-241.
Meyer, G. 1904-06: Die lateinischen Elemente im Albanesischen, neu bearbeitet von W. Meyer-Lübke, *Grundriss der romanischen Philologie. I. Abschnitt: Romanische Sprachwissenschaft*, Strassburg, 1038-1057.
Meyer-Lübke, W. 1935: *Romanisches etymologisches Wörterbuch*³, Heidelberg.
Ölberg, H. 1972: *Untersuchungen zum indogermanischen Wortschatz des Albanischen und zur diachronen Phonologie auf Grund des Vokalsystems*, Innsbruck.
Orel, V. 2000: *A concise historical grammar of the Albanian language*, Leiden – Boston – Köln.
Pedersen, H. 1900: Die gutturale im albanesischen, *Zeitschrift für vergleichende Sprachforschung* 36, 270-340.
Pekmezi, Dr. 1908: *Grammatik der albanesischen Sprache*, Wien.
Svane, G. 1992: *Slavische Lehnwörter im Albanischen*, Aarhus.
Thumb, A. 1909: Altgriechische Elemente des Albanesischen, *Indogermanische Forschungen* 26, 1-20.
Topalli, K. 1996: *Për historinë e hundorësisë së zanoreve në gjuhën shqipe*, Tiranë. [with a summary in English]
– 1998: *Zhvillimi historik i diftongjeve të shqipes*, Tiranë. [with a summary in English]
de Vaan, M. 1997: Una gramática de la lengua albanesa (review of: M. Sanz Ledesma, 'El Albanés. Gramática, historia, textos', Madrid, 1996), *Tempus. Revista de actualización científica* 17, 57-63.
– 1999: Review of: Bardhyl Demiraj: *Sistemi i numerimit të gjuhës shqipe: Vështrim diakronik*. In: *Kratylos* 44, 87-93.