THE LOW DEFINITE ARTICLE AND THE EVOLUTION OF THE ROMANIAN DP*

Alexandra Cornilescu (alexandracornilescu@yahoo.com)

1. BACKGROUND

1.1. Aim and outline of the paper

The Romanian definite article singles out Romanian among Romance languages since it is a suffixal enclitic form, rather than a free standing proclitic one, as elsewhere in Romance. On the other hand, this same property aligns Romanian with the Balkan languages where, as stressed by Tomić (2006: 3), there has been “grammaticalization of the category of definiteness through postpositive definite articles.”

The enclitic definite article is special not only through its position, but also through its more limited distribution, which has led to the emergence of a second, proclitic article, used when the enclitic one is not available. Romanian now possesses two definite articles: the enclitic bound morpheme –(u)l (Lat. ille), and the proclitic free standing morpheme cel, a reduced form of the demonstrative acel ‘that’, grammaticalized as an article within Old Romanian.

(1) a. băieți
   boys.DEF
   ‘the boys’

b. cei doi băieți
   CEL two boys
   ‘the two boys’

The aim of this paper is twofold: On the one hand, in the first part, we examine previous analyses and present a minimalist account of the syntax of the definite article(s), focusing on the enclitic definite article –(u)l. On the other hand, in the second part, we discuss the role of the enclitic definite article in the evolution of the Romanian DPs, as determined by what we have called the low definite article, a structure which is no longer available in Modern Romanian.

We propose that the loss of the low definite article is the consequence of stronger locality conditions on definiteness checking. All the many constructions involving the low definite article or related to it went out of the language when constraints on definiteness checking strengthened. The same locality constraints on definiteness checking account for the emergence of the free standing article cel.
1.2. The basic data

While the observation that Romanian has two definite articles goes back to the 19th century (see Nicolae (2013b) for the history of the problem), awareness that the two morphemes are in complementary distribution, spelling out the same definite determiner is of a more recent date (for an early extensive presentation see Cornilescu (1992)).

Before presenting their distribution, let us mention that both –(u)l and cel are φ-complete and vary for gender, number and case.

The distribution of the enclitic definite article may be summed up by saying that it is suffixed on nouns and on prenominal adjectives, therefore it is suffixed on [+N] constituents, as shown in (2a-c). Moreover, if a DP contains two [+N] constituents (2b, c), the definite article occurs only once on the first. As to cel, it selects quantificational constituents, usually described as φ-incomplete (3a, b) and it rejects [+N] constituents so that it cannot precede bare nouns or nouns modified only by adjectives (3c, d):

(2)  
a. băiatul  
    boy. The  
    ‘the boy’

b. frumosul băiat  
    beautiful. the boy  
    ‘the beautiful boy’

c. * băiatul frumos  
    boy. the beautiful  
    ‘the beautiful boy’

(3)  
a. cei doi băieți  
    CEL two boys  
    ‘the two boys’

b. cei câțiva băieți  
    CEL few boys

c. * cel băiat  
    CEL boy

d. * cel frumos băiat

Accepting that the functional structure of the DP is as outlined in Borer (2005), i.e. DP>QP>NumP>AP>NP, the complementarity of the two articles is immediately apparent if one considers the relevant part of the internal structure of the constituents headed by –ul, respectively by cel (in bold in (4)).

(4)  
* - (u)l  
    DP  (AP)  NP

    cel  
    DP  QP  (AP)  NP

Considering the distribution in (2) and (3) under Abney’s DP-hypothesis, a significant empirical generalization emerges, regarding the realization of definiteness in Romanian:
Definiteness realization in Modern Romanian (informal)

a. A definite article or a constituent bearing it occupies the DP-initial position.

b. The choice of the definite article is categorically determined. [+N] constituents (i.e. nouns and adjectives) select -l; [-N, +Q] constituents (i.e. quantifiers: cardinals, ordinals, degree quantifiers) select cel.

Concluding on the distribution of the two articles, -l selects [+N] constituents, while cel occurs elsewhere. From now on we reserve the term definite article for the enclitic morpheme, referring to cel as the ‘freestanding’ or ‘proclitic’ article.

1.3. An intriguing diachronic variation

In Old Romanian (=OR) the definite article variably occurs either on the first N or A, or on a lower N, so that another nominal constituent, for instance, a (non-definite) A(djective), may precede the definite N(oun):

(6)   * OR: spre ticăloase cuvintele mele audzul îți pleacă… (Cantemir)
       to vicious words.the my hearing your turn
       ‘Lend your ear to my vicious words.’
       (MR: spre ticăloasele cuvinte ale mele audzul îți pleacă…)

The article which appears on a constituent different from the first N or A will be referred to as “the low definite article” for convenience. It is the second aim of this paper to explore this type of variation and to examine its significance for the evolution of the Romanian DP.

2. PREVIOUS ANALYSIS RELEVANT TO THE PRESENT DISCUSSION

2.1. The status of the article: a clitic or a suffix ?

Even restricting ourselves to the generative studies of the past thirty years, the literature devoted to the Romanian definite article is considerable and mirrors both the evolution of generative theorizing and a better understanding of a more extended fragment of Romanian. Debates have centered on two problems, that of the morpho-syntactic status of the article (a (clitic) head in D or a suffix) and that of the syntax of the article, in other words whether the article and the noun or adjective come to be together by merge or move.

As to its status, the article has been variously argued to be either a second position clitic (e.g. Renzi (1993), (1997)) or a suffix (e.g. Halpern (1992)). This dual analysis is a consequence of the fact that its basic distribution (occurrence on the first [+N] constituent of the DP, as in (2) above) is consistent with both analyses. However, the examination of more distributional contexts tilts the balance to the suffix analysis.

Reviewing the evidence for the clitic hypothesis (e.g. Renzi (1993), (1997)), one first finds that the clitic hypothesis can accommodate the basic distribution of the article: assuming that it is projected under D, since it shows up on the first A or N of the DP, it clearly combines with the first [+N] constituent below D of the phrase. Depending on other assumptions, the article occupies its D position throughout syntax and is lowered on the first nominal constituent
following it at PF (as in (7)). Alternatively, the first N/A raise to D and adjoin to the article. Either way, the article will be the second constituent of the phrase, i.e. a second position clitic.

(7)  * [D ul [AP frumos]ul trandafir]]
     [D ul [[NP trandafir]ul] frumos]]

A second argument favouring the clitic interpretation is that, at least in stacking configurations where more than one adjective precedes the noun, the article is realized only once on the highest adjective:

(8)  
a. importanta recentă descoperire
    important.the recent discovery
b. * importanta recenta descoperire
    important.the recent.the discovery

Thirdly, clitic analyses of the definite article have also been proposed for other Balkan languages, as in Dobrovie & Giurgea (2006) for Bulgarian, and since enclitic definite articles represent a Balkan Sprachbund feature, a clitic analysis of the Romanian article was also indirectly supported.

Nevertheless, there are other properties of the definite article which place it in the class of suffixes, rather than clitics, as has been argued for in a number of discussions on this topic (Lombard (1974: 192), Halpern (1992); Ortmann & Popescu (2000); Dobrovie-Sorin & Giurgea (2006)), among others). At the same time, similarity with other Balkan languages is partial, detracting from the plausibility of the clitic hypothesis.

Halpern (1992) notices as a first problem for the clitic analysis the multiple occurrence of the article within the same DP – for instance, when there are coordinate adjectives (10b). Romanian and Bulgarian form a minimal pair on this point, since in Bulgarian the article occurs only once on the highest adjective (9a), while, in Romanian, the article occurs on all coordinate adjectival heads (10b).

(9)  
a. novata i interesna kniga
    new.the and interesting book
b. * novata i interesnata kniga
    new.the and interesting.the book
c. * nova i interesnata kniga
    new and interesting.the book

(10)  
a. * noua şi interesantă carte
    new.the and interesting book
b. noua şi interesanta carte
    new.the and interesting.the book
c. * nouă şi interesanta carte
    new and interesting.the book
Moreover, the article *does not always occupy the second position*. Degree words (11a, b) and other modifiers of an adjective (11c) may intervene between the D position and the adjective to which the article attaches (Dobrovie-Sorin & Giurgea (2006: 75-76)).

(11) a. *atât de lungile drumuri*  
    such of long. the roads  
    ‘the so long roads’

   b. *foarte lungile drumuri*  
    very long. the roads  
    ‘the very long roads’

   c. *realmente plicticoasele dezbateri*  
    really boring. the debates  
    ‘the really boring debates’

Had the article been a clitic, *it would not have been sensitive to the category* of the head or phrase it cliticizes on. From this point of view, the Bulgarian definite article behaves like a clitic attaching to nouns, adjectives, quantifiers or possessives, in other words, to any constituent which occurs in front position in the DP. In particular, the Bulgarian article may occur on cardinals (12a), which are φ-incomplete, contrasting with the Romanian article which does not combine with cardinals (12b).

(12) a. *pete knigi*  
    five. the books  
    (Bulgarian)

   b. *cinci cărți*  
    five. the books  
    (Romanian)

In contrast with the Bulgarian article, the Romanian one shows *selectivity* for a [+N] stem (a noun or an adjective), behaving like an affix, not a clitic. Moreover, adjectives should be φ-complete (13b); when they are not, (i.e. they are invariable, like *ditamai*, ‘big, huge’), the article cannot appear on them (13c), showing up on the following noun instead. (13c’). In conclusion, the definite article shows *high selectivity for its host*, behaving like an affix. (Ortmann & Popescu (2000: 784), Zwicky and Pullum (1983: 504)).

(13) a. *omul*  
    man. the

b. *marele om*  
    big. the man

c. *ditamaiul om*  
    huge. the man  
    c’. *ditamai omul*  
    huge man. the

Ortmann & Popescu’s (2000) main argument for rejecting the clitic analysis is the definite article’s *allomorphy*. It is instructive to contrast the allomorphy of the definite article, presumably a suffix, with the allomorphy of its (etymologically) cognate 3rd person pronominal clitic *l-/îl* ‘him’. The allomorphs of pronominal clitics are *phonologically conditioned* (Ortmann & Popescu
and preserve the integrity of the host. Thus, the form îl appears before consonants, the form l- is proclitic or enclitic on vowels:

(14) a. îl văd
    him (I) see
    ‘I see him’

b. l-am văzut
   him- (I) have. Seen
   ‘I have seen him’

c. nu-l văd
   not-him (I) see
   ‘I do not see him.’

The distribution of the clitic is less constrained than that of the suffixal definite article. The pronominal clitic occurs on any head of the extended projection of the verb (for instance, it cliticizes on the main verb in T (14a), on an auxiliary verb (14b), on negation (14c), on the subjunctive mood particle (să in (15a)), on the complementizer (că in (15b)). This shows that pronominal clitics are independent syntactic heads.

(15) a. să-l văd
    SĂ(subj)-him (I) see
    ‘that I should see him’

b. că-l văd
   that-him (I) see
   ‘that I see him.’

The allomorphy of the article is very different from that of a clitic. In the first place, which particular allomorph is selected depends on the morpho-phonological properties of the stem (according to Ortmann & Popescu (2000)); for instance, masculine nouns ending in a consonant take the basic allomorph –(u)l, as in copil (child) → copil-ul (child.the), while masculine nouns ending in -e take the allomorph -le, as in perete (wall) → peretele (wall.the). Secondly, the article does not preserve the integrity of the host. It may produce changes in the vocalic and syllabic structure of the stem (GBLR (2010: 88-90)). For instance, the suffixation of the article turns semivocalic endings into full vowels leu [leu] → le-ul [leul], and full vowel endings into semivowels floare [floare] → floarea [floarə]. It may also produce stress shift: radio → radioul. These morphologic facts are strong arguments in favour of analysing –(ul) as a suffix, a view also adopted in this paper.

Conclusion and consequences. Researchers now agree on the suffixal nature of the definite article. This result is significant for the syntax of the article, especially when these (mostly) morphologic considerations are correlated with Danon’s (2010) definiteness parameter, discussed below. The article as a suffix hypothesis has also been strengthened by the “discovery” of the low definite article construction in Old Romanian.
2.2. On the syntax of the article. The Noun+ article structure

2.2.1 N-to-D

Within GB, the most widespread analysis of the Romanian definite article, first proposed by Dobrovie-Sorin (1987), and adopted in important studies on Romanian (Giusti (1993), Coene (1999) among many) is the N-to-D analysis. It claims that the article is projected under D, but being an affix, it must be supported by a nominal stem, so that the head N raises and adjoins to D, as shown in (16).

(16) a. copacul
   tree.the

   b. [DP [D -ul] [ NP [N copac]]] (basic structure)

   c. [DP [D [N copac]+ [D –ul]]] [ NP [N copac]] (derived structure)

Adjectives may also move to the D-domain, but they move as phrases (Grosu (1988)), since when the AP is the complement of a Degree head, the Degree head cannot be left behind, but is pied-piped with the adjective.

(17) a. [DP [DegP foarte inteligentul] D [DegP foarte [AP inteligent] [NumP student]]]
   very intelligent.the

   b. * [DP [D inteligent +ul [DegP foarte [AP inteligent] [NumP student]]]]
      intelligent.the very

   student

N-to-D was consistent with a set of assumptions regarding the syntax of the DP, in particular with the hypothesis on how to derive the post-nominal order of adjectives.

In a structure like (18b=18a), under Cinque’s hypothesis of partial N-Movement to Number (N-to-Num), AP1 becomes post-nominal when N targets the Num head. The second AP2 is [+N] and can in principle target Spec, DP, with the result represented in (18c).

(18) a. fostul președinte foarte popular
   former.the president very popular
   ‘the former very popular’


The N-to-D analysis was abandoned not only because of the theoretical difficulties surrounding head movement (Chomsky (1995)), but also because of its inability to cope with certain empirical problems such as the mirror-image order of Romance postnominal adjectives in contrast with their English counterparts (see Shlonsky (2004), Cinque (2010), among money). Here is one example:
2.2.2. The Def-lowering hypothesis

The Affix lowering hypothesis has been formulated by Dobrovie-Sorin & Giurgea (2006) in the general framework of early Minimalism, combined with Distributed Morphology (DM) (Halle & Marantz (1993), Embick and Noyer (2001)) and it presents a very different view of the matter.

Under DM assumptions suffixation is the effect of a postsyntactic PF operation and it need not be uniform across languages. In some languages suffixation could be based on N-to-D raising, whereas in others it could involve Det-lowering. Dobrovie & Giurgea (2006) argue that in Romanian there is no evidence of N-to-D movement, basically because there are no word order contrasts between contexts where N-to-D was claimed to have applied and contexts where there is no N-to-D, since the DP is indefinite. Adjectives retain their subcategorized position irrespective of definiteness. They have the same position with respect to the noun, whether the DP contains a definite article or not. For instance, always prenominal adjectives remain so, irrespective of definiteness and word order is unchanged for other categories of adjectives as well.

(20)  
a. o fostă școală  
a former school  
b. * o școală fostă  
a school former

(21)  
a. fostă studentă  
former.the student  
b. * școală fostă  
school.the former

On the strength of such data, the authors conclude that Ns suffixed with a definite article do not occupy a position that is higher than that occupied by Ns governed by other determiners (indefinite, demonstrative, cardinals, etc.), so there is no evidence in favour of N-to-D in Romanian. A further indication that article suffixation is the result of a morphological rule rather than of a syntactic movement is the strict locality of the relation between the article and its host.

The analysis in fact, the claim against N-to-D is much less radical than it appears, since the two authors accept that there is N-Movement to the Number head (N-to-Num), though not further. Since QPs block the definite article, NumP will be the projection immediately below the D, which hosts the article.

Suffixation of the definite article is analysed as an instance of Affix Hopping. Affix Hopping is a Lowering rule, which has the following properties:
a) It affects heads.
b) It attaches a head to the head of its complement.
c) It is sensitive to syntactic structure, so that it applies at an early stage of PF derivation, prior to Vocabulary Insertion and linearization, where some structural information is still available. Essentially the analysis makes the claims in (22).

\[(22) \quad \text{Def-lowering to Num in Romanian.}\]

(i) Def Lowering Targets the Number head.
(ii) Def lowers to a morpheme marked with full φ-features.
(iii) φ-features attach to Number.

These assumptions are sufficient to derive the basic facts presented regarding definite nouns and definite adjectives in Romanian.

\[(23) \quad \text{Definite nouns}\]

a. The N raises to Num in (syntax) (N to Num). This is an instance of head movement and occurs in narrow syntax, resulting in the LF below:

\[a' \quad \text{LF} [\text{DP}[\text{D}-(u)l] [\text{NumP}[\text{Num}[N]+[\text{Num}]] [\text{NP} N]]] \]

b. At PF, the article targets Num (Def-lowering), since Num contains at constituent that has φ features (the head noun); the following PF is thus derived:

\[b' \quad \text{PF} [\text{DP}[\text{D}-(u)l] [\text{NumP} [\text{Num}[N]+[\text{Num}-(u)l]] [\text{NP} N]]] \]

\[(24) \quad \text{Definite adjectives}\]

a. If an adjective is prenominal, it is the AP rather than the N which raises to Spec, NumP. The adjective values its φ-features against the Number head, in narrow syntax.

b. At PF the article targets the adjective which has a complete φ-set.

Consider now DPs containing quantifiers, for instance cardinals. Adopting Borer (2005), the authors assume that cardinals merge in the intermediate QP position between D and NumP. As such they can host neither φ features nor the definite article. Def lowering is blocked since it is a local rule which can only target the head of the complement phrase, i.e., the cardinal in (25) and the latter is not phi complete. Hence article lowering is impossible and there must be last resort “morphological support” by the strong definite article cel.

\[(25) \quad \text{DP} \quad \text{QP} \quad \text{NumP} \quad \text{NP} \]

\[\text{* -ul} \quad [\text{φcardinal}] \quad [\text{Num} \ φ] \quad \text{NP} \]

\[\text{cel} \quad [\text{φ cardinal}] \quad [\text{Num} \ φ] \quad \text{NP} \]
In addition to its elegance, the Def-lowering analysis also has good explanatory power. It simplifies syntax, placing suffixation where it should, i.e. in the morphological component; it can account for the complementarity of –(u)l/ cel, and for the occurrence of the article on prenominal, not on postnominal adjectives.

Moreover, it can be extended to complex DPs which check definiteness by means of a definite constituent in Spec DP, such as a superlative or an ordinal in Spec, DP; see example in (26a), analysed as in (26b). The definite constituent in SpecDP is said to be “full”, containing its own empty nominal head. In such examples definiteness is valued by the full definite DP in Spec, DP, so the head noun is indefinite, as shown in (26b):

(26) a. \([\text{cea mai frumoasă]} \text{ printeşă}\]
   the more beautiful princess
   ‘the most beautiful princess’

   b. \([\text{DP [DP cea mai frumoasă]} [\text{D [def]}] [\text{NumP [ Num printeşă]]}]\]

Interestingly, if the definite constituent is not in Spec, DP but appears post-nominally in an adjunct position, the definite article shows on the head noun by Def-lowering again, as in (27), which is a double definite construction, showing two definite articles (in bold below, from Dobrovie-Sorin & Giurgea (2006: 95)):

(27) a. \([\text{printeşta} \text{ cea mai frumoasă}\]
   princess.the the more beautiful
   ‘the most beautiful princess’

   b. \([[[\text{D-a}] [[\text{Num printeşă}] [\text{NP [N printešă ] [DP cea mai frumoasă]]}]\]
   (from Dobrovie & Giurgea.

More problematic for the lowering approach are instances of multiple definites, where several copies of the definite article occur in the same DP (e.g. multiple definite in stacked structures), illustrated below:

(28) a. \([\text{săracul} \text{ copilul ăla}\]
   poor.the child.the that
   ‘that poor child’

   b. \([\text{săracul fratele Mariei.}\]
   poor.the
   ‘Mary’s poor brother’

Further research is needed to assess the empirical adequacy of this hypothesis.

3. AN AGREE ACCOUNT, DEFINITENESS AS A MORPHOSYNTACTIC FEATURE ON NOMINAL STEMS

3.1 The definiteness parameter

The analysis we propose here and have also proposed elsewhere (Cornilescu & Nicolae (2011), (2012)) relies on the intuition that in certain languages definiteness has been grammaticalized to the point of becoming a *morpho-syntactic feature*, an agreement category, rather than merely a
morpho-semantic feature with a role in determining reference. Languages where definiteness is syntactic (e.g. Romanian) are expected to present certain properties, say, double definite constructions, properties which are conspicuously absent in languages where definiteness is morpho-semantic (English).

The distinction between morphosyntactic and morphosemantic features concerns the phenomena in which the respective feature can participate, as shown in Kibort (2010: 80-81): “A morphosyntactic feature is a feature whose values are involved in either syntactic agreement or government”, in opposition to morphosemantic features “whose values are not involved in either syntactic agreement or government but are inherent only”. Reinterpreting Kibort in minimalist terms, as in Nicolae (2013:66), a morphosyntactic feature is a feature whose values are involved in c-command and/or Agree.

The fact that definiteness is a morphosyntactic feature in certain languages has first been discussed in connection with Semitic languages, by Hagit Borer (1999) in her analysis of construct state constructions. Her insight is that in certain languages the nominal stem is specified as [+definite] in the numeration, since, in those languages definiteness is a syntactic feature. Her idea was picked up and systematically developed by Danon (2010), who gives empirical evidence for differentiating between morphosyntactic and morphosemantic definiteness and proposes a series of tests which may be used to diagnose syntactic definiteness. Danon thus proposes the existence of a parameter, labelled the definiteness parameter, which states that in certain languages (e.g. Hebrew), N and A may be specified as [+definite] in the lexicon. Furthermore, in the parameter typology proposed by Roberts (2012), the definiteness parameter, which affects only [+N] heads is a nanoparameter (as opposed to mesoparameter (e.g. the head parameter) or a macroparameter, (e.g. the null subject parameter).

While Danon (2010) proposes the definiteness parameter as a UG dimension of variation, there are also other researchers who noticed the different status of definiteness across languages, reaching converging conclusions. In a series of papers comparing Romance and Germanic DPs, Rohers (2006, 2012) expresses the same intuition about the different role of the definite article in Romance and Germanic:

(29) a. Romance: Article insertion is a morphosyntactic process.
    b. Germanic: Article insertion is a semantico-pragmatic process.

We claim that Romanian is a language with morphosyntactic definiteness, a property made manifest with specific tests. Danon (2010) for Hebrew and Nicolae (2013a) for Romanian argue that there are specific phenomena which identify the presence of a morphosyntactic definiteness feature on nominal stems. Among these, the following are directly relevant for this discussion: definiteness agreement and multiple definites. Definiteness agreement simply says that adjectives agree with nouns in definiteness, as apparent in (30). The second phenomenon that of multiple definites regards the presence in the same DP of several definite articles or definite terms. Thus, the DP in (31) shows both of the definite article in the same DP:

(30) * săracul  fratele  Mariei.
    poor.the  brother. the  Mary’s
    ‘Mary’s poor brother’
(31)  * copilul cel fericit
child.the CEL happy
‘the happy child’

3.2 A local agree account of the suffixal definite article

An essential ingredient in the proposed analysis is Pesetsky & Torrego’s (2007) distinction between valuation and interpretability. Definiteness in D is always interpretable and may be valued or unvalued; D is also endowed with uninterpretable unvalued φ-features: D [i def [], uφ []]
Definiteness on nominal stems is always uninterpretable and valued:φ N [+iφ, u+def]. Since definiteness is marked on the nominal stem, it makes sense to allow the definite article to merge low, suffixed to the N, spelling out its [u+def] feature. In a language like English, [i def] is valued by merging the definite article the, while the φ-feature of D are valued via Agree with the c-commanded nominal restriction, as shown in (32).

(32)  * the rose

In Romanian, which has a suffixal definite article, in a simple structure where the DP consists only of a definite noun, the [u+def] feature of the noun, acting as a Goal, will value the [i def] feature of the D Probe via Agree:

(33)  a. elevul
pupil.the
‘the pupil’

b. DP  \rightarrow  c. DP

Let us concentrate on attributive adjectives now (34), assuming Cinque’s (2010) view that they merge as specifiers of functional projections. As mentioned, definiteness is among the
concord features of Romanian. The adjective in (35) is in a c-command position and may undergo Agree, i.e. *feature sharing* in the sense of Pesetsky & Torego (2007). In Romanian, the A and the N share both φ-features and definiteness. After agreement, the adjectival head is duly equipped with features matching those in D, and it is also in an appropriate configuration to serve as a Goal for the D-head, occupying the specifier position immediately under D: the D-head gets its definiteness and φ-features valued and the derivation is convergent.

(34) \[ frumosul \quad trandafir \\
\quad \text{beautiful.rose} \]

\[ \text{‘the beautiful rose’} \]

(35) \[
\text{DP} \quad \text{FP} \\
\quad \text{D} \quad \text{AP} \\
\quad \quad \quad \text{F} \\
\quad \quad \quad \quad \text{NP} \\
\quad \quad \quad \quad \quad \text{frumos} \quad \text{trandafirul} \\
\]

(36) \[
\text{DP} \quad \text{FP} \\
\quad \text{D} \quad \text{AP} \\
\quad \quad \quad \text{F} \\
\quad \quad \quad \quad \text{NP} \\
\quad \quad \quad \quad \quad \text{frumosul} \quad \text{trandafir} \\
\]

At this point it is important to distinguish between *the realisation of the definite article* and *the valuation of definiteness*. While definiteness is phonologically realized only once, on the highest [+N] constituent of the phrase, it is present on all prenominal [+N] heads in the extended projection of the noun, as shown in (36). Evidence for the presence of the definiteness feature on all [+N] heads is given by the existence of (exceptional) multiple realization of the definite article, both in Modern Romanian (30) and in Old Romanian (see below).

It is the highest [+N] constituent, the one immediately below D, which values the [idef] and the φ-features of D. The definiteness first requirement, informally stated above in (5), can now be restated more formally as below:
(37)  * **Definiteness valuation in Modern Romanian (Local Agree)**

The [+def] Goal Phrase which values [i-def] in D must be a [+N] phrase immediately below D.

When an appropriate A/N [+N, u+def] constituent is not found immediately under D, for instance because there is a QP below D, definiteness cannot be valued by a definite suffixed A/N and it will be valued by inserting the free standing article *cel*, operating just as in English. (for details regarding the insertion of *cel* see Cornilescu & Nicolae (2012)).

4. **THE LOW DEFINITE ARTICLE, A SECOND PATTERN OF DEFINITENESS VALUATION IN OLD ROMANIAN**

4.1 The basic configuration

Old Romanian (OR) (i.e. the period of the XVI\textsuperscript{th} to (end of) the XVIII\textsuperscript{th} century, according to Gheție (1975)) also allows Long Distance Agree, alongside of Local Agree. This means that the article, as well as other constituents which incorporate definiteness, does not necessarily occupy the first position in the DP, against the basic generalization in (5) above. Definiteness is valued *across an intervener*, in the following configuration:

(38)  * D[\text{[i-def:____]}] (XP) NP[u+def:val]

This variation in the distribution of the article had not been noticed in important histories of Romanian (Gheție (1975), Rosetti (1968), Densușianu (1961), Dimitrescu (1978), Brancuș (2004), Niculescu (1990)). The low article, illustrated in (39), shows that locality conditions on definiteness checking were more relaxed in OR, ruling in configuration (38). Definiteness valuation merely required Long Distance Agree, possibly formulated as in (40).

(39)  * **ca mare scrâşnetul roatelor** (Cantemir)

like great grinding.the of the wheels

‘like the strong grinding of the wheels’

(40)  * **Definiteness valuation in Old Romanian – Long Distance Agree (LDA)**

The goal that values the probe in D is a [+N] phrase (NP, AP) c-commanded by D, a phrase which need not be the first (nominal) phrase c-commanded by D.

This amounts to saying that the [+def] feature is realized either on the first or on a lower nominal constituent of the DP, that is, what we called “low definite article”.

4.2. The extension and range of the Low Definite Article

The low definite article is present from the earliest OR texts of the XVI\textsuperscript{th} century until the first half of the XVIII\textsuperscript{th} century, and it is uniformly distributed in the Romanian dialects.

(41) a.  *....au venit egumenul de Bistriță cu cinstăță cartea mării tale* (XVI\textsuperscript{th} – DÎR)

has come abbot.the of Bistriță with honoured letter.the highness.the Gen your

‘...the Abbot of Bistrița came with your highness’ honoured letter’
An interesting question is that of the range of the XP constituents that may intervene between the D position and the noun suffixed by the definite article. The following types of constituents have been identified: quantifiers, adjectives, pre-nominal genitives. From the point of view of MR, the most striking fact is the possibility of valuing definiteness across a QP, a fact which suggests that the specialization of cel as a second definite article in complementary distribution with –ul was required precisely by the stronger locality conditions on definiteness valuation. The examples below show cardinals (42a), ordinals (42b) and distributive quantifiers (42c) followed by the definite noun.

(42) Quantifiers

a. deade Dumnezeu [DP zeace cuvintele sale] gave God ten words.DEF his
   ‘God gave his ten commandments’ (CCat.1560: 4’t)

b. arătarea [a dooa venireei lui] showing.DEF the-second coming.DEF.GEN his
   ‘the showing of his second coming’ (CC2.1581: 536/16)

c. înă [câte trei morţii] nu sunt încipuiiri
   but each three deceased.DEF not are illusions
   ‘but the three deceased each are not illusions’ (AD.1722-1725:131’).

Consider now adjectives and pre-nominal genitives preceding the definite N. Of these, only the adjectival pattern has survived in certain circumstances (e.g. see below).

(43) * (Non-definite) adjective

Aşa, fără veste el în vrăjmaşi colţii crocodilului aflându-se (Cantemir)
thus suddenly he in inimical teeth the crocodile.DEF.GEN being
‘Thus, suddenly, he was in the crocodile’s inimical teeth’
In MR a pre-nominal Gen is followed by a non-definite N. Just as in English, a DP containing a pre-nominal Gen is interpreted as [+def], and it is the pre-nominal Gen DP which checks the [\text{\textit{idf}}] feature of the DP. The pre-nominal Gen DP in MR functions as a definite Determiner. It also occurs in \textit{the first position of the DP}, presumably in \text{[Spec, DP]}, since it incorporates the definite article (i.e. –\textit{l} in the genitival article \textit{al}).

Examples like (44) contrast with MR (45a), and are no longer found in MR. The low definite article raises several questions: 1. What is the interpretation of this phenomenon in the framework sketched above? 2. What are the contexts that favored the occurrence of the lower article? 3. What are the causes that led to the elimination of this pattern? 5. Are there other OR DP structures which relate to the existence of LDA? In the remaining part of the paper we supply tentative answers to some of these questions.

4.3. Interpreting the facts of Old Romanian in the Agree framework

Two morpho-syntactic properties of Romanian N/A have combined to produce the strict locality conditions on definiteness valuation in MR. The first is the suffixal nature of the article which allows Ns to be valued for definiteness, even if their definiteness feature is uninterpretable (i.e. [\text{\textit{u+def}}]).

The second significant property is that, at some point in the evolution of Romanian, as must have acquired the possibility of \textit{optionally incorporating an uninterpretable unvalued definiteness} feature [\text{\textit{undef} [ ] }]. This feature was valued through Agree with the N, as shown above. Since Agree was/ could be \textit{long distance}, and the nominal valued for definiteness did not need to be the first NP/AP below D, we expect the following alternations in OR, all of which are attested:

(i) The definite article shows on \textit{the first NP or AP of the DP} (the MR pattern, available in all attested stages of Romanian)

(46) * \text{au purces fără numai din vechea\ şi rânceda pizmăluire} (Cantemir)

(it) happened only out of old.the and rancid.the envy

‘It all happened out of the old and rancid envy.’
(ii) The definite article shows on an NP which is not the first phrase of the DP (this is the low article)

(47) * spre ticăloase cuvintele mele audzul îți pleacă… (Cantemir)
to vicious words.the my hearing your turn
‘Lend your ear to my vicious words.’

(iii) Sporadically, the definite article could also be present on more than one constituent, i.e. the DP shows multiple definites (see Croitor (2008)). This is consistent with the view that definiteness had become a concord feature in OR, and Romanian morphology allows it to be uninterpretable but valued and realized on both Ns and As. Multiple definites are present both in the order A+N (examples (48), from Croitor (2008)), and, highly infrequently, in the order N+A (examples (49), likewise from Croitor (2008)). Apparently, multiple definites were lost (in these patterns) at the end of the XVIIIth century:

(48)  
**Multiple definites: A+N**

a. Ce i-au tăiat atuncea curând puternica mâna lui Dumnedzeu zilele
that to-him have cut then soon mighty.the hand.the of God days.the
‘God’s’ almighty hand took his/her days’ (Costin)

b. …moaștele a sfintei prepodobei Paraschevei, în (Costin)
…relics.the of saint.the beautifully-adorned.the Parascheva, in…
‘…the relics of the holy, beautifully-adorned (Saint) Parascheva’

(49)  
**Multiple definites: N+A**

zidul cetății mari și frumoase în Spania (Cantacuzino)
wall.the city.the.Gen big.the.Gen and beautiful.the Gen in Spain
‘the wall of the big and beautiful city in Spain’.

It is reasonable to hypothesize that it was precisely the possibility of valuing definiteness on pre-nominal As that led to stricter locality conditions on Agree, that is, to the MR requirement that the constituent that values the [i\text{def}] feature of D should be the first AP/NP below D.

Gradually, as a result of a general tendency towards economy, the constituent that values [i\text{def}] in D (i.e. which has an LF effect) got to be the only one which phonologically realizes definiteness. This was either a definite pre-nominal A or a definite N. Multiple occurrences of the suffixal article are generally ruled out now.

Moreover, only pre-nominal As are ever suffixed by the definite article, since only pre-nominal As can be closer to D than the N. Definiteness thus turns into an exclusive property of attributive As as opposed to predicative ones. This means that As that merge as attributes (specifiers or pre-nominal adjuncts) are obligatorily specified as [u\text{def}]. This is what guarantees that if the head N is [u+\text{def}], and there is a pre-nominal A in the DP, definiteness will be realized on the pre-nominal adjective immediately below D.

4.4. Contexts of occurrence of the low definite article

One of the relevant questions formulated above is what contexts favor the use of the low article. To answer this question, a body of texts ranging from the earliest Romanian writings of the XVIth
century to the first half of the XVIIIth century was examined. The texts show variation between
the ‘high’ article, used in most cases, and the ‘low’ article. The low article predominantly occurs
when the definite head N is followed by another constituent, especially by a Gen(itive). Consequently, we have checked the relative frequency of the high vs. lower article (see table (50)
below) only for DPs which are modified or complemented by a Gen.

The examination of the data shows that a low definite article on N appears overwhelmingly
(87.5%) in contexts where the head N is followed by a Gen DP complement. Other post-
nominal modifiers may also sometimes trigger the presence of the lower article as in (56b), but
the incidence is low. If there is no complement or modifier, the article emerges on the pre-
nominal A, just as in MR.

(50)

<table>
<thead>
<tr>
<th>Text</th>
<th>A_DEF + N + Gen DP (high article)</th>
<th>A + N_DEF + Gen DP (low article)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miron Costin</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Letopiseţul Cantacuzinesc</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>R. Popescu</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>R. Greceanu</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Constantin Cantacuzino</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Cantemir</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>TOTAL</td>
<td>6</td>
<td>42</td>
</tr>
</tbody>
</table>

| High article: 6 (12.5%) | Low article: 47 (87.5%) |

(51) * A [+def] + N [-def] (no complement/modifier)

au purces fără numai din vechea și râncedă pizmăliure (Cantemir)

(it) happened only out of old. the and rancid. the envy

‘It all happened out the old and rancid envy’

(52) A [- def] + N [def] + GenDP

a. ...ca mare scrîşnetul roatelor (Cantemir)

like great grinding. the of the wheels

‘...like the strong grinding of the wheels...’

b. ...Corbul de uscate vinele goalelor ciolane clonţul și-ar ciocni (id.)
the Raven against dried-up veins. the empty. the Gen. bones. bill. his would knock

‘The Raven might knock his bill against the dried-up veins of his bones.’

c. ...de dulce otrava Hulpii tare se ameţiră (Cantemir)

...with sweet poison. the. of the Vixen much (they) got drunk

‘...they got quite drunk from the sweet poison of the Vixen’

(53) A [- def] + N [def] + Modifier (PP or AP modifier)

a. înşă nu puţine asupreluc despre vrâjitorii vremii trasă (Cantemir)

but not few injustices. the from the magicians of the times (he) endured.

‘...but he endured many injustices from the magicians of the times...’
b. *Neştiutor* omenesc gândul [...] la ce merge? (Cantemir)

Ignorant human thought. what is heading for?

‘What is the ignorant thought of man aiming at?’

This distribution signals a tight relation between the inflectional Gen (and other modifiers) and the lower article, a fact which should be accounted for.

### 4.5. Remnant low article

In MR, i.e. after 1780, the low article is not found any more. There are, however, two types of motivated exceptions.

(a) The lower article is still part of religious and other obsolete texts (54), and in such case it is again mostly followed by the Gen.

(b) Secondly, there is a small group of quantificational or evaluative As that may function as *definite quantifiers*, and may or must be followed by definite Ns (see Barbu (2004)): *întreg* (‘whole’), *singur* (‘unique’) *ditamai* (‘big, huge’), *gogeamite* (‘big, huge’), as in (55). No Gen modifier is required. Loss of the article on the prenominal A precisely shows a change of status, from adjective to determiner. Some of these As are invariable (ϕ-incomplete) and cannot be suffixed by the article, as shown above (55b).

(54)

*Miluieşte-mă, Doamne, după mare milă Ta!*

Have-mercy-on-me God, according to great mercy. the Your ‘God, have mercy on me, according to your great mercy.’

(55) a. *întreg orasul / întregul oras*

whole city.the / whole.the city

b. *ditamai prostul / *ditamaial prost*

big fool.the / big.the fool.

**Conclusion** one significant change in the syntax of the Romanian DP relates to the locality conditions of definiteness valuation. *The definite feature strengthens requiring to be valued by a strictly local nominal constituent (N or A). Long Distance Agree is ruled out.*

### 4.5. Accounting for the statistical correlation between the lower definite article and the inflectional Genitive

As shown in (50), in OR there is a strong preference for using the low article if the N is followed by an inflectional Gen. In our interpretation of the data, the correlation between the low article and the postposed inflectional Gen follows *from distributional facts combined with economy considerations*.

It is known that the inflectional Gen in Romanian is realized either by a bare inflected DP (= the *bare Genitive* (56b)), or as an inflected DP preceded by the *genitival article AL* (56a) (for a description of the Gen article see Cornilescu (1995), (2003)). The two forms are in complementary distribution.

Roughly, the AL Gen occurs whenever the head N is indefinite (56a), while the bare Gen occurs when the head N is definite *and* the Gen is *adjacent* to the head N (56b):
Let us focus on definite nominal heads and examine the use of AL-Gen with such heads. With *definite* heads, the AL Gen occurs in two situations:
(i) when there is an intervening modifier between the definite head and the Gen as in (57a);
(ii) when the Gen is adjacent to the head, but *the article is on a pre-nominal A*, as in (57b).

(57) a.  
\[
\text{prietenul bun al copilului} \\
\text{friend.the good child.theGen} \\
\text{‘the child’s good friend’}
\]
b.  
\[
\text{bunul prieten al copilului} \\
\text{good.the friend ALgenitival article child.theGen} \\
\text{‘the child’s good friend’}
\]
c.  
\[
\text{*bunul prieten copilului} \\
\text{good.the friend child.theGen.}
\]

Generalizing, the bare Gen occurs only when *it immediately follows a suffixed definite N*, while the AL Gen occurs *elsewhere*. The bare Gen is the strongly preferred form since it is more economical (economy of representation). It is likely that the preference for the bare Gen must have been felt in OR as well. This is what explains the use of the lower definite article with inflectional Gen, creating the proper context for the use of the bare Gen.

If, in contexts where there is a prenominal A, the article is placed on the N, instead of being placed on the higher A, the Gen is adjacent to the definite article and it is possible to employ the more economical bare Gen (68a), replacing the AL Gen used in MR (58b):

(58) a.  
\[
\text{OR: cu cinstită carte a mării tale} \quad \text{(Bare Gen)} \\
\text{with honored letter.the highness.theGen your}
\]
b.  
\[
\text{MR: cu cinstita carte a mării tale} \quad \text{(AL genitive)} \\
\text{with honoured.the letter ALgenitival article highness.theGen your.}
\]

The same preference for the bare genitive is responsible for some of the multiple definites that survive in MR:

(59) a.  
\[
\text{săracul fiul tău} \quad \text{poor.the son.the yours} \\
\text{‘your poor son’}
\]
b.  
\[
\text{??? săracul fiu al tau} \quad \text{poor.the son AL you.Gen} \\
\text{‘your poor son’}.
\]

4.5. What the Lower Article suggests about the emergence of the enclitic article
This present analysis relies on the contrast between languages which value definiteness in D by
merging a determiner and languages which value D by means of a *lexical category*
morphologically *marked as valued for definiteness*. From a Romance comparative diachronic
perspective, a famous puzzle is how the same Latin demonstrative *ille* (cf. Iordan & Manoliu
(1965)) led to *proclitic free-standing articles* in other Romance languages, but to an *enclitic
article* in Romanian.

The lower article tilts the balance to the hypothesis that the enclitic article developed out of
the *post-nominal* Latin Dem *ille* rather than out of a prenominal demonstrative, as in Giusti’s
(1998) analysis of the history of the definite article in Romance. Both proposals have been
advanced for Romanian as well (see Renzi (1993) for the prenominal demonstrative hypothesis,
and Graur (1967) for the postnominal demonstrative hypothesis).

As already stated, the low article suggests that the enclitic article emerged through the
reanalysis of the demonstrative *ille* in the context of the *postnominal demonstrative construction*;
reanalysis amounted to a change in its c-selection properties. This hypothesis is supported by the
existence of postnominal Demonstratives in all written phases of Romanian; a second essential
consideration is that the postnominal position of adjectives, including demonstrative adjectives,
which was possible in Latin, was reinforced by the contact with the local Dacian idioms, which
strongly preferred to postpose adjectives, including the (adjectival) demonstrative (Brâncuș
(2004), Graur (1967)). (For a possible path from the postnominal demonstrative to the suffixed
article in OR see Cornilescu & Nicolae (2011)).

The low article construction matters in this on-going debate on how the suffixed article was
born, since it is the so far only attested construction that *does not show the Romanian definite
article on the head of the first NP/AP of the DP*. As emphasized above, the standard “high
article” is consistent with analyses where the article merges in D and is post-syntactically
lowered on the first [+N] constituent, or where the article merges in D and there is movement into
the D-area (N-to-D, AP-to-Spec, DP).

The main analytical bite of the low article is that the lower article may be construed as *good
evidence that the article merges low*, and “moves” higher through Agree reaching the position
below D. Some of the proposed analyses have difficulties with long distance agree. For instance,
the post syntactic Affix Hopping rule, which is supposed to *observe locality and attach an affix
to the head of the next complement* is apparently incompatible with long distance agree, given the
locality conditions on morphologic rules.

The Agree account, coupled with the view that, being a spell-out of a property of the
nominal stem, i.e. [u+def], the article merges low, can naturally accommodate both the OR and
the MR facts. The passage from OR to MR, though spectacular when one examines the structure
of the DPs themselves, can simply be described as a *strengthening of the locality constraints on
the valuation of definiteness*.

5. CONCLUSIONS

(1) In OR, the definite article suffixed to the noun / adjective may occupy the first position of the
DP, but also a lower position. In particular, indefinite constituents such as indefinite adjectives or
indefinite Gens may precede the definite noun. OR alternates Local and Long Distance Agree in
the valuation of definiteness. Distributionally, the lower article is conditioned mostly by a post-
nominal Gen or modifiers.
(2) In OR the definite noun could even be preceded by a QP, i.e. the definite article valued definiteness in D across a QP: D[idef]> QP> NP[u+def]. The disappearance of the low article forced the specialization of demonstrative (a)cel as a second definite article of Romanian.

(3) At some point in the evolution of Romanian, definiteness became a concord feature for adjectives; the latter optionally entered the derivation with uninterpretable unvalued definite feature, valued by Agree with the noun. It is reasonable to assume that it was precisely the possibility of valuing definiteness on pre-nominal adjectives that ultimately led to the Modern Romanian requirement that the constituent that values the [idef] feature of D should be the first AP/NP below D.

(4) In OR, Long Distance Agree apparently co-exists with Local Agree, a factor that may cause ambiguity. Consequently, Romanian settles for Local Agree, selecting the more restrictive grammar.

(5) While in Old Romanian there is Long Distance Agree and Modern Romanian loses this option. Changes in this parameters lead to the disappearance of a number of DP structures involving Long Distance Agree. Thus, all patterns involving valuation of definiteness across another constituent (a genitive, an adjective) are lost.

(6) The Agree account apparently has the upper hand over rival accounts in accommodating the facts of OR.

REFERENCES


CORPUS

Biblia – Biblia sau dumnezeiasca scriptură (1688), București.
Cantacuzino C. „Istoria Țării Română” in Cronicarii Munteni.
Cronicarii Munteni (1964), Al. Piru (ed), București, Editura Tineretului.
Greceanu, R. „Cronică” in Cronicarii Munteni.
„Letopisul Cantacuzinesc”, in Cronicarii Munteni.
Popescu, R. „Letopisul Bălenilor” in Cronicarii Munteni.

* This work was generously supported by CNCSIS- UEFISCUS, PN II, IDEI 1979/2008.