
EMPHATIC ELEMENTS AND THE DEVELOPMENT OF DEFINITE ARTICLES: EVIDENCE FOR A LAYERED DP IN EARLY ROMANCE*

JUDY B. BERNSTEIN
WILLIAM PATERSON UNIVERSITY
FRANCISCO ORDÓÑEZ
STONY BROOK UNIVERSITY
FRANCESC ROCA
UNIVERSITAT DE GIRONA

ABSTRACT This paper examines the evolution of double article systems in two Romance varieties, Balearic Catalan and Picard. In these languages the two definite articles have different syntactic distributions. We account for these double article systems with a multi-layered DP analysis that maps semantically onto Löbner's (1985, 2011) Uniqueness Scale, further developed in Ortmann (2014). We trace the origins of the present-day definite articles to the Latin demonstrative *ille* (Catalan, Picard), emphatic pronoun and adnominal *ipse* (Catalan), and emphatic interjection particle *ecce* (Picard). The development of the present-day definite articles involved a stage where an emphatic form competed with the demonstrative. In the majority of the present-day Romance languages, a single definite article emerged from the competition that was already underway in Latin. In the Romance varieties considered here, two different forms emerged, but they took on different semantic interpretations. We show how the relevant semantic properties of the three Latin forms are discernible in the distribution of the definite articles that developed from them and that still exist in the present-day varieties.

* Many thanks to three anonymous reviewers for thoughtful comments and questions. We also thank audiences at these conferences: Diachronic Generative Syntax (DiGS) 21 in 2019, Antisymmetry and Comparative Syntax: Celebrating 25 Years of Antisymmetry and 75 Years of Richard S. Kayne in 2019, Linguistic Society of America (LSA) in 2020, and Linguistic Symposium on Romance Languages (LSRL) 50 in 2020. Thanks to Chris Collins, Ivy Sichel, Julie Auger, and Adam Ledgeway for discussion and helpful comments. Thanks to Madhuri Mukherjee for help with the French translations, Barbara Vance for help with the Old Picard translations and Julie Auger for help with the present-day Picard translations. All errors are our own. Parts of this research have been funded by the following research grants to the third author: SGR 1194 (Generalitat de Catalunya), FFI2017-87140-C4-2-P (MINECO); and an Advanced Research Collaborative fellowship of the Graduate Center (CUNY) to the second author.

©2021 Bernstein, Ordóñez & Roca

This is an open-access article distributed under the terms of a Creative Commons License (creativecommons.org/licenses/by/4.0).

1 INTRODUCTION

The definite articles in modern Romance languages like Spanish or French developed from the Latin distal demonstrative *ille* and display *l-* forms. This is a well-known pattern of development of definite articles in many languages (van Gelderen 2011: 219–224):

- (1) *ille* > *l-*
 (pronoun) (definite article)
- (2) (a) *el niño* (Spanish)
 the.M boy
 (b) *le garçon* (French)
 the.M boy

Perhaps less well known, Romance definite articles also developed from the Latin emphatic pronoun and nominal modifier *ipse* in Romance varieties such as Sardinian (Aebischer 1948, Vincent 1997, Ledgeway 2012). This language displays *s-* forms (example from Jones 1993: 34, his (3)):

- (3) *ipse* > *ipse* > *s-*
 (emphatic pronoun) (nominal modifier) (definite article)
- (4) *su babbu de Maria* (Sardinian)
 the father of Maria

These two sources for definite articles resulted in a double article system in Balearic Catalan, where both *ipse*-derived articles and *ille*-derived articles are found.¹ Interestingly, some varieties of Picard also display a double article system: one article derives from *ille*, like Balearic Catalan, but the other one derives from a different element, the Latin emphatic interjection *ecce*.²

In this article we study the historical evolution of Romance definite articles that descended from Latin *ipse*, *ille*, and *ecce* in double article systems such as Balearic Catalan and Picard. Building on work on contemporary Romance varieties (Bernstein, Ordóñez & Roca 2018), we claim that the double system of articles in Balearic Catalan and Picard corresponds to a multi-layered DP structure that started to develop in the transition from Latin to early Romance.

1 We discuss the development of a third type, the so-called personal articles, in Bernstein, Ordóñez & Roca (2019).

2 Vincent (1999), for example, discusses the emphatic nature of both *ipse* and *ecce*.

In this proposed structure, the more emphatic elements grammaticalized into forms occupying a higher layer (DP1) and the demonstrative element, lacking emphasis, was later generalized in a lower layer (DP2):

- (5) [DP1 (*ip*)*se*, (*ec*)*ce* [DP2 (*il*)*le* [NP]]]

The diachronic evolution of the forms includes a stage where the emerging articles competed with each other (Kroch 1989, 1994), ultimately yielding the specialized functions found in the present-day languages.

The article is organized as follows: in Section 2 we describe the synchronic data from two Romance varieties that display a double article system and adopt a multi-layered DP analysis; in Section 3, building on Ortmann (2014), we present a semantic mapping of the two different definite articles onto the DP structure. The next three sections of the article are devoted to the diachronic evolution of Latin *ille*, *ipse*, and *ecce* and the process of grammaticalization to definite articles in Late and Medieval Latin (Section 4), and their generalization in Catalan (Section 5), and Picard (Section 6). In Section 7 we summarize our analysis and provide some concluding remarks.

2 TWO DEFINITE ARTICLES IN BALEARIC CATALAN AND FRENCH PICARD

Before turning to the diachronic data, we provide an overview of the patterns in Balearic Catalan and Picard. Balearic Catalan and Picard are two Romance languages that display two forms of the definite article contemporaneously. Balearic Catalan has an *s*- article and an *l*- article. The *s*- article derives from the Latin emphatic pronoun *ipse* (see (6)), and the *l*- article derives from the Latin distal demonstrative *ille* (see (7)).

- (6) *Sa muntanya que veus és molt alta* (Balearic Catalan)
 the.F mountain that see.2PSG is very tall
 ‘The mountain that you see is very tall.’
- (7) *Anem a la muntanya*
 go.1PPL to the.F mountain
 ‘Let’s go to the mountain.’

Similarly, the relevant varieties of Picard also have two articles: a *ch*- form and an *l*- form.³ The *ch*- article derives from the Latin emphatic interjection *ecce*.

³ As far as we know, the *ch*- forms are not found in Picard varieties of Belgium. In France, the relevant varieties are concentrated in the western and southern areas of Picardie (Panwitz 2018).

The *l-* article, as in Balearic Catalan, derives from the Latin demonstrative pronoun *ille* (Boulogne Picard from Haigneré 1901: 264):

- (8) *chele école* (Boulogne Picard)
'the school (building)'
- (9) *l'école*
'instruction'

The distribution of these two forms of the definite article depends on different syntactic and semantic conditions (Ledgeway 2012, Bernstein et al. 2018), which are parallel in both languages.⁴ Both the *s-* article in Balearic Catalan and the *ch-* article in Picard “have truly deictic force and are able to identify both anaphorically and cataphorically definite referents” (Ledgeway 2012: 100 and references cited there). These articles appear in the contexts of common nouns (see (10) and (11)), relative clauses (see (12) and (13)), adjectives (see (14) and (15)), prepositional phrases (see (16) and (17)) and superlatives (see (18) and (19)) (Boulogne Picard from Haigneré 1901: 267, 286, 378; Amiens Picard from Debrie 1983: 79).⁵

Common nouns:

- (10) *es ca, sa casa* (Balearic Catalan)
the.M dog the.F house
- (11) *che fu, chele porte* (Boulogne Picard)
the.M fire the.F door

Relative clauses:

- (12) *sa dona que mira sa porta* (Balearic Catalan)
the.F woman that looks the.F door
'the woman that looks at the door'
[*Arbre que mira farola*, song by J. M. Oliver]

⁴ The distribution of the forms is subject to “idiosyncratic lexicalization” (Ledgeway 2012: 103, writing about Balearic Catalan) as well as to variation (Panwitz 2018, writing about Vimeu Picard).

⁵ In Boulogne Picard (Haigneré 1901: 264), the masculine form is *che* (from Latin *ecce*) and the feminine form is *chele* (reflecting Latin *ecce+ille*) and the plural (both genders) is *ches*. *Chel(e)* is also the form that precedes vowel-initial nouns. We do not know whether the absence of *-l* with the masculine and plural forms is explained by phonology, but these patterns recall parallel French ones.

Emphatic elements and the development of definite articles

- (13) *Eche nocheau k't'o vu il é t abiyé in maske.* (Amiens Picard)
'The reveler that you have seen (he) is disguised.'

Adjectives:

- (14) *ses màquines perfectes Macintosh* (Balearic Catalan)
the.F.PL machines perfect.F.PL Macintosh
'the perfect Macintosh computers'
[*Astronauta rimador*, song by Antònia Font]
- (15) *ech pauvre garcon* (Boulogne Picard)
'the poor (= pitiable) boy'

Prepositional phrases:

- (16) *es músics de jazz* (Balearic Catalan)
the.M.PL musicians.M of jazz
'jazz musicians'
[*Patxanga*, song by Antònia Font]
- (17) *... parmi chés wardeux d'vaques* (Boulogne Picard)
among the herders of cows
'among the cow herders'

Superlatives:

- (18) *sa dona més feliç* (Balearic Catalan)
the.F woman most happy
'the happiest woman'
- (19) *ch' pus riche* (Boulogne Picard)
the most rich
'the richest'

In both Romance varieties the *l-* article is used with "intrinsically referential entities such as countries and proper names ..." and "... in conjunction with unique, abstract, and generic referents ..." (Ledgeway 2012: 103).^{6,7} As

⁶ The DPs with a unique referent are generally not modified and do not take complements.

⁷ In some examples the unique referents might use the *s-* article in Balearic Catalan. For instance, in a context in which there is a unique hospital and a unique church, speakers use the *l-* article for 'church' (*església*) but not for 'hospital' (*hospital*): *l'església*, but *s'hospital*. This pattern suggests that the *s-* article is extending to what we label DP2.

pointed out above, the *l-* articles in both varieties descended from Latin *ille*. Below we illustrate the use of the *l-* article in Balearic Catalan and Picard with proper names (see (20) and (21)) and generic/unique referents (see (22) and (23); (21 b) from PICARTEXT corpus).

Proper names:

- (20) (a) *L'Havana, les Índies* (Majorcan Balearic Catalan)
 the-Havana the.F.PL Indies.PL
 (b) *el senyor Vidal, el rei*
 the.M mister Vidal the.M king
Lluís XIV, el bisbe Miralles
 Louis XIV the.M bishop Miralles
 'Mr. Vidal', 'king Louis XIV', 'the bishop Miralles'
- (21) (a) *l'Italie, le Mexique, la Belgique* (Boulogne Picard)
 the-Italy the.M Mexico the.F Belgium
 (b) *D'abord, édpi qu'és première feume,* (Abbeville Picard)
la Reine Astrid a s'avouot tuèe in auto ...
 'First, since his first wife, Queen Astrid was killed in an auto
 accident ...'

Generic/Unique referents:

- (22) (a) *el cel, la terra, la muntanya* (Majorcan Balearic Catalan)
 'the sky', 'the earth', 'the mountain'
 (b) *l'hospital, la parròquia, el rei*
 'the hospital', 'the parish', 'the king'
- (23) (a) *el roi d' Suède* (Boulogne Picard)
 'the King of Sweden'
 (b) *la mer, le pape, le lunne*
 'the sea', 'the Pope', 'the moon'

In both languages, we can find minimal pairs in which the use of one article or the other indicates a difference in meaning like *la Terra* 'earth' vs. *sa terra* 'land' in Majorcan Balearic Catalan or *che monde* 'earth' vs. *le monde* 'entirety of the world' in Boulogne Picard (Majorcan Balearic Catalan from [Grimalt 2009](#): 63 and [Ledgeway 2012](#): 102; Picard from [Haigneré 1901](#): 264).

To account for these patterns, [Bernstein et al. \(2018\)](#) propose a mapping from each definite article to a separate D head in a multi-layered DP structure

(cf. [Zamparelli 1995](#)). Specifically, they claim that the *s-* (Balearic Catalan) and *ch-* (Picard) definite articles correspond to a higher DP1 and the *l-* articles to a lower DP2:

- (24) (a) [DP1 *s-* [DP2 *l-* [NP]]] (Balearic Catalan)
 (b) [DP1 *ch-* [DP2 *l-* [NP]]] (Picard)

[Bernstein et al. \(2018\)](#) provide some evidence for this structure based on the existence of nominals in which there is co-occurrence of two different articles in one nominal expression as in (25) and (26). In these contexts, the higher DP corresponds to the *s-* or *ch-* article, while the lower one corresponds to the *l-* article.⁸

- (25) *es l'amo, es l'avi* (s- + l-) (Balearic Catalan)
 'the mister', 'the grandfather'
 (26) *ch' l'esprit, ch' l'infant* (ch- + l-) (Boulogne Picard)
 'the spirit', 'the child'

These examples convey a meaning (anaphoricity, etc.) linked to the higher definite article located in DP1. Some evidence for this assumed structure comes from the fact that in Balearic Catalan *es l'amo* and *es l'avi* are excluded from vocatives, a type of expression that might not correspond to higher DPs (see [Bernstein et al. 2018](#)). [Bernstein et al. \(2019\)](#) provide further evidence that the two D heads can be present in the same DP in Catalan. In (27) the higher D corresponds to the anaphoric value, expressed via the definite article *es*, and the lower D corresponds to the personal article, expressed via *en*.

- (27) *Es conco en Toni* (Balearic Catalan)
 the.M uncle ART Toni
 'Uncle Toni'

⁸ We note that the lower form cooccurs with the upper one only in environments preceding vowels: *amo, avi* (Catalan); *esprit, infant* (Picard). Julie Auger (p.c.) suggests that the orthography in (26) may reflect phonology, namely, the syllabification of *l-* (from *chele*) with the following word. According to [Haigneré \(1901: 265\)](#), however, *ch-* plus *l-* lacks an intervening vowel in these contexts and the two are syllabified with the following noun (i.e., [ʃl-]). [Corblet \(1851: 99\)](#) considers the following example to simultaneously contain an article (*l-*) and the demonstrative pronoun (*ch-*), what we take to be the *ch-* definite article: *ch'l'ékielle* ('the ladder'), from *Après c's armanos lo on put tirer ch'l'ékielle* ('After the almanacs one can pull the ladder', meaning that after the almanacs it couldn't be bettered; thanks to Julie Auger for the translation). It is also important to highlight the fact that the source of *l-* in all these examples is indeed *ille*, the demonstrative that eventually yields the *l-* definite article.

Amiens Picard also displays an interesting set of patterns, one that seems to further support the layered DP structure we have adopted. To start, Amiens Picard displays examples parallel to (26), though the masculine *ch-* article has a different form, *èche* (examples from [Debrie 1983: 66](#)):

- (28) (a) *èche l'achuchonaje* (Amiens Picard)
 'the concubinage' (practice of keeping a concubine)
 (b) *èche l'intèrmin*
 'the funeral'

In these examples, unlike the ones in (26), prenominal *l-* cannot be related phonologically to *èche*, since *èche* has no allomorph displaying *-l* (compare Boulogne Picard *chele*). The feminine article in Amiens Picard is *èle* and for vowel-initial masculine nouns, the pattern is as in (28) above ([Debrie 1983: 16](#)). [Debrie \(1983: 55, 67\)](#) furthermore states that for examples like (28), *èche* is sometimes replaced by *èle*:

- (29) (a) *Ele l'éfan lo ward ché vake* (Amiens Picard)
 the child CL watches the cows
 'the child watches the cows'
 (b) *èle l'èstrémizasyon*
 'the last-rites'

Seen through the lens of a layered DP analysis, examples (28) and (29) suggest a simultaneous spelling out of DP1 and DP2, equivalent to what we saw above for Balearic Catalan. Though explicit in Amiens Picard and Balearic Catalan because two instances of definite articles are expressed, there is syntactic evidence for this layered structure in Sardinian (*s-* forms only) and French (*l-* forms only) as well (see [Bernstein et al. 2018](#)).⁹

Our examination of two Romance varieties displaying two different definite articles (different in both form and use) motivated a layered DP analysis: DP1, higher in the structure, is associated with anaphoricity; DP2, lower in the structure, is associated with genericity/uniqueness. Our approach shares some similarities with [Wiltschko \(2013\)](#). In her study of relative clauses in Bavarian German, Wiltschko finds that this variety has two different forms of the definite article with semantic differences akin to those we discuss here for Balearic Catalan and Picard. The two analyses share the idea that determiners appearing with uniques/generics are closer to NP than contextualized

⁹ [Kayne & Pollock \(2010\)](#) analyze French *ce* as a definite article.

anaphoric determiners. However, the approaches also differ. Our analysis allows for both DPs to be realized and the examples discussed above (see also Footnote 8) provide compelling evidence that this is the case. It is unclear to us whether Wiltschko's analysis allows for such cooccurrences of D.¹⁰

3 SEMANTIC MAPPING OF DP1 AND DP2

The idea proposed in [Bernstein et al. \(2018\)](#), that one definite article corresponds to an upper layer (DP1) while another corresponds to a lower one (DP2), finds a correlate in the Uniqueness Scale of [Löbner \(1985, 2011\)](#), further developed in [Ortmann \(2014\)](#) (see also [Sichel to appear, Schwarz 2013](#)). [Ortmann \(2014: 296\)](#) proposes "a continuum of types of uniqueness" ([Sichel to appear; Schwarz 2013](#)), where 'semantic uniqueness' involves unambiguous reference due to a noun's inherent meaning (e.g., 'the sun'). 'Pragmatic uniqueness' comes from the context, which could be, for example, deictic or anaphoric (e.g., 'the man'). Löbner's scale is "defined according to the degree of invariance of reference of nominal expressions" ([Ortmann 2014: 297](#)) and is presented here:

- (30) Scale of uniqueness: deictic SN < anaphoric SN < SN with establishing relative clause < definite associate anaphors < IN/FN¹¹ < proper names < 3rd person personal pronouns < 2nd and 1st person personal pronouns

Ortmann provides several examples of languages exhibiting a lexical split in their definite article systems, one lexical form corresponding to the upper functions of the scale and another to the lower functions. Fering, a Northern Frisian variety spoken on the island of Föhr, is an example of a language with such a split article system ([Ortmann 2014: 300](#); see [Ebert 1971](#)). Several other Germanic varieties are also discussed.¹²

The Uniqueness Scale and its applications elaborated by Ortmann (building on Löbner) relate straightforwardly to the multi-layered DP analysis we

10 Another question we have is, what counts as an argument in [Wiltschko's \(2013\)](#) analysis? Is it only the larger nominal expression, or can it also be the narrower one? We address these issues for our own analysis in [Bernstein et al. \(2018, 2019\)](#).

11 Individual Nouns (INs; 'sun', 'prime minister'), Functional Nouns (FNs; 'father', 'head'); Sortal Nouns (SNs; 'dog', 'stone', 'water') "classify the universe" ([Ortmann 2014: 295](#)).

12 [Ebert \(1971\)](#) introduced the term 'Situatives Unikum' for the class of expressions that denote a single referent in a particular communicative/pragmatic context, that is, when there are no other possible referents, as would be the case for examples like (22)–(23) above (see also Footnote 7) and (33) below. We thank an anonymous reviewer for directing us to this point.

have adopted for Balearic Catalan and French Picard. In particular, the *s-* article in Balearic Catalan and the *ch-* article in Picard would map to the higher part of the scale with deictic and anaphoric Sortal Nouns (SN). In contrast, unique referents would correspond to Individual and Functional Nouns (IN/FN), lower on the scale, and to the *l-* article in Balearic Catalan and Picard. Thus, there seems to be a clear correspondence between the Uniqueness Scale and the multi-layered DP analysis detailed above: the deictic and anaphoric SNs correspond to DP1 and the IN/FNs correspond to DP2.

Ortmann's (2014: 297) implementation of the Uniqueness Scale makes two important predictions. The first is that definite articles become more obligatory as we move higher up the scale and less obligatory as we move down the scale. In other words, a language is more likely to use definite articles in anaphoric contexts than in generic contexts. The likelihood of use relates to the fact that in generic and unique contexts there is a redundancy in the use of the article (particularly with proper names and pronouns), so cross-linguistically it is less likely to appear; in anaphoric contexts, the referent is identified via the discourse and so the definite article serves to connect the nominal expression to the discourse. The second prediction relates to diachrony. Specifically, over time the use of definite articles will generalize downward along the scale, not upward. In other words, a language that uses definite articles in anaphoric contexts may eventually come to use them also in generic and semantic unique contexts, perhaps eventually covering also those areas where they are functionally redundant (as with proper names and pronouns).

This second prediction is particularly important for understanding the transition from Latin, a language lacking definite articles, to Romance languages, which fully developed them. Applying Ortmann's predictions to the multi-layered DP, we expect overt content in DP1 earlier than in DP2. In other words, we expect to find definite articles expressed in anaphoric contexts before we find them expressed in generic contexts. This is because the article derives from a demonstrative form, retaining the demonstrative's anaphoric nature but not its deictic nature (Recall that the articles that we claim occupy DP2 are associated with genericity or uniqueness, not anaphoricity.) We will see this in Section 4 when we examine the development of definite articles in Latin.

We follow van Gelderen (2011) and others and assume that the grammaticalization process proceeds from a phrase (XP) to a head (X) and then (optionally) to a clitic. This process also involves loss of semantic features as well as morphological reduction. This path is consistent with the transition from a demonstrative (XP) to a definite article (X), and yields a DP1 deter-

miner. The step to a DP2 determiner involves a different phenomenon, one more akin to “input generalization” (Roberts 2019: 5). That is, the demonstrative grammaticalizes into a definite article (a syntactic head) and then this article can occur in new positions, also syntactic heads, which would be lower positions in our system.

With these ideas and predictions in mind, we turn to the Latin forms that will eventually give rise to the present-day definite articles in Romance languages. In particular, we are interested in exploring the following questions:

- a) What are the two major sources of definite articles in Latin and how do they fit into the multi-layered DP analysis?
- b) What are the stages of evolution from demonstrative, emphatic pronoun and interjection to the modern double determiner systems in Balearic Catalan and Picard?

4 LATIN *IPSE* AND *ILLE*

In Classical Latin the form *ipse* was used DP-externally as an adjunct emphatic pronoun related to a nominal expression to which it is not necessarily adjacent, as shown in the following example (taken from Hertzberg 2011: 173, her (1)):

- (31) *Ancillae tuae credidi (...) tu mihi non* (Classical Latin)
handmaid your believed.1SG you me not
credis ipsi
believe.2SG myself
'I believed your handmaid and you won't believe me (myself)?'
[Cicero, *Orat.*; 1st c.]

As Aebischer (1948) and Vincent (1997), among others, point out, this marker of emphasis started to be used like a demonstrative in Late and Medieval Latin when restructuring of the deictic demonstrative system occurred. In the restructured system, *iste*, originally the 2nd person medial demonstrative, expanded its use, subsumed the content of *hic*, and was used to indicate proximity to both the speaker and hearer. That is, it became a proximal demonstrative as opposed to the distal demonstrative *ille* (and its related forms) (see Vincent 1997, among others). As its emphatic value weakens, *ipse* enters into the paradigm of demonstratives and is used mainly to establish anaphoric relations. It undergoes morphological reduction and starts to appear in a DP-internal position, eventually yielding the present-day usage of the *s-* definite

article in Sardinian and Balearic Catalan. *Ille* was also the object of a weakening process. The semantic weakening involves the loss of deictic content, a first step in its evolution towards the Romance definite article.

In the preceding section we have seen that in contemporary varieties, Catalan *ipse*-derived articles (as well as Picard *ecce*-derived articles) convey different interpretive effects than *ille*-derived articles. These differences correspond to the upper and lower layers in the Uniqueness Scale (see Section 3). There is, in fact, evidence that *ille* and *ipse* did not express exactly the same content when used DP-internally in Late and Medieval Latin. Trager (1932) observed these differences (see also Ledgeway 2012: 92, Ferrazzano 2013, among others) in examples like the following:

- (32) *ergo quarta pervenimus in summitatem illam* (Late Latin)
 thus fourth arrived.1PL in summit that.F.
montis Dei sancti Syna, ubi data est lex (...)
 mount God saint Sinai where given is law
persubissemus in ipsa summitate
 reach.1PL in that-very.F summit
 ‘The fourth day we arrived at the summit of Mount Sinai, where
 law is given (...) we reached the very summit’
 [*Peregrinatio Silviae ad loca sancta*, 4th-5th c.]

In (32) *illam* identifies a unique referent, the summit of Mount Sinai, whereas *ipsa* is discourse-oriented and is interpreted anaphorically to that referent. This indicates that these demonstrative-like elements had different functions and semantic features. One element (*ille*) identifies unique referents and corresponds to ‘semantic uniqueness’, low on the Uniqueness Scale; the other element (*ipse*) participates in anaphoric relations and corresponds to ‘pragmatic uniqueness’, high on the Uniqueness Scale.

The same pattern is attested a few centuries later in the Latin of medieval texts in Catalonia. In the example in (33) the nouns introduced by *ipsa* are linked anaphorically to referents introduced previously in the discourse: in the case of *ipsa parroechia* to the parish identified in the preceding sentence and in the case of *ipsa ecclesia* to the church the whole paragraph refers to:¹³

¹³ Moran i Ocerinjauregui (1984: 43) and Villanueva (1850: 236–7) give slightly different transcriptions of this document. In (33) we reproduce Villanueva’s with the exception of *ecclesias* (between brackets), which is added by Moran in the space that Villanueva left as unidentified.

- (33) *Ideo ego Nantigisus Episcopus* (Medieval Latin)
 therefore I N. bishop
consecro ipsa ecclesia et trado illi parrochia
 consecrate.1SG the.F church and commit to-it parish
per illam strata de Caridona usque in Villafreda et
 by the.F road of Cardona up-to in Vilafreda and
usque in Villa Castellani et per illas ... [ecclesias] ...
 up-to in Vilacastellà and by the.F.PL churches
usque ad illa Guardia et per illo ribo de Clerano
 up-to to the.F Guardia and by the.M river of Clarà
usque ad illa strata. Trado ipsa parrochia ad ...
 up-to to the.F road commit.1SG the.F parish to
 ‘Therefore, I, bishop Nantigís, consecrate the church and
 commit a parish for it through Cardona’s road up to Vilafreda
 and Vilacastellà, and through the [churches] up to Guardia, and
 through the Clarà river up to the road. I entrust the parish to ...’
 [Bisbat d’Urgell; 9th-10th c.]

In contrast, most of the occurrences of the article derived from *ille* introduce unique referents: a particular road (*illam strata de Caridona*), river (*illo ribo de Clerano*) or toponym (*illa Guardia*). The evolution of the two forms eventually led to a stage where both were used in the same way. In the above example there is an instance of *illa* that can be interpreted as anaphoric because it is connected to the previously mentioned road to Cardona, which is a unique referent: *ad illa strata* ‘to the road’. Moran i Ocerinjauregui (1984: 36–40) provided several examples where the two determiners alternated with the same meaning and the same kind of nouns in documents from the 11th and 12th centuries.¹⁴

- (34) (a) *de ipso honore, ipsum mansum, ipsas vineas* (Medieval Latin)
 ‘from the honor’, ‘the authority’, ‘the vineyards’
 (b) *illo suo honore, illum mansum, illam vineam*
 ‘his honor’, ‘the authority’, ‘the vineyards’
 [Cartulari Sant Cugat, 11th–12th c.]

¹⁴ In the same collection of documents there are instances of the use of *ipsa* with toponyms:

- (i) *ipsa villa de Mazquefa, ipsa Guardia, ipsa Rocha*
 ‘the village of Masquefa’, ‘the Guardia’, ‘the Roca’
 [Cartulari Sant Cugat, 11th–12th c.; examples from Moran i Ocerinjauregui (1984: 36)]

This would indicate that, at this stage, both elements (*illa* and *ipse*) also alternated with this kind of unique referent (recall *illa Guardia* in (33)).

This suggests that, at a particular stage, both elements coincide in their uses. Thus, we can trace, as a first sketch, the evolution of these two Latin noun modifiers according to the schema below, which illustrates a process of grammaticalization in which a phrasal XP (emphatic pronoun and demonstrative) evolves into a head X (definite article), as detailed in van Gelderen (2011) (see also Giusti 2001, Iovino 2015, and Renzi 1992).

- (35) (a) *ipse* → *ipse* → *ipse*
 (emphatic pronoun) (noun modifier) (article)
- (b) *ille* → *ille* → *ille*
 (deictic pronoun) (noun modifier) (article)

In a way consistent with van Gelderen (2011), we hypothesize that Latin *ipse* and *ille* were originally DP-external elements (pronouns), and then, DP-internal XP noun modifiers that underwent a process of grammaticalization and eventually became heads that filled a D position in Late/Medieval Latin or Early Romance. At an initial stage, we envision the following split DP structure despite absence of lexical content:

- (36) [DP1 [D1] [DP2 [D2] [NP]]] (Latin)

Certain questions arise from the proposed structure in (36): Which elements fill each of the two D heads? Is one D head filled earlier chronologically than the other and if so, which one? We address these questions in the next section.

5 THE DEVELOPMENT OF DEFINITE ARTICLES IN CATALAN

In the previous sections we considered that *ille* and *ipse* were external modifiers to the DP. In the following subsections we will examine how these elements grammaticalize as heads in D1 and then generalize to D2 as we saw above in Catalan.

5.1 The development of DP1

Recall that in our multi-layered DP analysis the *s-* article of present-day Balearic Catalan corresponds to DP1 (see (24) in Section 2), which is the D position that encodes pragmatic uniqueness, the upper part of the Uniqueness Scale (see Section 3). After losing their deictic force and emphatic force respectively, both *ille* and *ipse* maintained their anaphoric value and were used to recover referents previously introduced in the discourse. We consider that

this anaphoric property, which is one of the prototypical uses of definite articles, is codified in the DP1 position of the complex syntactic structure that we propose. We can express this overlap between Latin *ille* and *ipse* and definite articles in the following way:

- (37) (a) *ipse* (Latin) → *s-* article (Medieval Catalan)
 [+anaphoric, +emphatic] [+anaphoric]
 (b) *ille* (Latin) → *l-* article (Medieval Catalan)
 [+anaphoric, +deictic] [+anaphoric]

We are assuming, then, that *ille* and *ipse* coexisted in Latin as two determiner-like elements that would serve to express anaphoric relations within the discourse. This would eventually lead to competition between them in the evolution of the determiner structure in Romance. Medieval Catalan shows that this was indeed the case. In this language, unlike other Romance languages such as Spanish or French, both the *ille*-derived article and the *ipse*-derived article survived in a robust way. The following examples show that in Medieval Catalan the two articles co-occurred with the same meaning in the same text:¹⁵

- (38) (a) *Met.se en convenenza ab zos* (Medieval Catalan)
 put.REFL in agreement with the.M.PL
cavalers e ab los homes
 knights and with the.M.PL men
 ‘He get into an agreement with the knights and the men’
 [*Greuges dels homes de Sant Pere Graudescale*; 12th c.]
 (b) *des mas d’en Ramon de sa Scara sa jova (...)*
 of.the farm of-ART Ramon of the.F Scara the.F plow.duty
del mas de Pere de sa Guardia sa jova
 of.the farm of Pere of the.F Guardia the.F plow.duty
 ‘from R.de sa Scara’farm, the plow duty (...) from P. de sa
 Guardia’s farm, the plow duty’
 [*Capbreu de Castellbisbal*; 12th c.]

¹⁵ Catalan medieval texts adopted the norms of the *Cancelleria Reial*. The documents generated by this administrative body were taken as a model for the Catalan written at that time (13th century on). Concerning the use of definite articles, the norms of the *Cancelleria* imposed the use of the *l-* article. For this reason, there are very few instances of the *s-* article in Medieval Catalan texts. See also Footnote 17. Medieval Catalan examples are taken from the *Corpus Informatitzat del Català Antic (CICA)*; we provide the [CICA](#) source document in brackets.

- (c) *del mas de sa Cortada so quart de so*
of.the farm of the.F Cortada the.M quarter of the.M
que.n eiz (...) d'en Venrel Fossar lo quart
that-of.it go.out of-ART Venrel Fossar the quarter
e.l brassage
and-the tithing
‘rom sa Cortada’s, the quarter of it produces (...) from Venrel Fossar’s, the quarter and the tithing’
[*Capbreu de Castellbisbal*; 12th c.]
- (d) *la sanc de Jhesu Crist (...) En la sogona vegada*
the.F blood of J. C. in the.F second time
sa sua sanc fo escampada (...) En la terça manera
the.F his.F blood was spread in the third way
Jhesu Crist la sua sanc escampà quant ...
J. C. the.F his.F blood spreaded when
‘the blood of Christ (...) in the second time the his blood was spread (...) in the third time Christ the his blood spread when ...’
[*Vides de Sant Rosselloneses*; 13th c.]

In (38a) the two coordinated phrases refer to the knights (*zos cavalers*) and the men (*los homes*) that were identified by the context. In (38b, 38c) both articles introduce the same kind of noun in the same context: a particular farm (unique referent) and a portion (a quarter) of the production. In (38d) the nominal expressions headed by *sa* and *la* refer anaphorically to Christ’s blood, which has been introduced previously (*la sanc de Jhesu Christ*).

This type of competition is the same as we have seen for *illum* and *ipsum* in Latin texts from Catalonia during the same centuries (see examples in (34) above). The competition between these two determiners gave way to the two systems that we find today in some dialects of Catalan. In present-day mainland Catalan the result is a uniform system favoring the *l-* article in DP1:

- (39) (a) [DP1 *s-/l-* [NP]] Medieval Catalan
(b) [DP1 *l-* [NP]] present-day mainland Catalan

However, in Balearic Catalan the resulting system favors the *s-* article in DP1:

- (40) [DP1 *s-* [NP]] present-day Balearic Catalan

The development of a uniform system in DP1 implied that one of the determiners was lost: the *s-* article in present-day mainland Catalan; the *l-* article in Balearic Catalan.

Internal and external factors intervened in this loss. Internally, the two determiners compete for the same D position resulting in the substitution of one by the other (Kroch 1989, 1994). Externally, the different contact situations and the historical context played a crucial role. The preservation of the *s*-article in Balearic Catalan in DP1 (recall that Balearic Catalan *l*- is in DP2 in our system) was favored by the fact that Catalan was introduced in the Balearic Islands in the second half of the 13th century, that is, before mainland Catalan became a variety with a uniform *l*- article system in most of the territory. The predominance of the *l*- article in mainland Catalan was, presumably, due to contact with Romance varieties where this determiner was the only element that occupied the DP1 position. Evidence for the extension of the *s*- article in Catalan comes from toponyms.¹⁶ Toponyms with the *s*- article remain in almost all of Catalonia, but they are relatively scarce in the western areas, to the extent that in the most western territories (the part closest to Aragon), not a single one is found (Moran i Ocerinjauregui 1984: 44). This suggests that the extension of the *l*- article took place progressively from west to east, probably because of contact with Aragonese and Spanish, two languages displaying an absence of the *s*- article. During the process of generalization of the *l*- definite article, the *s*- article was used in the vernacular for several centuries, mainly in eastern Catalan.¹⁷ Balearic Catalan was not subject to such continuous contact because of its geographic isolation and preserved the use of the *s*- article. In addition, the people who settled the Balearic Islands came from eastern Catalan-speaking areas where the *s*- form predominated and is still used today in some restricted areas.¹⁸

16 Ledgey (2012: 103–105) points out that, historically, the article derived from *ipse* was prevalent in several southern Romance varieties (Catalan, Occitan, Sardinian, Abruzzese, etc.). In terms of our analysis, this means that the *s*- article was also used in the DP2 position in Medieval Catalan (see Section 5.2) before being substituted by the *l*- article, which appears with all types of referents in Modern Catalan. Its presence in toponyms is interpreted as historical evidence of this generalized use.

17 As indicated in Footnote 15, it is hard to find evidence of the survival of the *s*- article in written documents from the 13th–14th centuries and beyond because of the influence of written norms, but some texts reflect that speakers were actually using it (see Moran i Ocerinjauregui 1984, Rabella i Ribas 1998).

18 Nowadays, the *s*- article in mainland Catalan (oral registers) is found in a few locations of the Costa Brava (Begur, Cadaqués, Tossa de Mar, etc.), on the northeastern coast of Catalonia (see Griera 1931, Veny & Pons 2001, or Busquet Isart 2010, among others). The *s*- article is also found in Tárkena, a small town near Alicante, in the southwestern limit of the Catalan linguistic boundary. The reasons for the presence of this form in this area are different from those for the northeastern part. This town and its surroundings were repopulated by people from the Balearic Islands in the 17th century.

5.2 *The development of DP2*

In our analysis of present-day Balearic Catalan (recall Section 2), DP2 is spelled out as *l-* and corresponds to semantic uniques and those involving kind reference, the lower part of the Uniqueness Scale. As we have seen in Section 4, the pattern with *ille* for semantic uniques and *ipse* for pragmatic uniques was already present in Medieval Latin. The following examples illustrate both uses within the same phrase or sentence:

- (41) (a) *ipsas parrochias de illo Obago* (Medieval Latin)
 the.F.PL parishes of the.M Obago
 (b) *et est ipsa terra ad illum Camporuvio*
 and is the.F land to the.M Camporuvio
 ‘and the land belongs to Camporuvio’
 [Bisbat d’Urgell; 9th c.; examples from Moran i Ocerinjauregui (1984: 42–43)]

The referents of *ipsas parrochias* and *ipsa terra* are found through anaphoric relations. That is, they correspond to pragmatic uniqueness, relying on text anaphoricity (i.e., higher on Ortmann’s scale). In contrast, *illo Obago* and *illum Camporuvio* denote particular places (they are toponyms). They express semantic uniqueness, relying on the inherent meaning of the lexical item, rather than on the linguistic context. This kind of referent is lower on Ortmann’s scale and, consequently, we map it onto the lower determiner position, DP2.

There was a preference for using *ille* with toponyms (see (41 a)) and with descriptive noun phrases like *illa parrochia d’Alaso* (see (41 b)), those that denote a unique referent in a way similar to toponyms.¹⁹ In either type of example, no anaphoric or deictic relation is needed to identify the referent:

- (42) (a) *illa Noz, illa Corre, illa Curriz,* (Medieval Latin)
 the.F Noz, the.F Corre, the.F Curriz,
illa Tentellagine
 the.F Tentellagine
 (b) *illa parrochia d’Alaso*
 the.F parish of-Alaso
 ‘Alaso’s parish’
 [Bisbat d’Urgell; 9th–10th c; examples from Moran i Ocerinjauregui (1984: 42–43)]

¹⁹ Adam Ledgeway (p.c.) points out that the use of *ille* with toponyms is parallel to the use of personal articles (e.g., Balearic Catalan *en-na*) with personal names, both sets of elements preceding proper names. In our system, both sets would correspond to DP2 (see Bernstein et al. 2019).

- (46) (a) *del remedi d'infern e de* (Medieval Catalan)
of.the avoidance of-hell and of
la glòria de paradís
the.F glory of paradise
‘of the avoidance of hell and of the glory of paradise’
[*Homilies d’Organyà*; 13th c.]
- (b) *saviea volia ésser gran, en ell, per manera de*
wisdom wanted.3SG be big in him by way of
gran fortalea e ...
big strength and
‘wisdom was intended to be big in him by means of big
strength and ...’
[Llull, *Llibre de meravelles*; 14th c.]
- (c) *Lo philosoph dix que lamp es cremament*
the.M philosopher said that lightning is burning
soptós de ...
sudden of ...
‘The philosopher said that the lightning is a sudden burning of
...’
[Llull, *Llibre de meravelles*; 14th c.]
- (d) *Meraveyllà’s lo bon hom com paciència sa*
was.amazed-REFL the.M good man when patience his
ira no vensia
anger not overcome
‘The good man was amazed when his anger didn’t overcome
patience’
[Llull, *Llibre de meravelles*; 14th c.]

In present-day Catalan all these examples of generics and unique referents require the presence of the definite article. In fact, anticipating the modern usage, Medieval Catalan texts show that these kinds of referents alternate between an absence and presence of an accompanying definite article. The first two of the following examples correspond to older documents, and the third (later) one is from the same source (*Llibre de meravelles*) as (46b, 46c, 46d):

Emphatic elements and the development of definite articles

- (47) (a) *les malesfeites autres que zel* (Medieval Catalan)
 the.F.PL crimes other.F.PL that that.M
fa en la onor
 makes in the.F honor
 ‘the other crimes that he commits against honor’
 [*Greuges dels homes de Sant Pere Graudescale*; 12th c.]
- (b) ... *que si l malfeytor la justícia contendrà*
 that if the criminal the.F justice stop.FUT.3SG
 ‘... if justice stops criminals’
 [*Usatges de Barcelona*; 13th c.]
- (c) ... *hac dites estes paraules et moltes altres del lamp*
 have.1SG said these.F words and a.lot.of.F.PL other.F.PL
 of.the lightning
 ‘... I have said these words and many others about lightning’
 [*Llull, Llibre de meravelles*; 14th c.]

Oblique contexts favor the absence of the definite article in Old Catalan:

- (48) (a) *avant no gós affermar o procurar* (Medieval Catalan)
 before not dared confirm or procure
de metre alcun misatge ab senyor
 of put some message with master
 ‘before s/he dared to confirm or procure giving some message to the master’
 [*Manual de consells de la Ciutat de Valencia*; 14th c.]
- (b) *vos tremeitem una letra qui ·s’adreça a mestre Jacomí*
 you send.1PL a.F letter that is.directed to master
 Jacomí
 ‘we send you a letter addressed to master Jacomí’
 [*Epistolari Ferran I*; 13th c.]

All these contexts correspond to our DP2. These data support the idea that the DP2 head can be empty or occupied by the *l*- definite article in Medieval Catalan:

- (49) [DP1 [DP2 \emptyset /*l*- [NP]]] Medieval Catalan

5.4 Interim Conclusion

We have seen that in earlier stages of Catalan DP1 had to be filled while DP2 could remain empty in generic, unique and oblique contexts. Recall Ortmann's observation that definite articles become more likely as we go higher up the scale. This also means that the likelihood diminishes with lower positions on the scale.

From our analysis this implies that, diachronically, DP1 fills before DP2. This is consistent with the predictions at the end of Section 3, although Ortmann does not specifically analyze the internal structure of DPs. So, the structures in (50) and (51) for Catalan show the generalization of expression of definite articles from DP1 to DP2:

- (50) [DP1 *s-/l-* [DP2 \emptyset /*l-* [NP]]] Medieval Catalan
- (51) (a) [DP1 *s-* [DP2 *l-* [NP]]] present-day Balearic Catalan
- (b) [DP1 *l-* [DP2 *l-* [NP]]] present-day mainland Catalan

The change illustrates two ways of resolving the competition between the *l-* and *s-* definite articles: a) in mainland Catalan (with the exceptions noted in Footnote 18) the evolution led to the suppression of one of the forms (the *s-* article) and the generalization of the other (the *l-* article), rendering the whole system isomorphic; b) in Balearic Catalan, on the contrary, the grammar maintained the distinction and, consequently, led to the double definite article system we still see in the modern variety. In this variety, the choice between one or the other definite article finds its origins in the original distinction between *ipse* and *ille* in Late and Medieval Latin, a distinction we represented as DP1 and DP2 in our multi-layered analysis.

As we proposed in Section 5.1 (see (37)), *ipse* and *ille* had an anaphoric value, a value associated with pragmatic uniqueness and corresponding to the upper layers in Ortmann's scale. This value would map onto our DP1. In contrast, DP2 is associated with semantic uniques and *ille* is the best candidate for filling this position.

Our syntactic analysis, coupled with the Uniqueness Scale, explains why the following two scenarios are not attested:

- (52) (a) * [DP1 \emptyset [DP2 *l-* [NP]]]
- (b) * [DP1 *l-* [DP2 *s-* [NP]]]

Example (52 a) would correspond to a system that has definite articles only for unique referents, but not for anaphoric relations. This type of system is

excluded because if DP2 is instantiated this would entail that DP1 must also be instantiated. The impossibility of a system like (52 b) would be related to the etymology of the elements involved: *ipse* was emphatic and anaphoric and so would have been initially linked to DP1, whereas *ille* would have corresponded to both DP1 and DP2 as we saw in Latin in example (32) in Section 4.²⁰ That is, this unattested system wouldn't easily conform to the hypothesis that the features of *ille*-derived articles match those of DP2, but those of *ipse*-derived articles at that time do not.²¹

6 THE STATUS OF *CH-* IN OLD PICARD

In the limited work on Old Picard *ch-*, the adopted view is that it developed from the emphatic interjection (Latin *ecce*) in the 17th century (Panwitz 2018, citing Eloy 1986). Panwitz notes that the *ch-* form seems to have displaced the *l-* form in the relevant contexts, since the latter was already in use. Recall that for us present-day Picard *ch-* corresponds to DP1 and *l-* to DP2 (see Section 2).

(53) [DP1 *ch-* [DP2 *l-* [NP]]] (Picard)

In line with the view of Eloy and others and writing about *ch-* as definite article in Old Picard, Gossen (1951: 100) states, “La scripta ne le connaît pas, à moins que, dans une expression telle que *empoursivant che jour devant dit*, on ne veuille considérer *che* comme article, l'idée démonstrative étant exprimée par *devant dit*.” (‘It is unknown within Old Picard texts, unless, in an expression such as *empoursivant che jour devant dit* “following the aforementioned day” one considers *che* as an article, the demonstrative aspect being expressed by *devant dit*.’) Gossen’s comment deserves further investigation because in the 13th–14th century *Chronique Artésienne* (published in Funck-Brentano 1899), *l-* appears regularly with *devant dit* and *ch-* never does:

(54) *li roys devant dit, li devant dit roys,* (Old Picard)
le samedi devant dit
 ‘the king aforementioned’, ‘the aforementioned king’, ‘the
 aforementioned Saturday’
 [*Chronique Artésienne*]

20 In present-day standard Catalan, which lacks the *s-* article, both DP1 and DP2 can be filled with an *l-* form, the latter reserved for generics and uniques. Also appearing in DP2 is the personal article *en*. This function is realized with an *l-* form in some varieties of Catalan (see Bernstein et al. 2019).

21 This doesn't exclude the possibility that an *ipse*-derived article can occupy DP2, just not in double definite article systems. In Sardinian, an *ipse*-derived article can occupy DP2, and also DP1 (see Bernstein et al. 2018).

Putting the two sets of data together (the one provided by Gossen and the one from the *Chronique* above), we suspect that the more likely explanation is that at a certain stage both *ch-* and *l-* shared a demonstrative or anaphoric function. This recalls the patterns seen above for Medieval Catalan, and strongly suggests that the two Old Picard forms were in competition for DP1.

Consistent with this idea and in contrast to Gossen's claim about *ch-* as demonstrative, Brébion (1907: §16) provides an example from the early 14th c. of *ch-* in its article function:

- (55) *Si avint qu'il assanla ches* (Old Picard)
 thus happened that-he assembled the/these
gens de ches mestiers
 people from the/these trades
 'It happened that he assembled the people from the trades.'
 [*Chronique Artésienne*]

Brébion (1907: §16) states, "On trouve déjà au moyen âge des exemples de l'extension de l'emploi de ces adjectifs..." ('Extension of the use of these [demonstrative] adjectives is already found in the Middle Ages...'). Gossen (1951: 100), in response to the Brébion example, writes: "Il faut croire que les scribes hésitaient à employer ces formes qui leur semblaient par trop populaires." ('One must believe that the scribes hesitated using these forms that seemed to them too colloquial.')

²² These differing non-contemporaneous views raise the issue of distinguishing between oral and written registers and more generally, informality vs. formality of texts. For our purposes here, we wonder if Picard *ch-* as a definite article was actually in use earlier than the 17th c. The work of Brébion and the examples from the *Chronique Artésienne* suggest that it was. We discuss here some of the Old Picard data.

The *Chronique Artésienne* (1295–1304) contains many examples with *ch-* forms, though it is often difficult to categorize individual examples definitively as demonstratives or definite articles (for discussion see Adams 2013: 490ff and Lyons 1999). Nevertheless, a not insignificant number of the examples are good candidates for articles. In (56 a), for example, the interpretation of *ches nouvelles* ('the news') strongly suggests a definite article function; in (56 b), *chele bataille* ('the battle') serves as head of a restrictive relative clause,

²² It is difficult to fully assess the incidence of *ch-* as a definite article in Old Picard because it doesn't appear in most medieval texts. Our examples so far are limited to the *Chronique Artésienne*. Gossen's comment is consistent with the idea that Old Picard texts reflect formal or literary language, rather than informal or colloquial speech. This might account for the scarcity of *ch-*.

again indicating a definite article function; in (56 c), a PP modifier following *ches gens* ('the people') also supports a definite article function for *ch-*.²³

- (56) (a) *Et li quens de Savoie, qui estoit de par* (Old Picard)
 and the count of Savoie who was of side
le roy se retraist et porta ches nouveles au
 the king REFL retreated and carried this news to.the
roy que li Flamenc avoient dit.
 king that the Flemish had said
 'And the count of Savoie, who was on the side of the king,
 returned and carried to the king the news that the Flemish had
 communicated.'
- (b) *et demoura Willaumes de Julers en chele bataille qui*
 and remained Willaumes de Julers in that battle that
estoit si drue et si espesse ens u camp
 was so tight and so thick in the encampment'
 'and Willaumes de Julers remained in the battle that was so
 tightly packed and so thick in the encampment'
- (c) *Et envoia grant plenté de ches gens du teroir*
 And sent great plenty of these people of.the region
de Bruges et de le castelerie d'Yppre devant le
 of Bruges and of the castle of-Yppre in-front-of the
castel de Cassel
 castle of Cassel
 'And he sent a great number of the people from the region of
 Bruges and from the jurisdiction of the castle of Yppre to the
 front of the castle of Cassel.'
 [Chronique Artésienne]

At the same time and at first glance, many instances of *ch-* in the *Chronique Artésienne* seem to involve a demonstrative element rather than a definite article. Upon closer inspection, however, many of these involve textual anaphora, a function that bridges demonstratives and definite articles and would correspond to the upper layer of Löbner's (1985, 2011) Uniqueness Scale as elaborated in Ortmann (2014). We saw this pattern earlier (in Section 4) for Medieval Latin *ipse* and *ille*, and in the following examples we see the same gen-

²³ The *ch-* demonstrative (from Latin *ecce*) requires *-l* (from Latin *ille*) in its feminine form, at least in some varieties, as in (56 b) *chele bataille*. The form *chele* also precedes vowel-initial words (regardless of gender), as in (57 a) *chele abeye*. We saw this pattern above in contemporary varieties as well.

eral system for the descendants of *ecce* and *ille* (namely, *ch-* and *l-* forms) in the late 13th and early 14th centuries:

- (57) (a) *Et fu toute l'abaye reubée et leur* (Old Picard)
and was all the-abbey pillaged and their
moustiers et leur tresorie et tout *quankes*
monastery and their treasury and everything that
il avoient en chele abeye de Lo.
they had in that abbey of Lo.
'And the whole abbey was pillaged, and their church, and their
treasury, and everything that they had in that abbey of Lo.'
- (b) ... *pluiseurs viles et viletes* *dessi Aloes - et*
many towns and small-towns below Aloes and
Loeces; et les gens de ches viles s'assanlerent
Loeces; and the people of those towns REFL.gathered-3PL
'... many towns and small towns below Aloes and Loeces; and
the people of those towns gathered.'
- (c) *Et avoit-on mené .V. engiens qui getoient pierres*
and had-one brought five catapults that launched stones
et espringales pour traire as Flamens, et li quens
and crossbows for to-shoot the Flemish and the count
de Boulongne estoit derrière ches engiens pour warder.
of Boulogne was behind those catapults for to-guard
'And they had brought five catapults that launched stones, and
large crossbows with which to shoot at the Flemish, and the
count of Boulogne was behind those catapults to stand watch.'
- (d) *Et s'acorderent a le vile rendre du jour*
and REFL.agreed-3PL at the town to-surrender of.the day
sainte Crois en .x. jours- : en telle manière que se li
Saint Croix in 10 days in that way that if the
Flamens venoient courre sus le roy dedans ches
Flemish came-3PL running under the king within those
.x. jours, ...
10 days
'And they agreed to surrender the town 10 days after the
festival of Sainte Croix: in such a way that if the Flemish were to
overrun (attack) the king within those 10 days ...'
[*Chronique Artésienne*]

But the textual anaphora cases deserve further scrutiny as well, since additional examples plainly illustrate the *l-* forms playing an anaphoric role. In the

following example, the second instance of *le vile* ('the town') is used anaphorically, referring back to Dent:

- (58) ... *et rendi le vile du Dent (...)* *Et*
 and surrendered-3SG the town of Dent (...) and
ne demoura en le ville homme ne femme ...
 NEG remained in the town man NEG woman ...
 '... and he surrendered the town of Dent ... And there remained
 in the town not a single man or woman...'
 [*Chronique Artésienne*]

As mentioned above, we find the most plausible explanation for this apparent inconsistent behavior to be that the two emerging articles, *ch-* and *l-*, are in competition for DP1. At some point, however, *ch-* wins out.

Examples involving modification with adjectives seem to favor *ch-* in the *Chronique Artésienne*:

- (59) (a) *par ches presentes lettres* (Old Picard)
 'by the/those present letters'
 (b) *chele dolereuse desconfiture*
 'the/that painful discomfort'

Recall from Section 2 (present-day Picard and Balearic Catalan) that such modified nouns are expected to appear with *ch-* articles in a multi-layered DP system. Specifically, they would correspond to DP1. The *ch-* articles would also correspond to the upper layer of the Uniqueness Scale, overlapping with demonstratives,

In the *Chronique Artésienne* *l-* forms also display systematic patterns. For example, nouns related to the nobility appear exclusively with *l-* articles, never with *ch-*. The same is true for unique common nouns such as *mer* ('sea'):

- (60) (a) *li/les prinche(s), li viscontes, le/li duc, li quens, le contesse, li roy,*
le/la royne, etc.
 'the prince', 'the viscount', 'the duke', 'the count', 'the countess',
 'the king', 'the queen', etc.
 (b) *amiraus de le mer ...*
 'admirals of the sea ...'

These semantic uniques appear towards the bottom of the Uniqueness Scale and would be predicted to appear with an *l-* form in a split article system (and is in fact what we have seen in Balearic Catalan).

Eloy's (1986) claim that Picard *ch-* developed into a definite article in the 17th c. deserves further scrutiny. Examples from the *Chronique Artésienne* suggest that the process was well underway at a much earlier time. Although it would be premature to claim that demonstrative *ch-* in Old Picard had already fully grammaticalized into a definite article, we have seen some clear distinctions between the functions of the *l-* forms and those of the grammaticalizing *ch-* forms, the patterns already aligning with Löbner's Uniqueness Scale (as elaborated in Ortman 2014).

Recall example (35) in Section 4, where we presented a schema representing the grammaticalization of Latin *ipse* and *ille*. We can now add to it the development of the definite article in Picard from Latin *ecce* (ignoring here the morphological and semantic changes the three forms have undergone):

- (61) (a) *ipse* → *ipse* → *ipse*
 (emphatic pronoun) (noun modifier) (article)
- (b) *ille* → *ille* → *ille*
 (deictic pronoun) (noun modifier) (article)
- (c) *ecce* → *ecce* → *ecce*
 (emphatic interjection) (noun modifier) (article)

The evolution of *ch-* in Old Picard is striking, since it provides an example of the grammaticalization of an emphatic form (*ecce*) with anaphoric properties different from the Old Catalan one (*ipse*). Despite probable historical differences between the evolution of Catalan *s-* and Picard *ch-*, the end result is remarkably similar in the modern languages, consistent with both the multi-layered DP analysis and the Uniqueness Scale. This parallel cross-linguistic evolution suggests general processes are at play and bears out Ortman's predictions concerning the obligatory expression of definite articles and the diachronic evolution of forms.

7 CONCLUSIONS

In this paper we examine two Romance languages with double definite article systems. We trace the evolution of the present-day forms to Latin *ipse* and *ille* for Balearic Catalan and Latin *ecce* and *ille* for varieties of Picard. In both languages, the elements that were emphatic, *ipse* and *ecce*, were used to introduce nouns when they were interpreted anaphorically. This property was shared with *ille*. The result of competition between *ipse/ecce* and *ille* is that *ipse* and *ecce* won out in Balearic Catalan and Picard. Although *ille* was precluded in anaphoric contexts in both varieties, it became reserved for other contexts.

This evolution receives a natural explanation in a multi-layered approach to DP structure: a higher layer, DP1, corresponds to anaphoric uses of the definite article; a lower layer, DP2, corresponds to uses such as those associated with generics and uniques. Ortmann (2014) offers a semantic template for the syntactic analysis. According to his Uniqueness Scale, anaphoric interpretation would involve pragmatic uniqueness, corresponding to DP1. Generics and uniques would involve semantic uniqueness, corresponding to DP2. Generalization of these forms is unidirectional: DP1 to DP2, but not the other way around.

Our analysis raises interesting questions about Romance languages with single article systems, such as French and Sardinian. We predict that the *l*-forms in French and *s*-forms in Sardinian started as DP1 elements, given their development from *ille* and *ipse*, respectively. That these forms are also found in contexts that would correspond to DP2, such as with uniques and generics, is explained by the proposal that the definite article's reach generalized to contexts lower on the Uniqueness Scale.

REFERENCES

- Adams, James Noel. 2013. *Social variation and the Latin language*. Cambridge: Cambridge University Press.
- Aebischer, Paul. 1948. Contribution à la protohistoire des articles *ille* et *ipse* dans les langues romanes. *Cultura neolatina* 8. 181–203.
- Bernstein, Judy B., Francisco Ordóñez & Francesc Roca. 2018. Sardinian descendants of *ipse* in comparative relief. In Annamaria Chilà & Alessandro De Angelis (eds.), *Capitoli di morfosintassi delle varietà romanze d'Italia: teoria e dati empirici* (Supplementi al Bollettino, 21), 317–336. Palermo: Centro di Studi Filologici e Linguistici Siciliani.
- Bernstein, Judy B., Francisco Ordóñez & Francesc Roca. 2019. On the emergence of personal articles in the history of catalan. In Miriam Bouzouita, Anne Breitbarth, Lieven Danckaert & Elisabeth Witzhausen (eds.), *Cycles in language change*, 88–108. New York and Oxford: Oxford University Press.
- Brébion, Louis. 1907. *Étude Philologique: sur le Nord de la France (Pas-de-Calais, Nord, Somme)*. Paris: Librairie Honoré Champion.
- Busquet Isart, Núria. 2010. *El parlar salat. Descripció, àmbit geogràfic i ús*. Barcelona: Universitat de Barcelona dissertation.
- CICA. n.d. Corpus informatitzat del Català antic. <http://www.cica.cat/>.
- Corblet, Jules. 1851. *Glossaire étymologique et comparatif du patois picard, ancien et moderne*. Paris: Dumoulin.
- Debrie, René. 1983. *Eche pikar bèl é rade*. Paris: Omnivox.

- Ebert, Karin. 1971. *Referenz, Sprechsituation und die bestimmten Artikel in einem nord-friesischen Dialekt (Fering)*. Kiel: Christian-Albrechts-Universität dissertation.
- Eloy, Jean-Michel. 1986. *Les Formes de l'article dans un corpus picard contemporain*. Amiens: Publications du Centre d'Études Picardes.
- Ferrazzano, Lisa Reisig. 2013. *Demonstratives in motion: The grammaticalization of demonstratives as a window into synchronic phenomena*. New York: City University of New York dissertation.
- Funck-Brentano, Frantz. 1899. *Chronique Artésienne (1295-1304) et Chronique Tournaisienne (1296-1314)*. Paris: Alphonse Picard et Fils.
- van Gelderen, Elly. 2011. *The linguistic cycle: Language change and the language faculty*. Oxford: Oxford University Press.
- Giusti, Giuliana. 2001. The birth of a functional category: From Latin ILLE to the Romance article and personal pronouns. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*, 157–171. Amsterdam: North-Holland.
- Gossen, Charles Théodore. 1951. *Petite grammaire de l'ancien picard*. Paris: Librairie Klincksieck.
- Griera, Antoni. 1931. *Gramàtica històrica del català antic*. Barcelona: La Casa de Caritat.
- Grimalt, Josep A. 2009. Els articles en el parlar de Mallorca. *Estudis romànics* 31. 55–75.
- Haigneré, Daniel. 1901. *Le patois boulonnais*. Boulogne-Sur-Mer. Société Académique de Boulogne-Sur-Mer.
- Hertzenberg, Mari Johanne. 2011. Classical and Romance usages of *ipse* in the Vulgate. In Eirik Welo (ed.), *Indo-European syntax and pragmatics: Contrastive approaches, Oslo Studies in Language* 3.3, 173–188. Oslo: University of Oslo.
- Iovino, Rossella. 2015. Some observations on *ille* and *ipse* in the *Mulomedicina Chironis*. *Journal of Latin Linguistics* 14. 269–306.
- Jones, Michael A. 1993. *Sardinian Syntax*. London: Routledge.
- Kayne, Richard S. & Jean-Yves Pollock. 2010. Notes on French and English demonstratives. In Jan-Wouter Zwart & Mark de Vries (eds.), *Structure preserved: Studies in syntax for Jan Koster*, 215–228. Amsterdam: John Benjamins.
- Kroch, Anthony. 1989. Function and grammar in the history of English: Periphrastic DO. In Ralph Fasold & Deborah Schiffrin (eds.), *Language change and variation*, 133–172. Amsterdam: John Benjamins.

- Kroch, Anthony. 1994. Morphosyntactic variation. In Kathryn Beals (ed.), *Papers from the 30th Regional Meeting, Chicago Linguistic Society*, 180–201. Chicago: Chicago Linguistic Society.
- Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Oxford: Oxford University Press.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Löbner, Sebastian. 1985. Definites. *Journal of Semantics* 4. 279–326.
- Löbner, Sebastian. 2011. Concept types and determination. *Journal of Semantics* 28. 279–333.
- Moran i Ocerinjauregui, Josep. 1984. *El Capbreu de Castellbisbal: Estudi filològic i lingüístic d'un text català antic*. Barcelona: Publicacions Universitat de Barcelona.
- Ortmann, Albert. 2014. Definite marker asymmetries and concept types. In Thomas Gamerschlag, Doris Gerlans, Rainer Osswald & Wiebke Petersen (eds.), *Frames and concept types*, 293–321. Berlin: Springer.
- Panwitz, Amber. 2018. The system is definitely divided: A variationist analysis of Picard's definite articles *CH* and *L*. Bloomington Indiana University. (Unpublished manuscript).
- PICARTEXT. n.d. Centre d'Études des Relations et Contacts Linguistiques et Littéraires de l'Université de Picardie Jules Verne. <https://www.u-picardie.fr/LESCLaP/PICARTEXT/Public/>.
- Renzi, Lorenzo. 1992. Le développement de l'article en roman. *Revue roumaine de linguistique* 37. 161–176.
- Rabella i Ribas, Joan Anton. 1998. *Un matrimoni desanvoigut i un gat metzinat. Procés criminal barceloní del segle XIV*. Barcelona: Institut d'Estudis Catalans.
- Roberts, Ian. 2019. *Parameter Hierarchies and Universal Grammar*. Oxford: Oxford University Press.
- Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7/10. 534–559.
- Sichel, Ivy. to appear. Weak and strong articles, diversity, and uniformity. In Solveiga Armoskaite & Martina Wiltschko (eds.), *Oxford handbook of determiners*, Oxford and New York: Oxford University Press.
- Trager, George L. 1932. *The use of the Latin demonstratives (especially ille and ipse) up to 600 A.D., as the source of the Romance article*: Columbia University dissertation.
- Veny, Joan & Lúdia Pons. 2001. *Atles lingüístic del domini català*. Barcelona: Institut d'Estudis Catalans.

- Villanueva, Jaime. 1850. *Viage literario a las iglesias de España*. Madrid: Real Academia de la Historia.
- Vincent, Nigel. 1997. The emergence of the D-system in Romance. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 149–169. Cambridge: Cambridge University Press.
- Vincent, Nigel. 1999. The synchrony and diachrony of deixis in Old Neapolitan (with special reference to the *Cronache e ricordi* of Loise De Rosa). University of Manchester. (Unpublished manuscript).
- Wiltschko, Martina. 2013. Descriptive relative clauses in Austro-Bavarian German. *Canadian Journal of Linguistics* 58(2). 157–189.
- Zamparelli, Roberto. 1995. *Layers in the Determiner Phrase*. University of Rochester dissertation.

Judy B. Bernstein
Department of Languages and Cultures
William Paterson University
300 Pompton Road
Wayne, NJ 07470 USA
bernsteinj@wpunj.edu

Francisco Ordóñez
Linguistics Department
Stony Brook University
Stony Brook, NY 11794
USA
francisco.ordonez@stonybrook.edu

Francesc Roca
Dept. Filologia i Comunicació
Universitat de Girona
Pça. Ferrater Mora, 1
17004 Girona Spain
francesc.roca@udg.edu