

ON THE EVOLUTION OF ARTICLES INTO AGREEMENT MARKERS IN ROMANIAN AND ALBANIAN

Ion Giurgea *

Abstract: This paper tries to elucidate the processes by which former determiners became preposed agreement markers in Romanian and Albanian. In both languages, these markers introduce genitives, agreeing possessors and ordinals. In Albanian the same forms are used as agreement prefixes on all old adjectives and participles and can precede cardinals in definite noun phrases. The fact that these items originate in definite determiners is proven not only by their forms, but also by the possibility of marking the matrix DP as definite when they occur in DP-initial position. I propose that the development definite determiner > agreement marker was made possible by the fact that these languages had specialized definite articles, a suffixal one and an independent, “strong” form which was used when suffixation was impossible. It is the strong form which evolved into specialized agreement markers. Another necessary condition for the reanalysis was the possibility for the strong form to appear in postnominal position, which I assume to have been provided by double- or poly-definite constructions. For Romanian, I propose that the reanalysis of *al* was made possible by the fact that it had restricted contexts of occurrence. For Albanian, where the strong forms must have also been used with adjectives, I adopt the view that a change in the unmarked adjective order from A-N to N-A was the main trigger of the reanalysis, starting from a stage in which postnominal adjectives could only appear in the double definiteness construction, where they were preceded by the article. A further possibility, for Albanian, is the (morphologically triggered) confusion between the strong article and a relativizer stemming from IE **yo-/*yā-*, used to introduce postnominal modifiers.

Keywords: genitive “articles”, preposed agreement markers, historical syntax, Romanian, Albanian

1. Introduction

1.1 Overview of preposed articles in Romanian and Albanian

Romanian and Albanian are peculiar, inside the Indo-European family as well as among the languages of Europe¹, by showing preposed agreement markers (PAM) introducing genitive noun phrases and agreeing possessors. In both languages, these markers, which agree with the “possessee”, introduce noun phrases marked for morphological dative case:

- | | | | |
|-----|----|--|--------|
| (1) | a. | o parte a oraşului | (Rom.) |
| | | a part-F PAM.F.SG city-the.DAT | |
| | b. | një pjesë e qytetit | (Alb.) |
| | | a part(F) F.SG.NOM city.the.DAT | |
| (2) | a. | aceşti prieteni ai mei | (Rom.) |
| | | these friends.M PAM.M.PL my.M.PL | |
| | b. | këta miqtë e mi | (Alb.) |
| | | these friends.M.the PL.NOM/ACC my.M.PL | |

* Institutul de Lingvistică “Iorgu Iordan – Al. Rosetti” and Universităt Konstanz, giurgeaion@yahoo.com.

¹ See Koptjevskaja-Tamm (2003).

Traditionally, these markers are called “articles” (Romanian *articol*, Albanian *nyjë*). However, it must be stressed that they do not have any determiner function in some of their uses, as can be seen from the examples (1)-(3) above. I use therefore the term “preposed agreement marker”, abbreviated PAM. In glosses, I only use PAM for Romanian because its agreement marker *al* can be decomposed into a root *a-* and a ϕ -feature morpheme; for Albanian, where such a decomposition is impossible, I only indicate the agreement features.

The distribution of these items is not identical in the two languages – it is much wider in Albanian – but they share a number of properties: the possibility to mark definiteness in the DP-initial position, and, as I will show in this paper, the origin: both come from the same item as the definite article.

In this paper, I discuss the origin of these markers. After arguing in section 2 that they both originate in strong forms of the definite article, I will try to explain in section 2 why they were reanalyzed into PAMs. In the rest of this section, I present the distribution of these markers in the attested stages of the two languages.

1.2 Genitival and possessive PAMs

With DPs other than personal pronouns, these markers are not attached at the word level, but to the whole noun phrase² (they are phrasal agreement markers), as can be seen, among others, from the fact that they can combine with a coordination of noun phrases³:

- (3) a. primul sindicat al [medicilor și
 first-the trade-union PAM.M.SG physicians-the.DAT and
 asistenților] (Rom.)
 nurses-the.DAT
- b. ministria e [arësimit dhe kulturës] (Alb.)
 ministry.F.the FSG.NOM education.the.DAT and culture.the.DAT

With pronominal possessors, such coordination is ruled out in both languages. In Romanian, pronominal possessors, both agreeing and dative-marker, have been shown to qualify as weak forms in Cardinaletti and Starke’s (1999) typology (see Dobrovie-Sorin and Giurgea 2011):

- (4) * primul sindicat al nostru și *(al) vostru /lor
 first-the trade-union PAM.M.SG our.M.SG and (PAM.M.SG) your.M.PL/them.DAT

In Albanian, the weak character of pronominal possessors has led to a greater differentiation between agreeing possessors and genitives: the PAM has been fused with the pronominal form in a part of the paradigm – see, e.g. the declension of the 2nd singular possessors (the forms representing the article, either fused or not, are boldfaced):

² As noticed by Faensen (1975) for Albanian, who gives the example under reproduced under (3)b here. For Romanian, see Dobrovie-Sorin and Giurgea (2005), Giurgea and Dobrovie-Sorin (2013).

³ The PAM can also be repeated before each conjunct. Traditional grammars of Romanian actually recommend to repeat it.

(5)		m. sg.	f. sg.	m. pl.	f. pl.
	Nom.	yt	jote	e/të tu	e/të tua
	Acc.	tënd	tënde	e/të tu	e/të tua
	Dat.	tënd	sate	të tu	të tua

In Romanian, the PAM is absent when the genitive immediately follows the suffixal definite article:

- (6) sfârșitul luptelor
end-the fights-the.DAT

Syntactic studies (Ortmann and Popescu 2000, Dobrovie-Sorin and Giurgea 2005, 2011, Beavers and Teodorescu 2012, Giurgea and Dobrovie-Sorin 2013) have shown that the absence of the article is a surface structure phenomenon and does not involve a different mechanism of genitive licensing. Thus, if the genitive immediately following the definite noun is coordinated with another genitive, the PAM normally appears on the second genitive, as in (see (7a)); moreover, the PAM must appear if the genitive phrase is modified by a focal particle, as in (7b), and if the definite article is separated from the genitive by a parenthetical, as in (7c). Finally, in the old language, which had case agreement with appositions, the PAM could appear in apposition to a genitive where PAM was absent because of the adjacency with the definite article, as in (7d):

- (7) a. Casa [[Mariei] și [?(a) surorii ei]] a fost vândută.
house-the Maria-the.OBL and AL sister-the.OBL her has been sold
- b. Este casa [chiar [*(a) mamei lui]].
is house.the even AL.F.SG mother.the.OBL his
- c. Începutul, așadar/de altfel, *(al) romanului era plictisitor
beginning-the thus /by-the-way AL novel.the.OBL was boring
- d. sfatul aceluia neam ales de Dumnezeu,
counsel-the that-OBL family chosen by God
al Băsarăbeștilor (Let. Cant. 97)
AL Basarabs-OBL

The obligatory absence of the article after (and only after) the definite article can be analyzed as a null contextual allomorph of the PAM, inserted when PAM is adjacent with suffixal definite D and the two share ϕ -features as a result of agreement (see Giurgea 2013)

The rule of dropping the genitival article after the definite article was not yet fully established in the old language (16th-17th centuries). Although the absence of the article was predominant, the texts show many instances of genitival articles after a definite article⁴:

⁴ See Cornilescu and Nicolae (2009), Giurgea (2013).

- (8) a. **înrimiloru ale** lor (PH 9.38)
 hearts-the.OBL AL.F.PL their
- b. **înțelepciunea a** lu Solomon (CT 25v)
 wisdom-the AL.F.SG OBL Solomon
- c. **oamenii ai** noștri (DÎR XCVII, 1593-1597, Moldavia)
 people-the AL.M.PL our.M.PL

In present-day Albanian, the PAM must appear if the genitive is introduced by a determiner. Bare nouns appear, without the PAM, in a special case form called the “ablative”, which is identical to the dative in the singular but has a distinct ending in the plural, except when the head noun is definite: in this case, both PAM and the suffixal definiteness marking on the bare noun must appear, presumably by a definiteness agreement phenomenon:

- (9) a. (një) shkollë vere
 a school summer.DAT
 ‘(a) summer school’
- b. shkolla e verës /*shkolla vere
 school.DEF.F.SG.NOM F.SG.NOM summer.DEF.F.SG.DAT
 ‘the summer school’

In old Albanian (16th-17th centuries, the oldest text belonging to the 16th century), the genitival article was often absent after a dative or ablative suffixal definite article:

- (10) a. t pârëvet priftënet (Buzuku, apud Çabej 1959)
 PL first.the.PL.DAT priests.the.DAT
 ‘principibus sacerdotum’ [= ‘to the leaders of the priests’]
- b. bririt lopësë (Dozon, apud Çabej 1959)
 horn.the.DAT cow.the.DAT
 ‘to the horn of the cow’

In old Romanian, possessors inside predicative bare NPs normally appeared without the article, in the so-called “adnominal dative” form. The label “dative” is based on the fact that personal pronouns which have special agreeing forms in adnominal context (the so-called “possessive adjectives”) do not take these forms, but take the dative form:

- (11) eu șerbul tău-s și **fecior șerbeei** **tale** (PH 115.7)
 I servant-the your-am and son handmaid-the.DAT your
 ‘I am thy servant, and the son of thy handmaid’
- (12) acel e **frate mie** și **soru mie** și **mumă-mi** easte (T 74r)
 that-one is brother me.DAT and sister me.DAT and mother-me.DAT is
 ‘That one is my brother and my sister and my mother’

1.3 Adjectival uses

In Romanian, the PAM also introduces ordinals:

- (13) uşa a doua
door-the PAM.F.SG second

In Albanian, not only ordinals, but all the old adjectives⁵, as well as participles (when in adjectival positions⁶), are introduced by PAMs, in all positions:

- (14) a. një njeri i mirë / Njeriu është i mirë
a man M.SG.NOM good man.the is M.SG.NOM good
b. një vend i lënë vakant
a place M.SG.NOM left vacant

Although the PAM has the same forms as with genitives, it has a different morphosyntactic status here: that of a word-level affix (a prefix). This is shown by two properties (see Faensen 1975, Campos 2009): (i) unlike the genitival article in (3b), it cannot attach to a coordination – see (15), and (ii) it must appear immediately before the adjective, even if the adjective is not the first constituent of its phrase – thus, in (16), when the adjective is preceded by the degree word *më* ‘more, -er’, the article must appear between the degree word and the adjective⁷:

- (15) një ditë e ngrohtë dhe *(e) bukurë
a day F.SG.NOM warm and FSG.NOM beautiful
‘a warm and beautiful day’

⁵ The conditions for the use of the article are today predominantly morphological (see Çabej 1959 and Buchholz and Fiedler 1987), but have the effect that most adjectives which do not take the article are relational/classificatory. Adjectives which do not take the article are for the most part borrowings (e.g. *aktual* ‘present’, *agresiv* ‘aggressive’), compounds (e.g. *dritëndjeshëm* ‘sensitive to light’, *pendëkuq* ‘with red feathers’) or have initially been nouns (e.g. *besnik* ‘faithful’), which led Riza (1982: 127) to conclude that the article is absent with items recently entered in the category of adjectives – probably, items which entered the class of Albanian adjectives **after** the adjectival article became an agreement affix, part of the morphological make-up of an adjective. Buchholz and Fiedler (1987: 321-322) list the following suffixes which trigger absence of the article: *-ac, -aç, -ak, -al, -an, -ant, -ar, -ash, -at, -cak, -çi, -ent, -esh, -ez, -ik, -il, -in, -ist, -it, -iv, -li, -lli, -man, -njoz, -oid, -or, -ot, -osh, -rak, -ror, -(ë)s/-as, -sor, -shí, -tar, -tor, -uk, -vec*.

⁶ Participles are not introduced by PAMs in compound tenses, but do take PAMs, like adjectives, when the verb *be* is the copula.

⁷ Romanian does not have such an article. The independent (i.e. non-suffixal) definite article *cel* is sometimes misleadingly called “adjectival article” when appearing in the double definiteness construction or in noun phrases with a null (elided) N. These uses, however, are not restricted to adjectives; other modifiers, e.g. prepositional phrases, can appear in those contexts – see (i)-(ii). Moreover, this article has a clear determiner feature [+def], unlike the Albanian adjectival articles, and is attached at the phrasal level. The lack of similarity between *cel* and the Albanian adjectival article has been pointed out by Vătăşescu (1987).

- (i) oamenii cei răi /cei fără prihanë
people-the the bad the without guilt
‘the bad people/people without guilt’
(ii) cele roşii /cele de ieri
the.FPL red/ the.FPL of yesterday
‘the red ones / those from yesterday’

- (16) një shtëpi më e vogël
 a house more FSG.NOM small
 ‘a smaller house’

Moreover, unlike genitive PAMs (on which see 1.3 below), adjectival PAM cannot mark definiteness of the DP in DP-initial position. In such cases, the suffixal definite article appears on the adjective:

- (17) a. i miri
 M.SG.NOM good.the.M.SG.NOM
 ‘the good one’
 b. [të ashtuquajturit] mbiemra pronorë
 PL so-called-the.PL adjectives possessive
 ‘the so-called possessive adjectives’
 c. [e vetmja] zonë e bukur e
 F.SG.NOM sole.the.F.SG.NOM zone F.SG.NOM beautiful F.SG.NOM
 qytetit
 city.the.OBL
 ‘the only nice area of the town’

1.4 DP-initial uses

Although in the uses presented so far PAMs have no determiner feature, there are some uses where they still keep the ability to mark the definiteness of the matrix noun phrase, which represents, as we shall see, a relict of their original status.

This property is found with genitives in both languages: if the genitive is placed at the beginning of a noun phrase, its PAM marks this phrase as definite, so that no suffixal article is necessary on the head noun. In both languages, this construction is current if the head noun is elided (for Albanian, see Riza 1982 for discussion):

- (18) a. Casa Mariei e mai mare decât a Ioanei. (Rom.)
 house.F-the Maria-the.DAT is more big than PAM.F.SG Ioana-the.DAT
 ‘Maria’s house is bigger than Ioana’s.’
 b. Shtëpia e Verës është më e madhe dhe
 house.F.the F.SG.NOM Verë.the.DAT is more FSG.NOM big than
 e Teutës. (Alb.)
 F.SG.NOM Teutë.the.DAT
 ‘Vera’s house is bigger than Teuta’s.’

In Romanian, this use extends to agreeing possessors, whereas Albanian, where the PAM became here a prefix, must use, like with adjectives (see (17a)), the suffixal article here:

- (19) a. Casa Mariei e mai mare decât a mea. (Rom.)
 house.F-the Maria-the.DAT is more big than PAM.F.SG my.F.SG
 ‘Maria’s house is bigger than mine.’

- b. Shtëpia e Verës është më e madhe dhe
 house-F.the FSG.NOM Verë.the.DAT is more F.SG.NOM big than
imja. (Alb.)
 F.SG.NOM-my.the.F.SG.NOM
 ‘Vera’s house is bigger than mine.’

With overt nouns, DP-initial genitives are current in Romanian only if they are *wh-* words. Otherwise they are nowadays obsolete, restricted to the poetic style; they were more frequent in the old language, but they were always rather rare – see (20). DP-initial pronominal possessors (both agreeing and genitive-marked) were quite frequent in the old language (see (20c)) and are still preserved in some north-western varieties (in the standard language, they can sporadically be found, but they are marked):

- (20) a. a căruî casă (Modern Rom.)
 PAM.F.SG which-DAT house
 ‘whose house’
 b. ale idolilor slujbe (Old Rom.: *CV* 79v 3)
 PAM.F.PL idols-the.DAT masses
 c. a mea parte (Old Rom.: 1586, Iași, *DÎR* LXVI 4)
 PAM.F.SG my part

DP-initial genitives are still alive in Aromanian and Meglenoromanian. Notice that the PAM became invariable in these varieties – therefore, I use the gloss *a*:

- (21) [[**ăl** feata noastră] bărbat] ăi doctor (Meglenorom.: Atanasov 2002)
 a + DAT girl-the our man is doctor
 ‘Our daughter’s husband is a doctor.’
 (22) [[**A bîrbasui** mumî] nu avea doru (Arom.: Koutsovlachika, *apud*
 a man-3SG.POSS.M.SG.DAT mother not had pain Campos 2005, (47a))
 ‘Her husband’s mother was not in pain.’

Although the agreement of the genitive marker with the head noun has disappeared in these varieties, the possibility of marking the definiteness of the whole matrix DP by being placed in DP-initial position can only be explained if the genitive marker had once been the variable element *al* in these varieties too, like in (Daco-)Romanian, which was initially the determiner of the matrix noun phrase (see the next section)⁸.

In Albanian, there are some uses of PAMs, not discussed yet, in which they always appear at the beginning of the matrix DP and are correlated with definiteness: before

⁸ Meyer-Lübke (1900), followed by Densusianu (1901), Pușcariu (1913: *DAR*), Papahagi (1937; 1963), Coteanu (1969a, b), considered that this marker (which is also used for the dative in Aromanian) continues the preposition *a* (< lat. *ad*). But this cannot explain the fact that genitives can mark definiteness by being placed in DP-initial position: in no modern Indo-European language do we find adnominal prepositional phrases marking the definiteness of the matrix noun phrase. Therefore in Giurgea (2012) I defend the view that these invariable markers continue the article *al*, which in Aromanian was perhaps contaminated with the preposition *a* and was extended to the dative under the influence of the Greek genitive-dative syncretism.

numerals and *gjithë* ‘all, whole’, PAMs appear optionally in definite noun phrases. In the present-day language, definiteness must be marked by the suffixal article if the noun is expressed, as in (23b), and the PAM seems to be used to underline the totality implied by definiteness (being translatable as ‘both’, ‘all three’, etc.), according to Çabej (1959). In the old language, however, definiteness could be marked exclusively by the PAM, as in (23c):

- (23) a. (të) dy gjuhët (Modern Alb.)
 PL two languages
 b. *të dy gjuhë (Modern Alb.)
 PL two languages
 c. të dymbëdhjetë dishipuj (Old Alb.: Buzuku, apud Demiraj 1986: 322)
 PL twelve disciples
 ‘the twelve disciples’
 d. të dy vëllazënë (Old Alb.: Bogdani, apud Demiraj 1986: 322)
 PL two brothers

With noun ellipsis, the PAM suffices to mark definiteness even in the present-day language:

- (24) Dolën të tre /të tria (Buchholz and Fiedler 1987: 353)
 got-out PL three.M PL three.F
 ‘All three got out’

Another use of this type involves agreeing possessors with kinship terms and *zot* ‘master’. Here even Modern Albanian allows absence of the article on the head noun. As we have seen, the article became fused with the pronominal stem in a part of the paradigm:

- (25) a. im vëlla, ime motër
 M.SG.NOM + my brother F.SG.NOM + my sister
 b. të mi vëllezër, të mi motra
 PL my.PL brothers PL my.PL sisters

This construction is restricted to singular possessors in the present-day language, but plural possessors too were allowed in the old language.

The preservation of an older syntax of possessors with kinship terms is a phenomenon attested in other languages: thus, in Italian and Romanian agreeing possessors allow the absence of the definite article with kinship terms – it. *mio fratello*, calabr. *frátimma*⁹, rom. *frate-miu* ‘my brother’. In Romanian, these agreeing possessors became clitics on the head noun (actually, in the present-day language they have the status of affixes, see Niculescu 2008).

For 3rd person singular possessors however, an unexpected construction is used: the kinship term bears the suffixal article and is preceded by the PAM:

⁹ See Rohlfs (1949, §430) on enclitic/affixal possessors in southern and central Italian dialects.

- (26) *i vëllau, e motra*
 MSG.NOM brother.the FSG.NOM sister.the
 ‘his/her brother, his/her sister’

In this construction, the PAM behaves as if it incorporated, without trace, a 3rd person agreeing possessor. I suspect that the article here has fused with a vocalic clitic representing the possessor, possibly the dative clitic *i* (Albanian has 3rd person clitics consisting in only a vowel: Acc. sg. *e*, pl. *i*, Dat. sg. *i*, pl. *u*), and the vowel of the clitic merged with the vowel of the article, leading to some forms which were identical to those of the article. For the fact that here, unlike with the other prenominal agreeing possessors, the suffixal definite article must appear on the noun, I have no explanation.

In Romanian, the capacity of marking the definiteness of the entire noun phrase is found with all uses of PAMs – thus, besides genitives and possessors, it extends to ordinals. Unlike for genitival *al*, this position is in fact the normal position of ordinals inside definite phrases (the postnominal position illustrated in (13) above is preferred only when the ordinal has a classifying function, e.g. *clasa a doua* ‘the second grade (in schools)’¹⁰:

- (27) *a doua întrebare*
 PAM.F.SG second question

In Albanian, PAMs cannot mark definiteness in NP-initial position with ordinals and adjectives (see (17)), which is probably correlated to the fact that they became agreement prefixes with these categories.

1.5 Syntactic analysis

Albanian prefixal PAMs are agreement morphemes. This analysis can extend to Romanian ordinal PAMs. Genitival PAMs can be analyzed as phrasal agreement morphemes. Giurgea (2011) adopts this analysis for Romanian. However, as the PAM *al* can be decomposed into *a-* and an inflectional morpheme (*a-l*, *a-a* > *a*, *a-i*, *a-le*, see Dobrovie-Sorin 1987, Grosu 1994), it is also possible to analyze *al* as a genitival K head endowed with agreement features. This analysis seems preferable because it can account for the loss of inflection of *al* in a large area of Romanian (northern Daco-Romanian and southern Balkan varieties): disappearance of agreement features is expected on a case head, but would be very surprising on an agreement marker. Note moreover that in Daco-Romanian varieties, the invariable *a* usually preserves the distribution of the PAM *al*.

For the possibility of marking definiteness in the DP-initial position, I propose that PAMs are endowed with a [+def] feature, which licenses a null definite D by agreement:

¹⁰ The ordinal is normally prenominal also in phrases with other determiners, in which case it appears between the determiner and the noun:

(i) *o a doua întrebare*
 a PAM.F.SG second question

- (28) [DP [OrdP a+udef doua]] [D + idef] [NP sală]]
 PAM.FSG second hall

As shown in Dobrovie-Sorin and Giurgea (2006), definiteness marking by placement in SpecDP is also found with superlatives. Moreover, the same mechanism can be adopted for adjectives marked with the suffixal definite article, which occupy a phrasal position; the fact that the definite “article” must be repeated on DP-initial coordinated adjectives or nouns although there is a single D (a single referent, hence a single DP) shows that definiteness is an inflectional **feature** in Romanian, the so-called “article” representing in fact a morpheme marking this feature rather than a clitic (as pointed out by Dimitrova-Vulchanova and Giusti 1998):

- (29) [DP[lunga și anevoioasa] [D [NP ascensiune]]]
 long-the and difficult.the climbing

2. The origin of the preposed agreement markers

The uses in which PAMs mark the definiteness of the entire noun phrase, in DP-initial position, offer a clear indication for the origin of these items: they were once definite determiners of the matrix DP. This immediately explains the agreement with the head noun: as determiners of the head noun, they must agree with it.

Indeed, the most widespread view on the origin of these articles, for both languages, is that the PAMs and the suffixal definite article have the same etymon. For Romanian, this is the Latin distal demonstrative *ille* (*illu(m)*) (as held by most authors, see Cihac 1870, Miklosich 1881, Meyer-Lübke 1895, Tiktin 1895, Pușcariu 1905, *DAR* 1913, Găzdaru 1929, Graur 1929, Gamillscheg 1936, Drăganu 1936-1938, *GLR* 1954: 167, Ciorănescu 1966, Coteanu 1969a,b, Ivănescu 1980, *DEX* 1975, *DELR* 2011, a.o.). A number of researchers saw in the initial vowel *a-* the Latin preposition *ad*, which would have combined with *ille* (Hasdeu 1887, Densusianu 1906, Candrea and Densusianu 1907, Papahagi 1937, Spitzer 1950, Rosetti 1968), but such an origin is impossible, because (i) agreement with the head noun cannot reach into the complement of a preposition (any determiner following the preposition would agree with the head noun of the phrase introduced by the preposition, and not with a noun external to the entire prepositional phrase – i.e. in a construction of the type N_1 [*ad* [*ille* N_2]], *ille* can only agree with N_2 and not with N_1 ¹¹) and (ii) the possibility of marking the definiteness of the matrix noun

¹¹ Among the supporters of the *ad* + *illum* etymology, only two addressed this problem, but their solutions cannot be accepted: Densusianu (1906) proposed that agreement with the possessee started in agreeing possessors; but (i) a construction of the type *ad illum nostrum* can only mean ‘to ours’, not simply ‘ours’ and (ii) prepositional marking never applies to agreeing possessors (all Indo-European languages which have developed prepositional genitives have *either* prepositional marking *or* agreeing possessors, never combine both). Lozbă (1969) proposed that the definite article of the head noun was encliticized on the preposition of a prenominal possessor: *illum [ad meum] parentem* > *ad-illum meum parentem* (> *al meu părinte*). But enclisis only targets stressed words (an enclitic needs the support of an independent phonologic word to its left), and

phrase is not expected for a prenominal prepositional phrase. Notice moreover that besides genitives, agreeing possessors and ordinals, there is another form in which *al* clearly comes from a definite article: the definite alternative old Rom. *alalt* ‘the other’, transparently composed of *al* and *alt* ‘other’. In the old language already, this compound was no longer analyzable, the first part having become uninflected. Therefore it was remarked with the new preposed definite article *cel*, yielding the modern-day forms *celălalt*, *cealaltă* etc. (otherwise, it is only preserved in compound temporal adverbs: *alaltăieri* ‘the day before yesterday’, *alaltăseară* ‘the evening before yesterday’). *Alalt* exists in the southern Balkan dialects – Aromanian *alantu* (< *alaltu* by dissimilation), *anantu*, Meglenoromanian *lalt(u)*, *lant(u)*. This shows that *al* as a preposed definite determiner existed in common Romanian, supporting the claims made in the previous section on the existence of the genitival PAM in Common Romanian, from which the invariable genitival markers of the southern Balkan Romanian dialects evolved.

The phonetic evolution *illu* (> **elu*) > *al(u)* can be explained by regular phonological processes. First, in Late Danubian Latin or Proto-Romanian, short *i* became *e*: thus, *illu(m)* became **elu*, a form preserved in the 3rd personal pronoun *el*¹². Then, as first proposed by Miklosich (1881), unstressed word-initial *e-* became *a-*¹³, as shown by *ericus* > *arici* ‘hedgehog’, *aoace* ‘there’ < *illac-ce*, *arunca* ‘to throw’ < *eruncare*, *ascuți* ‘to sharpen’ < **excotire*, *asmuța* ‘to stir up, incite’ < **exmucciare*, *andrea* ‘knitting needle’ < **endrella*, *amnar* ‘tool used to make fire by hitting the quartz’ < *ignarium* (all from *DELR* 2011)¹⁴. It is important to notice that this explanation leads to the conclusion that *ille* was already unstressed. This means that it had become an article before the *e-* > *a-* rule operated¹⁵.

the preposition *a(d)*, a highly grammaticalized item, cannot be assumed to have been stressed. Syntactically, the displacement *illum [ad meum] parentem* > *ad-illum meum parentem* is inconceivable because short/functional prepositions in Latin and Romance languages cannot be stranded.

¹² This item is pronounced with an initial *i*-glide which is unmarked in writing, due to a special orthographic rule of Romanian which applies to pronouns and forms of the verb *be* (otherwise, the initial glide is written *i*: *iese* ‘gets-out’, *iepure* ‘rabbit’, etc.).

¹³ This explanation for the origin of *al* was endorsed by Tiktin (1895), Găzdaru (1929), Ivănescu (1980). The *e-* > *a-* rule is also accepted by authors who do not consider *ille* as the etymon of *al* (Candrea 1902, Densusianu 1938) or do not discuss its origin (Sala 1970, Avram 2012: 83).

¹⁴ Some of the examples cited in the literature are not sure, because *a-* can represent an accident of Vulgar Latin: the deictic particle *ecce-*, *eccum-* in *acel* ‘that’, *acest* ‘this’, *aci* ‘here’, *acolo* ‘there’ etc. shows forms with *a-* in Ibero-Romance and Occitan (e.g. Sp. *aque!*); *ajuna* ‘fast’, Lat. *ieiunare*, shows *a-* in Sp. *ayunar* and Lat. *aiunus* (attested in Plautus). Other examples where *a-* can represent either Lat. *e-* or *a-* are: *aiępta* ‘throw, turn (towards), indicate’ < *ieictare* or **adiectare*, *asuda* ‘to sweat’ < *assudare* or *exsudare*. See Giurgea (2012, 2013) for a more detailed discussion of this phonetic development, with references.

¹⁵ Latin unstressed *ille* also yielded forms with aphaeresis in Romanian: *lu* (modern Rom. *l, îl*), **ĕ̃>o*, *l’i>i*, *le* and the definite article forms *-lu* (> *-l*), **-(ĕ)a* (**-a-(ĕ)a* > **-ā* > *-a*), *-l’i(>-i)*, *-le*. As I have argued in Giurgea (2012, 2013), the different evolution (aphaeresis in clitics and the suffixal article vs. absence of aphaeresis in *al*) probably follows from the difference between enclisis and proclisis: note that clitic pronouns, which were enclitic in Old Romance (see the well-known Tobler-Mussafia law, see Tobler 1875/1912; Mussafia 1886; Benincă 1995), have aphaeresis in all Romance languages, whereas the definite article, which is preposed in the other Romance languages, sometimes preserve its initial vowel (see Ibero-

In Albanian, the common origin of *i* and the suffixal article is supported by the identity or similarity of most forms:

(30)	PAM	suffixal definite article
m. sg. Nom.	i	-i, -u
sg. Acc.	të/e ¹⁶	-n(ë) (< *-n + -të) ¹⁷
m. sg. Dat.	të	-t
m. sg. Abl.	të, old Alb. also së	-t
n. sg.	të/e ¹⁶	-t
f. sg. Dat. + Abl.	së/të ¹⁸	-s(ë)
pl. Nom. + Acc.	të/e ¹⁶	-t(ë)
pl. Dat.	të	-t / Ø
pl. Abl.	të, old Alb. also së	-t

As several studies have shown (Mann 1977: 103-104, Bokshi 1980, Riza 1982, Orël 2000: 247), the PAM and the suffixal article come from the same item in different positions, and this item was a definite determiner, which in turn comes from a demonstrative. The forms in (30) or forms similar to them appear indeed in the demonstrative and 3rd person pronoun paradigms, preceded by the elements *kë-* (proximal) and *a-* (distal and 3rd person pronoun; it can fall after prepositions):

(31)	m. sg.	f. sg.	n. sg.	m. pl.	f. + n. pl.
Nom.	a-i, a-y; k-y	a-jo; k-jo	a-ta; kë-ta	a-ta; kë-ta	a-to; kë-to
Acc.	a-të; kë-të	a-të; kë-të	a-ta; kë-ta	a-ta; kë-ta	a-to; kë-to
Dat.	a-tij; kë-ti(j)	a-saj; kë-saj	a-tij; kë-ti(j)	a-tyre; kë-tyre	a-tyre; kë-tyre
Abl.	a-si; kë-si	a-so; kë-so	a-si; kë-si	a-sish; kë-sish	a-sosh; kë-sosh

Stressed forms of the proposed articles are maintained in agreeing possessors, where they fused with the pronominal stem at an early date and carried stress (probably by a stress placement rule which counted the distance – in syllables and/or morae – to the end of the word, which explains the variation inside the paradigm in singular possessors: 1st m. sg. Nom. *im*, old Alb. *em*, vs. pl. *të/e mi*, 2nd m. sg. Nom. *yt*, pl. *të/e tu*¹⁹). I give below, for

Romance – Sp. m. sg. *el*, Old Sp. f. sg. *ela* –, some Occitan and Rhaeto-Romance varieties, cf. Meyer-Lübke 1895, §101-105).

¹⁶ In several forms of the paradigm, *e* instead of *të* appears when the article immediately follows the definite article of the head noun.

¹⁷ This explanation for the origin of *-n(ë)*, proposed by Pedersen (1894: 249) and universally accepted afterwards (see Çabej 1959, Demiraj 1973, 1986, Orël 2000), implies that the article was suffixed at an early date, when the accusative Indo-European nasal ending (*-m/n*) was still preserved.

¹⁸ *së* appears in NP-initial position or after a noun carrying the suffixal definite article; otherwise *të* is used.

¹⁹ We find a similar stress shift between singular and plural in a few other words (see Çabej 1960): *dhëndërr/dhëndërrë* ‘son-in-law’, *veshtë* (< **vënëst-*)/*vreshhta* (< **venëst-*) ‘vineyard’, *tjetër* (< **jetër* < **ëter-*)/(*të*) *tjerë* (< **etër-*) ‘other’; in *gjarpër* ‘snake’ (< **sërpVno-*)/*shtërpinj* ‘reptiles’ (< **serpini*), he assumes that the original singular and plural forms evolved into two different words.

comparison with (30) and (31), the stressed forms found before the plural possessor stems *-në* 1st pl. and *-j* 2nd pl. in both numbers and before the 2nd sg. possessor stem *-t-* in the singular, in old Albanian (Buzuku, after Demiraj 1973: 148-150), because for these persons the decomposition into (stressed) article + pronominal possessor is transparent²⁰:

(32)

	m. sg.	f. sg.	n. sg.	m. pl.	f. + n. pl.
Nom.	y- ²¹	jo- ²²	ta-	ta-	to-
Acc.	ta-/tan- ²³	ta-/tan-	ta-	ta-	to-
Dat.	ti-	sa-	ti-	ti-	to-
Abl.	si-	so-	si-	si-	so-

All these paradigms ((30)-(32)) show an alternation between forms with *t* and forms without *t* which is familiar to any Indo-Europeanist, being found in the demonstrative reconstructed as **so*, **sā*, **tod* (Greek δ , η , $\tau\acute{o}$, Sanskrit *sa*, *sā*, *tad*, Gothic *sa*, *so*, *þat*); although the masculine and feminine singular nominative forms are difficult to trace back directly to **so*, **sā*²⁴, it is clear that at least a part of the paradigm continues the Indo-European demonstrative stem *to-/tā-*. As for the forms in *s-*, it is likely that they represent the result of *t-* + *-j-*, although it cannot be ascertained whether this *-j-* comes from endings of the type encountered in Sanskr. *táyā* (sg. f. Instr.), *tásya* (sg. m. Gen.), etc., through

²⁰ The 1st singular forms show forms with the same consonants *t-* and *s-* but with different vocalism: *e* instead of *y-*, *-a-*, *jo-*, *-i-* instead of *-a-* in the feminine dative and ablative. These forms must result from contractions between the article and an initial vocalic part of the pronominal stem (for an initial vocalic part in the 1st person root, see Ancient Greek Acc. $\acute{\epsilon}\mu\acute{\epsilon}$ ‘me.ACC’ vs. $\sigma\acute{\epsilon}$ ‘you.ACC’, agreeing possessor $\acute{\epsilon}\mu\acute{o}\varsigma$ ‘my’ vs. $\sigma\acute{o}\varsigma$ ‘your’).

²¹ Mann (1977: 117) also cites *uj-* in Buzuku.

²² The forms in *-o* are lengthened to *-uo-* before the *-j* of the 2nd plural. Likewise, we find *sā-* for *sa-*, m. pl. *tā-* for *ta-*.

²³ *-n* preserved before the 2nd person singular root *t-*, which becomes *-d-*: m. *tand*, f. *tande*.

²⁴ Pedersen (1894: 252) tried to derive *y-* and *jo-* from **so-* and **sā-*, by proposing that *-j-* was added as a hiatus glide after *s > h > Ø* between vowels. But forms with *-j-* are not restricted to the paradigms with a preposed particle (those in (31)); *j-* is also found word-initially – see (32), e.g. *jotë* ‘your.F.SG.NOM’, *jonë* ‘our.F.SG.NOM’; the element *j* can also explain the palatal vowel of the articles *i* and *e*; the definite article *-a* probably comes from **-ā-yā*. But an evolution *s- > j-* is not attested for Albanian: *s-* before stressed vowels, or at least some of them, evolved to *gj-* (see Çabej 1960; Orël 2000: 60-61; this evolution is expected after a front vowel, but it seems to appear in a few words after central or back vowels – the most convincing examples are *gjumë* ‘sleep’ < IE **supno-*, *gjallë* ‘alive’: Lat. *saluus* ‘safe, sound’, Gk. $\acute{\omicron}\lambda\omicron\varsigma$ ‘entire, complete’); in some words, it is reflected as *sh-* (e.g. *shi* ‘rain’, cf. Gk. $\upsilon\epsilon\iota$ ‘it rains’). Other researchers traced back these forms to the IE. **i-*, **yo-* stems (Meyer 1892, Orël 2000). But there is some evidence for IE *y-* > Alb. *gj-* word-initially (see Çabej 1960). Orël (2000) proposes that IE *y-* yielded *gj* before IE *a*, *e*, *ō*, but was preserved as *j-* before IE *ā* and *u* (besides the form *jo* discussed here, the only convincing example he cites is 2nd pl. *ju*), hence the form *jo*. As for the m. Nom. sg., he assumes the IE prototype **ei(s)/is* (cf. Lat. *is*, Sanskr. *ay-ām*), but this does not explain the *y* vowel, which seems to be older, judging from the possessive forms (see (32)). A possible explanation for the vowel *y* is that in **-os* the final *s*, before falling, triggered raising of the vowel (perhaps during a *h* stage), like in Slavic (see Halla-aho 2006 for a detailed argumentation in favor of such a law in Slavic); we would then have IE **(e)yos > Proto-Alb. *iju* or **ju > y*. The root **y(o/ā)-* might be also reflected in the 3rd person clitic forms, explaining the front vowel in Acc. sg. *e*, Acc. pl. *i*, Dat. sg. *i*. Mann (1977: 116) proposes an etymon IE **syo-*, **syā-*, but there is no independent evidence for an evolution *sy- > j-*.

loss of the thematic vowel, as proposed by Pedersen (1900), or was part of a special demonstrative stem (**t(i)̄io*), as proposed by Brugmann (1893: 321).

Summing up, in both languages the preposed and suffixal articles have a common origin, which is a former demonstrative. This etymology raises the following questions: why did the same determiner evolve into two items, a suffixal and a preposed article? How did the preposed articles lose the determiner feature, evolving into agreement markers for specific adnominal and predicative constituents (genitives and ordinals in both languages, and also adjectives in Albanian)? Is there a historical relation between these developments in the two languages?

The first question is easy to answer if we have a look at the present-day definite determiner system of Romanian. This language requires that definiteness should be marked on the first constituent of the noun phrase. Therefore, in contexts where definiteness cannot be marked by the suffixal article on the first constituent, for morphological reasons, or by a DP-initial phrase bearing *al*, an independent form of the definite article is used, *cel*, which comes from a short form of the distal demonstrative *acel*. *Cel* appears before cardinals, before an elided/empty noun (see Giurgea 2010 on empty nouns in Romanian), before a postnominal modifier (typically an adjective) in the double definiteness construction, before the comparative degree word *mai* ‘more, -er’ in superlatives and before ordinals preceded by *de*:

- (33) a. *cei doisprezece apostoli*
 the twelve apostles
 b. *modelele vechi, dar și cele [NØ] noi*
 models-the old but also the.F.PL new
 ‘The old models, but also the new ones’
 c. *modelul cel nou*
 model-the the new
 d. *cea mai bună soluție*
 the more good solution
 ‘the best solution’
 e. *cel de-al doilea caz*
 the of-PAM second case
 ‘the second case’

This shows that a language which normally marks the definite article by inflectional morphology may need in some contexts an independent form of the article. We are thus led to the idea that *al* once functioned as such an independent form, like *cel* in contemporary Romanian. The same holds for Albanian *i*. It is true that present-day Albanian does not use a special independent form, but resorts to suffixation in most counterparts of (33) (in cases of noun ellipsis, the distal demonstrative can play the role of a definite article, like in English *the essays on freedom and those on poverty*); this is possible because Albanian does not require that definiteness should be marked on the first constituent of the phrase:

- (34) a. *(të) dymbëdhjetë apostujt*
 (PL) twelve apostles.the

- b. modelet e vjetra, por edhe të rejat
 models.the PL old but also PL new.the

But, as the examples (23)c-d)) suggest, this has not always been the case. The preposed article in (34a)) allowed absence of the suffixal article on the noun in the old language. We can infer that at some previous unattested stage, the preposed article was used precisely because the noun was not capable of bearing the suffixal definite article, due to the fact that it did not occur in the first position of the noun phrase.

The parallel with Romanian *cel* is supported by the fact that *al* (< Lat. *illu(m)*) was once opposed, as a short distal demonstrative form, to *acel* (< Lat. *ecce illum*). A recurrent pattern in the evolution of demonstrative systems is that some forms become weakened and evolve into articles or anaphoric demonstratives unmarked for the distance contrasts, whereas new forms are created by remarking with deictic particles (see, e.g. French *celui*, which has become a strong form of the definite article used with empty N, and it must be reinforced with *-ci* and *-là* to be able to function as a demonstrative pronoun). In Romanian, after *al* was reanalyzed into a PAM, a new short demonstrative evolved into an article – the form *cel*, a shortened form of *acel*.

Summing up, we can conclude that the PAMs continue a previously preposed, independent (i.e. non-affixal) form of the definite article. The emergence of a special preposed article is conditioned by the existence of a suffixal article, which is a common property of both languages, because in languages with a suffixal article, in environments where suffixation is blocked for syntactic reasons, an independent determiner form is needed in order to express definiteness.

Finally, it must be stressed that the syntactic environments which impose a strong form of the article need not be the same nowadays and in the unattested stage of the two languages which we are interested in. We have already seen this for Albanian in the discussion around (34), and in the next section we shall see further differences.

For Romanian, based on the distribution of the *al*-forms in the attested phases of the language, we are led to assume that the suffixation was impossible when the noun was preceded by genitives, agreeing possessives, ordinals and the alternative *alt* ‘other’²⁵. For prenominal genitives and ordinals, this impossibility is expected: prenominal genitives blocked suffixation because they were phrasal constituents, full DPs with their own ϕ -features, to which another ϕ -element, with different ϕ -features, could not be added; ordinals already contain a suffixal article, which is used not to express definiteness but to derive ordinals from cardinals (in the present-day language, it is only in the feminine that the suffixal article and the ordinal suffix are identical, i.e. *-a*; the masculine has a new form *-lea*; but the old language had *-lu* and *-le* used according to the final vowel of the cardinal exactly like the masculine definite article: *patru-lu* ‘fourth’, *şase-le* ‘sixth’, *optu-lu* ‘eighth’, etc.). Agreeing possessors may have blocked suffixation because they were often weak forms, possibly lacking word stress, and enclitics must attach to a stressed word.

²⁵ This claim holds for the unattested stage of the language when *al* functioned as an article. In the attested Old Romanian period, when *al* already functioned as a genitival PAM, cases of co-occurrence of prenominal *al* and the definite article are attested (see Cornilescu and Nicolae 2011a), but they are much less frequent than the type in (20), in which the DP-initial *al*-phrase suffices to mark the DP as definite.

For Albanian, since the PAM also appears with adjectives, it seems that the strong article was also required with prenominal adjectives (see the discussion in the next section). The possibility of a system in which the article can be suffixed on the noun but not on an adjective is attested in Scandinavian languages; notice that in most varieties a strong form is used when an adjective appears before the noun, exactly as I propose for Proto-Albanian:

- (35) a. flickan (Swedish)
 girl.the
 b. den vackra flicka
 the beautiful girl

3. The reanalysis of determiners into agreement markers

3.1 Preliminaries

What is more difficult to explain is how an independent definite article came to lose definiteness and become obligatory with certain types of adnominal phrases, functioning as an agreement marker.

In sum, what we have to explain is how a structure of the type (36), where X is a genitive, agreeing possessor or ordinal, has been replaced by the structures in (37) (*al* is the Romanian PAM, *i* is the Albanian one: I give the masculine singular nominative form); for Albanian, we have to add X = adjective, for (37b'); as we have seen, in Albanian ((37a)), where the article is capable of marking definiteness of the entire noun phrase, is only found with genitives, possessives and cardinals:

- (36) [DP [D *al/i*] [X [N...]]]
 (37) a. [DP [XP[_{Agr}*al/i*]_{+def} X] [[D_Ø_{+def}] [N...]] (PAM marking definiteness)
 b. [DP D [... ([XP [al/i] X ..] .. N .. ([XP [al/i] X ..])..]] (phrase-level PAM, no +def)
 b'. [DP D [... ([X₀[_{Agr}*al/i*]-X) .. N .. ([X₀[_{Agr}*al/i*]-X) ..]] (word-level PAM, no +def)

In Romanian, (37b) applies to genitives (today only postnominal; the old language allowed prenominal *al*-genitives between D and N, in a marked order) and (37b') to ordinals. As for Albanian, if definite article suffixation is analyzed as licensing of a null [+def] D by agreement with an N or A marked [+def], the structure in (37b') also applies to the word order [Art-A_{+def}] N (on which see section 1, example (17).

The crucial step separating (36) from (37) is the reanalysis of *al/i* from D into Agr – or, for genitival *al* in Romanian, K + Agr as shown in section 1.4; moreover, this Agr is specialized for certain categories (X), whereas in (36) the definite D is realized as an independent article due to the syntactic context which blocks the suffixation.

Regarding (37a), one could also assume, instead of a null D licensed by a [+def] PAM, that the items *al* and *i* are ambiguous between Agr and definite D. However, given the fact that these items are specialized for certain adnominal constituents and that

Romanian has a different independent definite article (*cel*), I prefer to analyze them uniformly as Agr and to add, for these environments, a [+def] feature on Agr which licenses a null [+def] D.

3.2 Romanian

Discussing only the Romanian case, in Giurgea (2012) I proposed the following circumstances which supported the reanalysis of *al*: (i) *al* + X constituents could appear in postnominal position, co-occurring with another D, in the double definiteness construction; (ii) the independent article appeared in a restricted number of contexts, more restricted than those of the present-day independent article *cel*, and in some of these contexts, but not all, it came to be in functional competition with the distal demonstrative.

Among these conditions, only (i) can be assumed to have held for Albanian. (ii) is a point on which the two languages differ: whereas in Romanian PAMs are not found with adjectives, in Albanian they are, which implies that the distribution of the strong article was wider, probably similar to that of Romanian *cel* nowadays.

Let us first examine Romanian. In the oldest attested phase of the language (16th century; some manuscripts may have 15th century lost originals), the reanalysis of *al(u)* is already completed. Based on the traces of *al(u)* in Aromanian and Meglenoromanian, which show the same specialization as in Romanian (genitives and possessives, ordinals and *alalt* ‘the other’), we can assume that the reanalysis took place in Common Romanian already (an unattested stage of the language in which the ancestors of the present-day North- and South-Danubian dialects were not separated by communities speaking other languages). Therefore, the conditions which promoted the reanalysis of *al(u)* cannot be directly observed, but must be reconstructed with the help of the old language and of the southern Balkan dialects.

On this basis, it can be argued that the reanalysis D > Agr specialized for some constituents was favored by the fact that the preposed article had limited contexts of occurrence and in some of its contexts, but not in all, it came into competition with the distal demonstrative (which, under the shortened form *cel*, was to become the new independent definite article). These two circumstances had the effect that *al* co-occurred very frequently with agreeing possessors, genitives, ordinals and *alt* ‘other’.

More precisely, comparing the reconstructed distribution of the article **alu* with the modern independent article *cel*, there is some evidence that **alu* was not needed in one of the contexts in which *cel* is very often found now, namely with N-ellipsis + adjective. It is possible that common Romanian allowed article suffixation in a context such as (38a) – (see (33b) above), like Modern Albanian, realizing it as (38b):

- (38) a. *cel* [[_NØ] nou]
 the new ‘the new one’
 b. *noul*
 new-the

Old Romanian shows constructions in which the suffixal definite article appears on adjectives and participles which are not nominalized, as shown by the preservation of the

adjectival or verbal syntax, which means that they rely on N ellipsis; in the modern language, *cel* is obligatory in such contexts:

- (39) nece îmblaiu cu nalții, nece cu [**mai minunații**
 nor walked.1SG with tall.the.M.PL nor with more wonderful-the.M.PL
de mine] (CP 255v 14-15)
 than me
 ‘I didn’t walk with the tall ones or with those more wonderful than me.’
- (40) Și [**răstignitul cu nusul**] împută lui (T 107r)
 and crucified-the.M.SG with him scolded him.DAT
 ‘And the one crucified with him scolded him.’

Likewise, when the missing noun can only be inferred from the context, as there is no stable association of the adjective with some nominal concept which could be encoded in the lexicon, we must assume N-ellipsis rather than nominalization. We find in some contexts of this sort the suffixal article in the old language, which is impossible in the modern language:

- (41) lua-se-va cusătura ei noa de spre veachea (T 72r)
 take-REFL-will seam-the.F.SG its.F. new-the.F.SG from on old-the.F.SG
 ‘Its new seam will take away from the old one.’
- (41) *Cusătura ei nouă se va lua de pe vechea (Modern Rom.)
 seam-the its.F. new REFL will take from on old-the

Another context in which *cel* is followed by adjectives in the present-day language is the double definiteness construction. Here again, Old Romanian attests the possibility of using the suffixal article (see Croitor 2008 on double and multiple definiteness in Old Romanian):

- (42) Ascunde-i-veri ei în furișea feței tale de **învăluitura**
 hide-them-will.2SG them in secret.the face.the.DAT your from adversity-the
omenreasca (PH 30.21)
 human-the
 ‘Thou wilt hide them in the secret of Thy face from the human adversity.’

No strong article appeared in the superlative because the superlative was not yet grammaticalized as a special construction; like in modern Italian and Albanian, the superlative meaning was simply conveyed by a comparative inside a definite DP. This stage is still preserved in Old Romanian:

- (43) a. mai bunul pământului Egiptului eu voiu da
 more good-the earth-the.DAT Egypt-the.DAT I will give
 voao (PO 161, Genesis 45.18)
 you.DAT
 ‘I’ll give you the best of Egypt’s land.’

- b. ucise mai mulții lor (CP 148v 14-15)
killed.3SG more many.the their
'He killed most of them.'
- c. mai marea parte a .Austrii (Ist. T.R., 71)
more big.the part PAM.F.SG Austria-OBL
'most of Austria'

In predicative position and with adverbs, the comparative could be used in order to convey a superlative meaning:

- (44) Aceasta să fie mai sfântă la voi (PO 284, Exodus 30.36)
this SĂ be.3SG.SUBJ more holy at you.PL
'This (one) should be (holier=) the holiest among you.'
- (45) cine au fost **mai mult** și **mai întâi** pricina micșorării și
who has been more much and more before cause-the decline-the.OBL and
stricării monarhii romane (Ist. T.R., 60)
destruction-the.OBL monarchy-the.OBL Roman
'who was for the most part and in the first place the cause of the decline and
destruction of the Roman Empire'

Marking of definiteness by the suffixal article on adjectives in noun ellipsis, double definiteness and superlative contexts is also found in Aromanian:

- (46) a. Feata ațea marea easti profesoarâ → (Ațea) marea easti
girl-the that big-the is teacher (that) big-the is
profesoară. (Caragiu-Marioțeanu and Saramandu 2005: 170)
teacher
'The elder daughter is a teacher.' → 'The elder is a teacher.'
- b. unlu lo pričili mascuri, anantu **0eaminli** (Papahagi 1905: 49)
one-the took animals-the male other-the female-the
'One took the male animals, the other the female ones.'
- c. ma búnlu caș eára ațél. (Grămostean dialect, Livezi: interview
more good-the cheese was that *Hoara a noastă Livădz*, Nevaci 2013)
'The best cheese was that one.'

Another context in which it is likely that common Romanian did not resort to the preposed articles is that of cardinals in definite phrases (where Albanian has the article, as we have seen). In Aromanian, the definite article is suffixed on the numeral (e.g. *doil'i* 'two.the.M.PL', *doauli* 'two.the.F.PL'). Old Romanian shows already the construction with *cel*, but also suffixation on the noun (e.g. *doaosprădzeace neamurele lu Israil* 'twelve tribes.the DAT Israel' = 'Israel's twelve tribes', CV 55r 8-9) and absence of any marking²⁶. Suffixation on a noun which is not NP-initial was probably not possible in common Romanian, because otherwise it could have prevented the emergence of a

²⁶ See Nicolae (forthcoming) on quantifier + definite N constructions in Old Romanian.

definite article. Therefore, I suppose that the Aromanian construction continues the common Romanian use.

Besides the restricted contexts, the reanalysis of *al(u)* may have been favored by the competition with the distal demonstrative in a part of the contexts of *al(u)* – namely, in phrases without an overt noun (based on noun ellipsis or a null non-anaphoric N). Such a use of the demonstrative can indeed be found in English and in Modern Romanian:

- (47) a. The article on the oil crisis has already been published. The one/**That** on Tibet was not accepted by the editor.
- b. Articolul despre criza petrolieră a apărut deja. Cel/**Acela** despre Tibet nu a fost acceptat de către redactor.
- on Tibet not has been accepted by editor

In some languages, distal demonstrative forms become the only way of expressing the definite article in some N-ellipsis contexts – thus, Italian must use *quello* ‘that’ in the context in (47). French went a step further in this direction: the distal demonstrative *celui* became a form of the definite article specialized for N-ellipsis contexts²⁷, because in the demonstrative uses, it was reinforced with the deictic particles *-là* and *-ci*. The development can be clearly seen if we put side by side Old French, where the definite article could appear with empty N, and Modern French, where *celui* must appear in such contexts:

- (48) a. Mon non, douce dame, vous di, /et si vous ai dit
my name sweet lady you.DAT tell.1SG and so you.DAT have.1SG told
le mon pere (Old French: Gerbert de Montreuil, La continuation de
the my father *Perceval*, 3080-1, apud Foulet 1980: 53)
- b. Mon nom, chère dame, je vous le dis, et je vous ai dit **celui** de mon père.
my name dear lady I you.DAT it tell.1SG and I you.DAT have.1SG
told that of my father

An explanation of this use of the distal demonstrative is beyond the scope of this article²⁸.

As a result of these facts, **alu* was the only way to express the definite article in phrases with prenominal possessors (agreeing as well as dative-marked NPs), ordinals and *alt* ‘other’, whereas in phrases with an empty N and postnominal constituents other than adjectives, both **alu* and the distal demonstrative **(a)celu* could be used in order to

²⁷ Corblin (1995) analyzes *celui* as D+pro-N. I prefer a D + [_NØ] analysis because it can be shown that French nominal ellipsis relies on deletion, being an instance of ellipsis with internal structure, or “surface anaphora” in Hankamer and Sag’s (1976) terms), as I argue in Giurgea (2010).

²⁸ This use is probably related to the use of distal or unmarked demonstratives with the same semantic effect as the definite article in cases of relatives or post-nominal modifiers which may be analyzed as reduced relatives, discussed in Wolter (2006), e.g. *That student who scored one hundred on the exam is a genius* (Wolter 2006: 111, example (1)) = *The student who scored one hundred on the exam is a genius*.

express the meaning of the definite article. This situation provided a basis for the reanalysis of **alu* as a marker specialized for possessors, ordinals and *alt*.

The second condition for the reanalysis was the possibility of the *al* + X string appearing in postnominal position, co-occurring with another determiner. I hypothesize that this possibility was offered by the double definiteness construction. In this construction, besides the definite article of the entire noun phrase, which appears as a suffix, there is a second definite article or demonstrative introducing a postnominal modifier (adjective, PP, relative clause)²⁹; in present-day Romanian, only the independent definite article *cel* is allowed, but in the old language, with adjectives, the suffixal article could appear on the adjective, sometimes co-occurring with the articles or demonstratives *cel* and *cela* (see also (42) above); in the southern Balkan dialects, the distal demonstrative is always used, and sometimes it co-occurs with the suffixal article on the adjective (see also (46a) above):

- (49) a. omul cel bun (Modern Rom.)
man.the the good
- b. împăratul cel mare (Old Rom.: PH 135.17)
emperor.the the big
- c. omul cela rromileanul (Old Rom.: CV 22 4-5)
man.the the Roman.the
- d. aspida surda ce-și astupă urechile
asp-the deaf-the that-REFL.3SG.DAT covers ears-the
sale (Old Rom.: PH 57.5)
POSS.3SG.F.PL
'the deaf asp that covers its ears'
- e. omul ațel bun(lu) (Arom.) (ALR I, map 185)
man-the that good(-the)
- f. fițoru țela marle (Meglenorom.) (ALR I, map 185)
boy-the that big-the

Although this construction resembles an apposition, it cannot be analyzed as an apposition, because the modifier preceded by *cel* can be followed by restrictive modifiers and complements of the noun (if they are heavy: genitives, PPs):

- (50) a. teoria cea nouă a multîmilor (Modern Rom.)
theory-the the new PAM sets-the.DAT
'the new set theory'
- b. numele cel svânt a lui (Old Rom.: PH 32.21)
name-the the holy PAM he.DAT
'his holy name'

²⁹ The semantics of this construction is not easy to pin down. For recent syntactic analyses, see Cornilescu and Nicolae (2011b), Giurgea (2013).

Since this construction exists in old Romanian and the southern Balkan dialects, it can be assumed to be old. Moreover, since it is not found in the other Romance languages, but is current in Greek (where more than two definite articles can appear, hence the name “polydefiniteness”), where it is found already in the ancient language, it can be assumed to be one of those Balkan features of Romanian which have evolved under the influence of Greek, in the period in which these languages were in contact, in the Eastern Roman Empire (see Sandfeld (1930) on other Greek-based features of the Balkan *Sprachbund*).

We have thus reasons to assume that the double definiteness construction existed in the common Romanian stage we are interested in. With modifiers which did not allow the suffixal article, i.e. other than adjectives, we expect **alu* to have been used, since this was the independent article of the time. We can thus reconstruct the following constructions:

- (51) a. N-def (...) *alu* X, with X= genitive/possessive or ordinal
 b. N-def (..) A-def

Now, I believe there are sufficient conditions to trigger the reanalysis (36) →(37): **alu* mostly appeared with a restricted number of X’s – genitives, possessives, ordinals, *alt* – and the string *alu*-X could also appear in postnominal position, co-occurring with another determiner. The reanalysis is completed as soon as this determiner is not restricted to the definite article. The oldest texts already reached this stage, as *al* can be found in indefinite noun phrases and noun phrases with no determiner and non-referential interpretation:

- (52) a. Pătru, apostol a lu Isus Hristos (CV 69v 5)
 Pătru apostle PAM the.DAT Jesus Christ
 b. multe trupure moarte ale sfinților (T 64v)
 many bodies dead PAM holy-the.DAT

In Giurgea (2013), I also develop an account for the failure of *al* to generalize in certain contexts: as shown in section 1.2 (examples (6)-(8)), *al* can be absent immediately after the definite article and, in the old language, with predicative bare nouns (which can take the so-called “adnominal dative”, see (11)-(12) above). For space reasons, I will just mention here the main idea of the account I propose: the generalization of *al* first appeared with possessives, which could appear as weak forms only in a specifier immediately below D, and otherwise required the use of *al*, reanalyzed as a way to form a strong possessive form. Analyzing article suffixation as raising to D, this implies that possessives without *al* could appear after the N only when the N bears the suffixal article. Therefore, postnominal *al* was not generalized in the environment [N + D]³⁰. As for the adnominal dative, pronominal possessors could take dative forms, instead of the agreeing forms (which, being weak, would have required *al*), as a result of the reanalysis of a Latin

³⁰ For the fact that ordinal *al* is not dropped after the article (e.g. *anul al treilea* ‘year-the PAM third’ not **anul treilea*), I assume that the ordinal was always pronominal – it couldn’t appear after N raised to D. The postnominal ordinal, thus, originates in the double definiteness construction exclusively, which explains the fact that it always has *al*. This assumption is supported by the fact that ordinals are postnominal only in definite DPs (as noticed by Grosu 1994), except when the ordinal is reanalyzed as a classifying adjective (compare **o problemă a doua* ‘a second problem’ with *o clasă a doua* ‘a second grade’).

construction involving a relational dative, the type *ille mihi nepos est* ‘to me, he is a nephew’.

3.3 Albanian

Turning now to Albanian, it is clear that an explanation based on the restricted syntactic contexts of the strong article cannot work, because in this language PAMs have a wider distribution, being used with adjectives and cardinals.

Riza (1982) proposed the following evolution: first, the normal placement of adjectives was prenominal and the definite article appeared as a preposed article in this case, yielding the order Art-Adj-N (e.g. **e madhe dorë* ‘ART big door’³¹). The string Art-Adj could appear after the noun as a second noun phrase in apposition (with an empty noun, in the terms used here; he talks about nominalization). Later, in this construction Art lost its definiteness feature and the Art-Adj changed its status from an apposition to a restrictive modifier.

This last step of the evolution is not sufficiently justified, in my opinion. What I would like to retain from this explanation is the idea that the unmarked placement of adjectives was once prenominal. This is in agreement with what we know about old Indo-European languages. Moreover, as Riza himself notes, evidence for an old prenominal placement of modifiers comes from the prenominal placement of pronominal possessors with kinship terms (see section 1, where I mentioned the parallel with Romanian and Italian, where possessors with kinship terms reflect an older syntax).

Regarding the use of the postnominal Art-Adj string, I assume that it represented the double definiteness construction, which is something distinct from apposition: as we have already seen for Romanian, the ‘definite’ modifier is included in the noun phrase of the head noun, rather than forming a separate DP (see(50)); moreover, the modifier can be restrictive, and in Greek polydefinites it is necessarily restrictive (see Alexiadou 2001, Kolliakou 2004, Campos and Stavrou 2005, Alexiadou et al. 2007). Although Modern Albanian lacks this construction, we can assume, based on the many syntactic similarities with Romanian and Greek (see the literature on the Balkan *Sprachbund*, Sandfeld 1930, Solta 1980, Tomić 2006, a.o.), that pre-literary Albanian did have this construction at some point of its history^{32, 33}. Since the modifier in the double definiteness construction is

³¹ “*” here indicates reconstruction; such a sequence is impossible in Modern Albanian but assumed to have existed at a previous stage of the language.

³² Campos (2009) points out that 16th century Albanian (Buzuku’s *Meshari*) did have a double definiteness construction, but the examples he gives are superlatives, which are known to allow more easily double definiteness (see French *le jour le plus beau* ‘the day the more beautiful’, but **le jour le beau* ‘the day the beautiful’). Note that double definiteness at this stage of the language already implies the use of the suffixal article on the adjective, because the adjectival article is already a prefixal agreement marker in Buzuku (see its placement after the degree word *mâ* ‘more, -er’):

(i) Jakobnë birrë e saj mâ të vogëlinë (Buzuku, Gen. 27:15, apud Campos
Jacob.the.ACC son SG.ACC her more SG.ACC little.the.ACC 2009: 48a)
‘Jacob, her youngest son’

³³ Mann (1977: 103) considers the Hellenistic Greek construction *ho lógos ho emós* as an “antecedent” for the Albanian genitival and adjectival articles, but he treats this issue together with the issue of the postposition of the definite article (so that he also mentions Late Latin as an “antecedent” – cf. Vulgata *sermo ille dei* –

syntactically integrated and can have a restrictive interpretation, this construction is a better candidate for the missing link between the old *i*-Adj sequence with *i* = D and the new *i*-Adj sequence with *i* = Agr.

But why did Albanian reanalyze the double definiteness construction with adjectives, whereas other languages which have it, such as Romanian and Greek never did so? Here I think Riza's hypothesis that adjectives were once prenominal is a good starting point for an explanation. Note that in Greek, adjectives are normally prenominal, except in the polydefiniteness construction; in a definite DP, they must be prenominal, unless polydefiniteness is used (Alexiadou et al. 2007: 287). Suppose that Albanian was once similar to Greek: in order for an adjective to appear in postnominal position, it had to be introduced by the article. Now, let us suppose there was a tendency towards a postnominal placement of the adjective. This tendency may have been due to the close contacts with Latin and Early Balkan Romance (Proto-Romanian); in classical Latin the N-Adj order was already much more common than in the other Indo-European languages of a similar age (see Bauer 2009: 263-265, and references therein), and in the evolution towards the Romance languages this order became unmarked in most varieties. The tendency towards N-Adj order in Albanian may have led to the loss of the special interpretation associated to double definiteness³⁴. However, the article *i* was still kept as a means of attaching an adjectival modifier after the noun. If postnominal adjectives in languages with normal Adj-N order are indeed reduced relative clauses (see Larson 1998 for English), it may be assumed that *i* was reanalyzed as a functional element introducing reduced relatives. From this moment, the restriction to definite noun phrases was lost, and *i* + Adj sequences started to appear in postnominal position with any determiner. The last step in the evolution was also triggered by the bias towards the N-Adj order. Since for intersective modifiers reduced relatives and direct modification have the same semantic contribution, postnominal adjectives replaced prenominal ones completely and *i* changed from a relativizing marker into an agreement marker. The last step of the evolution is the transformation of *i* into a prefix, a word-level agreement marker (see section 1.2).

Like in Romanian, a necessary condition for the reanalysis of *i* has been the availability of distal demonstratives for the empty N context. In the present-day language, the distal demonstrative is indeed the only way of expressing definiteness with empty N followed by a PP or relative clause (with genitive and adjectives, it is optional, because, as we have seen in (17)a above, the preposed genitival article can license definiteness and adjectives allow suffixation of the article):

- (53) Rritje e kompetencave të Parlamentit, duke përfshirë
 increase F.SG.NOM competencies.the.DAT PL Parliament.the.DAT including

although this is not a double definiteness construction). The two phenomena must be neatly distinguished: first Albanian acquired the definite article, which has been established as a suffix, whereby a preposed (independent) article appeared in cases suffixation was impossible. Later, the independent article evolved into an agreement marker, and it is for this development that double definiteness is relevant.

³⁴ In Greek, this special interpretation has been described as narrow focus by Campos and Stavrou (2005). In Romanian, this is a possibility, but it is by no means necessary. A thorough description of all the meaning nuances of this construction in Romanian is still to be made. See Cornilescu and Nicolae (2011b), Giurgea (2013).

- ato për emërimin e Presidentit të**
 those for nomination.the.ACC SG.ACC president.the.DAT M.SG.DAT
Komisionit ([http://eeas.europa.eu/delegations/albania/documents/more_](http://eeas.europa.eu/delegations/albania/documents/more_commission.the.DAT_info/newsletter)
 commission.the.DAT info/newsletter)
 ‘an increase of the competencies of the Parliament, including those for the nomination of the president of the Commission’
- (54) Nga problemet e shumta do të trajtohen vetëm **ato që lidhen me historinë e gjuhë.**
 of problems.the PL.NOM/ACC many will SUBJ treat.M.PAS.3PL only those that bind.M.PAS.3PL with history.the.ACC SG.ACC language.the.DAT
 (Buchholz and Fiedler 1987: 294)
 ‘Of the many problems, only those which are connected with the history of the language will be examined.’
- (55) Duhet të zhvillohet industria energjetike, sidomos **ajo e naftës**
 must SUBJ develop.M.PAS.3.SG industry.F.the energetic especially that.F.SG.NOM F.SG.NOM oil.the.DAT
 (Buchholz and Fiedler 1987: 294)
 ‘The energy industry must be developed, especially the oil one.’

Summing up, I tried to explain the fact that the double definiteness construction gave rise to a prefixal agreement marker on adjectives in Albanian but not in Romanian or Greek by the hypothesis of a shift in adjective placement from Adj-N to N-Adj which took place in Albanian at a time when the language already had articles and double definiteness. Romanian had N-Adj order at least as one of the unmarked orders from the very beginning, judging from Latin³⁵ (in any case before the emergence of the definite articles), and Greek kept the A-N order as the basic, unmarked order, to this day.

4. Open issues: A possible second source of the preposed articles in Albanian and the relation with Romanian

Because of the fact that no form of Albanian older than the 16th century is attested, we cannot completely exclude a second source for the emergence of preposed articles, although there is no independent evidence for it in the language. I have proposed in the end of the previous section that the independent article with postnominal adjectives was reanalyzed, at some stage, as a relativizer. But what if there was such a relativizer from the very beginning, and the forms of the two items (the preposed article and the relativizer) were confounded? The confusion might have been rooted in the forms and enhanced by the existence of the double definiteness construction.

The parallel which immediately comes to mind is the Iranian *ezāfe*. In several Iranian languages, postnominal constituents (adjectives, possessives) require the attachment of a suffix (modern Persian *-e*, with the variant *-ye* after vowels) to the

³⁵ According to Gianollo (2012), the percentages of N-Adj and Adj-N are the same in the Late Latin texts *Passio Perpetuae et Felicitatis* and *Peregrinatio Egeriae*, which are known for representing the spoken language.

constituent preceding them³⁶. This suffix originates in a relative pronoun which could introduce non-finite constituents, showing agreement with the head noun, already in Old Persian:

- (56) a. kāram tayam hamiçiyam (Old Persian: Schmitt 1989: 75)
 army.M.ACC REL.M.SG.ACC rebellious.M.SG.ACC
- b. upa tam çarətaṃ jaṃ
 on that.ACC course.ACC REL.ACC
 darəgəṃ (Avestan: YT 19.77, apud Hewson and Bubeník 2006: 139)
 long.ACC

In Avestan, the relative Indo-Iranian pronoun *ya-* is preserved, but Old Persian has forms coming from the sequence demonstrative (Indo-Ir. *sa-/ta-*) + relative pronoun (Indo-Ir. *ya-*) – see Schmitt (1989), and (56a) above. The origin of the construction must have been a finite sentence with a nominal predicate ('(the) army, which (is) rebellious'), but case agreement spread to the nominal predicate, turning the construction into a non-finite one. Thereby the relativizer was extended to reduced relatives.

This construction resembles the Albanian construction of postnominal modifiers in several respects: first, in Old Persian, it contains the demonstrative *sa-/ta-*, like the Albanian article, at least in a part of its paradigm (see the discussion in section 2); if the Albanian forms beginning with a vowel or *j-* contain the IE **-yo-/yā* (see the discussion in footnote 24), we can also draw a parallel between the Albanian forms and the second part of the Iranian relativizer. The *s-* forms of the article have been traced back to **tya-* (see section 2), which closely resembles the Iranian *ta-ya-* forms. Syntactically, this construction resembles the Albanian one by the fact that it is used to introduce postnominal adjectives and genitives. Notice that Iranian languages remained head-final to this day (they are OV languages). In these conditions, the unusual head-initial order in noun phrases has been made possible by taking relative constructions as a starting point (relative clauses, due to their weight, are regularly postposed in various Indo-European OV languages, such as West Germanic, Ancient Greek, Latin). This situation resembles the one I proposed for the old, unattested stage of Albanian before the reanalysis (normal prenominal placement of modifiers, special marker needed for the postnominal construction).

There is however an important difference between the Albanian construction and the Iranian one, which may be the reason why this parallel has remained almost totally unnoticed in the literature on the history of Albanian³⁷: as we have seen, Albanian *i* certainly continues a definite article, at least in the prenominal use. Iranian languages do not have a definite article. The *haya-/taya-* forms and the *ezāfe* are not used to express definiteness.

This is why I talked about a possible confusion between a relativizing marker and an independent article. For the forms without initial consonant of the old demonstrative,

³⁶ Syntactically, there is evidence that the *ezāfe* forms a constituent with the adnominal phrase it introduces, being only phonologically attached to the preceding word, as shown by Ruff (2009).

³⁷ The only exception I know of is Hasdeu (1879: 672-676).

which appear as *y-* and *jo-* in agreeing possessors, several researchers (Meyer 1891: 5, Orël 2000: see f.n. 24) proposed an IE etymon **ei-/y(o)-/yā-* (cf. Lat. *i-s*, *e-a* ‘he/she, that’ < **eyā*, Sanskr. *ay-am*, *iy-am*, *id-am* ‘this’, Old Slavic 3rd person Acc. *i*, *j*³, Lith. 3rd person Nom. *jis*, *ji* ‘he, she, it’, Acc. *jì*, *jà*, Gothic *is* ‘he’, *ita* ‘it’, etc.). The stems **yo-/yā-* are also those of relative pronouns in Indo-Iranian, Greek, and (combined with the enclitic particle *-že*) in Old Slavic. Thus, Old Slavic shows that the same root can yield both demonstratives/personal pronouns and relatives in one and the same language. We may hypothesize that (Proto-) Albanian had once forms of the **ei-/y(o)-/yā-* stems as both demonstratives and relativizers, the demonstrative ones being distinguished from relativizers by having **to-/tā-* (and possibly also **tyo-/tyā-*) forms in a part of a paradigm. With postnominal modifiers, both the preposed article coming from the demonstrative and the relativizer could both be used – the former in the double definiteness construction –, which led to a confusion of the two paradigms. Thereby the double definiteness construction disappeared but gave some of its forms to the relativizer, which acquired the *t-/s-* forms.

Possible evidence in favor of this hypothesis comes from the unexplained alternation between *e* and *të* in a part of the paradigm of the preposed articles: when the preposed articles (adjectival and genitival alike) immediately follow the suffixal definite article, it has the form *e* instead of *të* for the singular accusative and plural nominative and accusative:

- (57) a. (këtë /një) ditë të bukur
 this.ACC/a day SG.ACC beautiful
 b. ditën {e/*të} bukur
 day.the.ACC e/të beautiful
 c. të mi vëllezër
 PL.NOM/ACC my.MPL brothers
 d. vëllezërit {e/*të} mi
 brothers.the.NOM/ACC e/të my.MPL
 e. yjet {e/*të} mëdha dhe {të/*e} nxehta
 stars.the.NOM/ACC e/të big and të/e hot

The *e* forms, whose distribution is today completely dictated by the (surface structure) morpho-phonological environment, might represent the descendants of the **yo-/yā-* relative, which, after the merger of this paradigm with that of the independent article, remained as variants used for euphonic purposes, in order to avoid *-të të* sequences (where the first *-t* is the definite article). What remains unexplained is why *e* was not generalized to all the forms where the preceding definite article had a *t(ë)* forms (e.g. why *njeriut të mirë* ‘man.the.DAT ART good’ is allowed, and there is no **njeriut e mirë*).

If the preposed articles in Albanian have this second source which allowed their extension to the postnominal use correlated with the absence of any determiner feature, then the issue of the relation between the developments in Albanian and Romanian might be reconsidered. In the previous section, I tried to explain the evolution which led from articles to agreement markers taking into account the systems of the two languages and

without resorting to external factors, with the only exception of the change in adjective placement in Albanian.

However, if the generalization of the article in postnominal position in Albanian came as a result of the confusion of the preposed article and the relativizer introducing reduced relatives into a single item, an asymmetry appears between the languages: Romanian never had such a relativizer (we can be sure about this, because there is no such item in Latin or any Romance language, and *al* cannot come from a Latin relative pronoun). One may suppose that the generalization of *al* in postnominal position in Romanian possessors and ordinals was favored by the Proto-Albanian construction, or, perhaps, by a lost language which had a construction similar to the Proto-Albanian one³⁸. However, the phenomenon cannot be reduced to a simple syntactic borrowing. This can be seen from the fact that PAMs are only used with possessors and ordinals in Romanian, whereas Albanian uses them also with adjectives and cardinals. Moreover, as we have seen in 1.2 and 3.2, genitival PAMs have not been generalized to all contexts in Romanian, which is not the case for Albanian. This leads to the conclusion that the internal syntactic conditions of the language have been crucial for the reanalysis, which cannot be reduced to a syntactic borrowing.

If it is true that the distribution of genitival *al* is due to the generalization of the pattern of possessives (see 3.2 above), one might say that the loss of the definiteness feature of *al* in postnominal and predicative possessives was influenced by the existence, in Proto-Albanian or a related language, of a similar element functioning as a definiteness marker in the DP-initial position but lacking any determiner feature in the postnominal and predicative position.

Summing up, due to the very late date at which the Albanian branch of the Indo-European family is attested, we cannot be sure whether the only source of the PAM was the strong definite article (originating in a demonstrative), or there was also a second source for the postnominal use, a relativizer whose forms were confused with those of the article. For Romanian, the only source of the preposed article is the Latin distal demonstrative *ille*. Because it is certain that Romanian and Proto-Albanian or a closely related language were in contact at some unattested stage of these languages, it is possible that one language has influenced the other, although the reanalyses show different results – each operating according to the specific syntactic conditions of the language. As for the direction of this influence, in case the loss of the definiteness feature of the preposed article in Proto-Albanian is due to the confusion with the relativizer, it can be assumed that Proto-Albanian (or another closely related Old Balkan language) has been the source. It is also possible that there was no influence, but rather the similar development was due to the similarities between the initial conditions of the two systems: each language had a strong and a suffixal article and the double definiteness construction.

³⁸ Based on the geographical distribution, the substratum of Romanian is widely believed to be Daco-Moesian, a language of the Thracian family, spoken both North and South of the Danube (see Poghirc 1969 for discussion), whereas Albanian is spoken on a territory where Illyrian was spoken in Antiquity. Therefore, although there is an important number of substratum elements in Romanian which have close Albanian counterparts, the possibility of an influence from a language distinct from Proto-Albanian must be taken into consideration.

5. Conclusions

In Romanian and Albanian, preposed agreement markers specialized for certain types of constituents have developed from former definite article forms. I have tried to reconstruct the conditions which made this evolution possible: (i) These languages had specialized forms of the definite article: the most common was a suffixal form; when suffixation was impossible for syntactic reasons, special independent forms were used (which I called “strong forms”); as these forms often co-occurred with certain adnominal constituents, they could be reinterpreted as specialized introductory elements of those constituents; (ii) A necessary condition for the reanalysis of the strong article forms was the possibility of co-occurring with other determiners; this possibility was provided by the double definiteness construction, in which [Art + XP] constituents could appear in postnominal position; (iii) In one of the environments typical for the strong forms, namely the N-ellipsis environment, strong forms could enter competition with distal demonstrative forms. Besides these general conditions, specific factors operated in each language.

In Romanian, the strong forms may have had a quite restricted distribution because article suffixation was possible with adjectives (either prenominal, postnominal with N-ellipsis or in the double definiteness construction) and possibly even cardinals. This favored the reanalysis of the strong forms. Moreover, the emergence of the genitive *al* may have been due to the restriction of possessives to weak forms, *al* being first reanalyzed as a strong possessive marker and later extended to inflectional genitives.

In (Proto-)Albanian, the strong forms must have had a wider distribution, as they yielded the adjectival PAM and the PAM used with cardinals. Here, the reanalysis of the strong forms was probably triggered by a shift of the unmarked order of adjectives from Adj-N to N-Adj. At the first stage, postnominal adjectives had to be preceded by the preposed article, which functioned as a determiner in the double definiteness construction and possibly also as a relativizer. As this order became more frequent, possibly under Latin or Proto-Romance influence, the preposed article has been reinterpreted as an introductory element for adjectives. It is also possible that in postnominal position the preposed article is the result of a formal confusion between the strong forms of the definite article in the double definiteness construction and a relativizer used to introduce reduced relatives, stemming from IE **yo-/*yā-*, with a parallel in Old Persian *haya-/taya-* and the modern Persian *ezāfe*.

Corpus

ALR I 1942 = S. Pop (coord.), *Atlasul lingvistic român*, vol. 2, Leipzig, Harrassowitz, 1942.

Ist. Ţ.R. = Stolnicul Constantin Cantacuzino, *Istoria Ţării Rumâneşti*. In M. Gregorian (ed.), *Cronicari Munteni*, vol. I, 1-80. Bucharest, Editura pentru Literatură, 1961.

CP = Coresi, *Psaltire slavo-română*, Braşov, 1577, edited by S. Toma, Bucharest, Editura Academiei, 1976.

T = Coresi, *Tetraevanghel*, Braşov, 1561, edited by F. Dimitrescu, Bucharest, Editura Academiei, 1963.

CV = *Codicele Voroneţean*, [1563-1583], edited by M. Costinescu, Bucharest, Minerva, 1981.

DÎR = *Documente şi însemnări româneşti din secolul al XVI-lea*, edited by G. Chivu, M. Georgescu, M. Ioniţă, A. Mareş and A. Roman-Moraru, Bucharest, Editura Academiei, 1979.

- Let. Cant.* = *Letopiseșul Cantacuzinesc* [c. 1690]. In M. Gregorian (ed.), *Cronicari Munteni*, I, 81-224. Bucharest: Editura pentru Literatură, 1961.
- Papahagi, P. 1905. *Basmе aromâne și glosar*. Bucharest: Institutul de arte grafice „Carol Göbl”.
- PH* = *Psaltirea Hurmuzaki* [c. 1490-1516], edited by I. Gheție, M. Teodorescu, Bucharest, Editura Academiei Române, 2005.
- PO* – *Palia*, Orăștie, 1582, edited by V. Pamfil. Bucharest, Editura Academiei, 1968.

References

- Alexiadou, A. 2001. Adjective syntax and noun raising. Word order asymmetries in the DPs as the result of adjective distribution. *Studia Linguistica* 55: 217- 248.
- Alexiadou, A., Haegeman, L., Stavrou, M. 2007. *Noun Phrase in the Generative Perspective*. Berlin · New York: Mouton de Gruyter.
- Atanasov, P. 2002. *Meglenoromâna azi*. Bucharest: Editura Academiei.
- Avram, A. 2012. *Studii de fonetică istorică a limbii române*. Bucharest: Editura Academiei.
- Bauer, B. 2009. Word order. In P. Baldi and P. Cuzzolin (eds.), *New Perspectives on Historical Latin Syntax*, vol. I, *Syntax of the Sentence*, 241-316. Berlin · New York: Mouton de Gruyter.
- Beavers, J., Teodorescu, A. 2012. Repetition effects in the syntactic domain. A new analysis of Romanian possessives. Paper presented at the Annual Conference of the English Department, May 31-June 2, University of Bucharest.
- Benincă, P. 1995. Complement clitics in medieval Romance: The Tobler-Mussafia law. In A. Battye and I. Roberts (eds.), *Clause Structure and Language Change*, 325-344. Oxford: Oxford University Press.
- Bokshi, B. 1980. *Rruga e formimit të fleksionit të sotëm nominal të shqipes*. Prishtina: Akademia e shkencave dhe e arteve e Kosovës.
- Brugmann, K. 1893. (with B. Delbrück) *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen*, vol. II, *Lehre von der Wortformen und ihrem Gebrauch*. Strasbourg: Karl J. Trübner.
- Buchholz, O., Fiedler, W. 1987. *Albanische Grammatik*. Leipzig: Verlag Enzyklopädie.
- Çabej, E. 1959. Unele probleme ale istoriei limbii albaneze. *Studii și cercetări lingvistice* X (4): 527-560.
- Çabej, E. 1960. *Hyrje në historinë e gjuhës shqipe: Fonetika historike e shqipes*. Prishtina: Enti i teksteve dhe i mjeteve mësimore i Krahinës Socialiste Autonome të Kosovës. Romanian translation by Adriana Ionescu. Bucharest: Editura Universității din București, 1997.
- Campos, H. 2005. Dative and genitive variations in two dialects of Greek Aromanian. In M. Karali, D. Koutsogiannis and N. Liosis (eds.), *Meletes gia tin elliniki glossa. (Studies on the Greek Language)*, 156-167. Thessalonica: Aristotelian University of Thessaloniki.
- Campos, H. 2009. Some notes on adjectival articles in Albanian. *Lingua* 119: 1009-1034.
- Campos, H., Stavrou, M. 2005. Polydefinite constructions in Modern Greek and Aromanian. In O. Milosevic, (ed.), *Balkan Syntax and Semantics*, 137-173. Amsterdam/Philadelphia: John Benjamins.
- Candrea, I.-A. 1902. *Les éléments latins de la langue roumaine. Le consonantisme*. Paris: Librairie Emile Bouillon.
- Candrea, I.-A., Densusianu, O. 1907. *Dicționarul etimologic al limbii române. Elementele latine (A-Putea)*, 1st part. Bucharest: Socec.
- Caragiu-Marioțeanu, M., Saramandu, N. 2005. *Manual de aromână – carti trā învītari armâneaști*. Bucharest: Editura Academiei.
- Cardinaletti, A., Starke, M. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In H. van Riemsdijk (ed.), *Clitics in the Languages of Europe*, 145-234. Berlin: Mouton de Gruyter.
- Cihac, A. de. 1870. *Dictionnaire d'étymologie daco-romane*, vol. I. *Éléments latins, comparés avec les autres langues romanes*. Frankfurt am Main: Ludolphe St. Goar; Berlin: A. Asher; Bucharest: Socec.
- Ciorănescu, A. 1966. *Diccionario etimológico rumano*. La Laguna, Tenerife: Biblioteca Filológica.
- Corblin, F. 1995. *Les formes de reprise dans le discours*. Rennes: Presses Universitaires de Rennes.
- Cornilescu, A., Nicolae, A. 2009. Evoluția articolului hotărât și genitivul în româna veche. In R., Zafiu, G. Stoica and M. Constantinescu (eds.), *Limba română: teme actuale*, 647-667. Bucharest: Editura Universității din București.

- Cornilescu, A., and Nicolae, A. 2011a. On the syntax of Romanian definite phrases: Changes in the patterns of definiteness checking. In P. Sleeman and H. Perridon (eds.), *The Noun Phrase in Romance and Germanic. Structure, Variation, and Change*, 193-221. Amsterdam/Philadelphia: John Benjamins.
- Cornilescu, A., Nicolae, A. 2011b. Nominal peripheries and phase structure in the Romanian DP. *Revue roumaine de linguistique* LVI (1): 35-68.
- Coteanu, I. 1969a. Româna comună. II. Morfologia. 3. Articolul; 5. Pronumele. In I. Coteanu (ed.), *Istoria limbii române*, vol. II, 232-254. Bucharest: Editura Academiei.
- Coteanu, I. 1969b. *Morfologia numelui în protoromână*. Bucharest: Editura Academiei.
- Croitor, B. 2008. Aspecte privind acordul în determinare în limba română veche. In G. Pană Dindelegan (ed.), *Limba română. Dinamica limbii, dinamica interpretării*, 213-218. Bucharest: Editura Universității din București.
- DAR 1913 = S. Pușcariu (coord.). 1913. *Dicționarul limbii române*, vol. I, A-B. Bucharest: Librăriile Socec & Comp. și C. Sfetea.
- DELRL 2011: Sala, M., Avram, A., Mihăiescu, D. (coords.). 2011. *Dicționarul etimologic al limbii române*, vol. I, A-B. Bucharest: Editura Academiei.
- DEX 1975 = Coteanu, I., Seche, L., Seche, M. (coords.). 1975. *Dicționarul explicativ al limbii române*. Bucharest: Editura Academiei Republicii Socialiste România.
- Demiraj, S. 1973. *Morfologjia historike e gjuhës shqipe*. Tirana: Universiteti i Tiranës, Fakulteti i Historisë dhe i Filologjisë.
- Demiraj, S. 1986. *Gramatikë historike e gjuhës shqipe*. Tirana: 8 nëntori.
- Densusianu, O. 1906. Report in *Buletinul Societății Filologice*, II, 16, Bucharest.
- Densusianu, O. 1938. *Histoire de la langue roumaine*, vol. II. *Le XVI^e siècle*. Paris: Ernest Leroux.
- Dimitrova-Vulchanova, M., Giusti, G. 1998. Fragments of Balkan nominal structure. In A. Alexiadou and C. Wilder (eds.) *Possessors, Predicates and Movement in the Determiner Phrase*, 333-360. Amsterdam/Philadelphia: John Benjamins.
- Dobrovie-Sorin, C. 1987. *Syntaxe du roumain. Chaînes thématiques*. PhD dissertation, Université Paris 7.
- Dobrovie-Sorin, C., Giurgea, I. 2005. Romanian genitives and determiners. *Bucharest Working Papers in Linguistics* VII (1): 89-101.
- Dobrovie-Sorin, C., Giurgea, I. 2006. The suffixation of definite articles in Balkan languages. *Revue roumaine de linguistique* LI (1): 113-135.
- Dobrovie-Sorin, C., Giurgea, I. 2011. Pronominal possessors and Feature Uniqueness. *Language* 87 (1): 127-157.
- Drăganu, N. 1936-1938. Recensii. *Dacoromania* IX: 273-312.
- Faensen, J. 1975. Genitiv und Adjektiv im Albanischen. *Zeitschrift für Balkanologie* XI (2): 40-47.
- Foulet, L. 1980. *Petite syntaxe de l'ancien français*. Paris: Librairie Honoré Champion (1st edition: 1919).
- Gamillscheg, E. 1936. *Zum romanischen Artikel und Possessivpronomen*. Sonderausgabe aus dem Sitzungsberichten der preussischen Akademie der Wissenschaften. Phil.-hist. Klasse XXVII, Berlin.
- Găzdaru, D. 1929. *Descendenții demonstrativului latinesc ille în limba română*. Iași: Viața Românească.
- Gianollo, C. 2012. Prepositional genitives in Romance and the issue of parallel development. In C. Galves, S. Cyrino, R. Lopes, F. Sandalo, and J. Avelar (eds.), *Parameter Theory and Linguistic Change*, 281-303. Oxford: Oxford University Press.
- Giurgea, I. 2010. *Pronoms, déterminants et ellipse nominale. Une approche minimaliste*. Bucharest: Editura Universității din București.
- Giurgea, I. 2011. Agreeing possessors and the theory of case. *Bucharest Working Papers in Linguistics* XIII (2): 5-35.
- Giurgea, I. 2012. The origin of the Romanian “possessive-genitival article” *al* and the development of the demonstrative system. *Revue roumaine de linguistique* LVII (1): 35-65.
- Giurgea, I. 2013. *Originea articolului posesiv-genitival și evoluția sistemului demonstrativelor în română*. Bucharest: Editura Fundației pentru Știință și Artă.
- Giurgea, I., Dobrovie-Sorin, C. 2013. Nominal and pronominal possessors in Romanian. In A. Carlier and J.-C. Verstraete (eds.), *Genitive Case and Genitive Construction*, 105-139. Amsterdam/Philadelphia: John Benjamins.
- GLR, 1954. *Gramatica limbii române*, coord. D. Macrea. Bucharest: Editura Academiei.
- Graur, A. 1929. A propos de l'article postposé. *Romania* LV: 475-481.
- Grosu, A. 1994. *Three Studies in Locality and Case*. London, New York: Routledge.

- Halla-aho, J. *Problems of Proto-Slavic Historical Nominal Morphology*. Helsinki: Helsinki University Press.
- Hankamer, J., Sag, I. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7 (3): 391-428.
- Hasdeu, B. P. 1879. *Cuvente den bătrâni*, vol. II, *Cărțile poporane ale românilor în secolul XVI în legătură cu literatura poporană cea nescrisă. Studiu de filologie comparativă*. Bucharest: Noua Tipografie Națională C. N. Rădulescu.
- Hasdeu, B. P. 1887. *Etymologicum Magnum Romaniae*, vols. I-II. Bucharest: Socec și Teclu.
- Hewson, J., Bubenik, V. 2006. *From Case to Adposition. The Development of Configurational Syntax in Indo-European Languages*. Amsterdam/Philadelphia: John Benjamins.
- Ivănescu, Gh. 1980. *Istoria limbii române*. Iași: Editura Junimea.
- Larson, R. 1998. Events and modification in nominals. In D. Strolovitch and A. Lawson (eds.), *Proceedings of Semantics and Linguistic Theory VIII*, 145-168. Ithaca, NY: Cornell University CLC Publications.
- Kolliakou, D. 2004. Monadic definites and polydefinites: Their form meaning and use. *Journal of Linguistics* 40: 263-333.
- Koptjevskaja-Tamm, M. 2003. Possessive noun phrases in the languages of Europe. In F. Plank (ed.), *Noun Phrase Structure in the Languages of Europe*, 621-722. Berlin · New York: Mouton de Gruyter.
- Lozbă, M. 1969. „A” – prepoziție, articol posesiv, pronume posesiv de reluare. *Analele Științifice ale Universității „Al. I. Cuza” din Iași* III, XV.
- Mann, S. E. 1977. *An Albanian Historical Grammar*. Hamburg: Buske.
- Meyer, G. 1891. *Etymologisches Wörterbuch der albanesischen Sprache*. Strasbourg: Trübner.
- Meyer-Lübke, W. 1895. *Grammaire comparée des langues romanes*, vol. II, *Morphologie*. Translated from German by Auguste Doutrepoint and Georges Doutrepoint. Paris: H. Welter.
- Meyer-Lübke, W. 1900. *Grammaire comparée des langues romanes*, vol. III, *Syntaxe*. Translated from German by Auguste Doutrepoint and Georges Doutrepoint. Paris: H. Welter.
- Miklosich, F. 1881. *Beiträge zur Lautlehre der rumunischen Dialekte*. Vienna: Gerold.
- Mussafia, A. 1886. Una particolarità sintattica della lingua italiana dei primi secoli. In G. I. Ascoli et al. (eds.), *Miscellanea di filologia e linguistica. In memoria di Napoleone Caix e Ugo Angelo Canello*, 255-261. Florence: Successori Le Monnier.
- Nevaci, M. 2013. *Identitate românească în context balcanic*. Bucharest: Editura Fundației pentru Știință și Artă.
- Nicolae, A. forthcoming. Notă de sintaxă comparată: parametrul [+definit] în sintaxa grupului nominal românesc. *Limba română* LXII (2) [also available at ling.suf.net/lingbuzz/00174/current.pdf].
- Niculescu, D. 2008. *Mijloace lingvistice de exprimare a posesiei în limba română*. Bucharest: Editura Universității din București.
- Orël, V. 2000. *A Concise Historical Grammar of the Albanian Language: Reconstruction of Proto-Albanian*. Leiden, Boston, Köln: Brill.
- Ortmann, A., Popescu, A. 2000. Hapology involving morphologically bound and free elements: Evidence from Romanian. *Yearbook of Morphology*: 43-70.
- Papahagi, T. 1937. *Din morfologia limbei române*. Bucharest: Socec.
- Pedersen, H. 1894. Bidrag til den albanesiske sproghistorie. In *Festskrift til Vilhelm Thomsen*, 246-257. Copenhagen: Gyldendal.
- Pedersen, H. 1900. Die Gutturale im Albanesischen. *Zeitschrift für vergleichende Sprachwissenschaft* 36: 277-340.
- Poghirc, C. 1969. Influența autohtonă. In I. Coteanu (ed.), *Istoria limbii române*, vol. II, 313-365. Bucharest: Editura Academiei.
- Pușcariu, S. 1905. *Etymologisches Wörterbuch der rumänischen Sprache*, vol. I, *Lateinisches Element mit Berücksichtigung aller romanischen Sprachen*. Heidelberg: K. Winter.
- Riza, S. 1982. Probleme të njave të shqipes. *Studime Filologjike* 36 (1): 121-139.
- Rohlf, G. 1949. *Historische Grammatik der italienischen Sprache und ihrer Mundarten*, vol. II, *Formenlehre und Syntax*. Bern: A. Francke Ag. Verlag.
- Rosetti, A. 1968. *Istoria limbii române. De la origini până în secolul al XVII-lea*. Bucharest: Editura pentru Literatură.
- Ruff, J. 2009. Locus and linkers. http://www.langsci.ucl.ac.uk/linguistics/reading_groups/syntax.
- Sala, M. 1970. *Contribuții la fonetica istorică a limbii române*. Bucharest: Editura Academiei.
- Sandfeld, K. 1930. *Linguistique balkanique. Problèmes et résultats*. Paris: Édouard Champion.

- Schmitt, R. 1989. Altpersisch. In R. Schmitt (ed.), *Compendium Linguarum Iranicarum*, 56-85. Wiesbaden: Dr. Ludwig Reichert Verlag .
- Solta, G. R. 1980. *Einführung in die Balkanlinguistik mit besonderer Berücksichtigung des Substrats und des Balkanlateinischen*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Spitzer, L. 1950. Romanian “un prieten al meu” ‘a friend of mine’. *Word* 6: 141-159.
- Tiktin, H. 1895. *Rumänisch-Deutsches Wörterbuch*, vol. I. Bucharest: Imprimeria Statului.
- Tobler, A. 1875. Review of Jules Le Coultre, de l'ordre des mots dans Crestien de Troyes. *Göttingische Gelehrte Anzeigen* 34: 1057-1082.
- Tomić, O. M. 2006. *Balkan Sprachbund Morpho-syntactic Features*. Dordrecht: Springer.
- Vătăşescu, C. 1987. Quelques remarques concernant l'article possessif et l'article démonstratif dans le roumain et en albanais. *Studia albanica* 24 (2): 141-149.
- Wolter, L. K. 2006. That's That: The Semantics and Pragmatics of Demonstrative Noun Phrases. PhD dissertation, University of California, Santa Cruz.

