

Modern Romance split DP systems in their historical context

Judy B. Bernstein, William Paterson University
Francisco Ordóñez, Stony Brook University
Francesc Roca, Universitat de Girona

Definite articles in Romance languages like Spanish developed from the Latin distal demonstrative *ille* and display *l-* forms (1). Definite articles also developed from the Latin emphatic *ipse* in other varieties such as Sardinian (Aebischer 1948, Vincent 1997, Ledgeway 2012) and display *s-* forms (2).

- (1) *el* niño (Spanish) (2) *su* babbu de Maria (Sardinian; Jones 1993)
'the.M boy' the father of Maria

More striking are varieties displaying a split system, such as Balearic Catalan (*s-* and *l-*) and Picard (*ch-* and *l-*). Picard *ch-* forms derive from Latin emphatic *ecce*:

- (3) a. *Sa* muntanya que veus és molt alta. b. Anem a *la* muntanya. (Balearic Catalan)
the.F mountain that see.2PSG is very tall go.1PPL to the.F mountain
- (4) a. *ch'* pus riche b. *le* roy (Boulogne Picard; Haigneré 1901)
the.M most rich the.M king

In the modern varieties with a split DP system each definite article is specialized: *l-* appears with 'semantic uniques' (Ortmann 2014 building on Löbner 1985, 2011), those involving reference due to a noun's inherent meaning (e.g., 'the sun'); *s-/ch-* appears with 'pragmatic uniques', those involving anaphoric or deictic context (e.g., 'the chair'). Our claim is that the split systems of Balearic Catalan and Picard correspond to a multi-layered DP, schematized in (5):

- (5) [DP1 *s-/ch-* [DP2 *l-* [NP]]]

In this structure *s-* and *ch-* occupy a higher DP layer than *l-*. Evidence for this structure comes from the co-occurrence of both determiners in a single nominal expression:

- (6) *es l'*amo, *es l'*avi (s- + l-) (Balearic Catalan)
'the mister', 'the grandfather'
- (7) *ch' l'*esprit, *ch' l'*infant (ch- + l-) (Boulogne Picard; Haigneré 1901)
'the spirit', 'the child'

There are also syntactic differences between these determiners. For example, in Balearic Catalan DP1 *s-* forms may co-occur with a specialized personal article (*en/na*) that precedes only [+human] proper names (see (8)). We take *en/na* to correspond to a lower position akin to DP2.

- (8) *es* conco *en* Toni (Menorcan Balearic Catalan)
the uncle *en* Toni
'Uncle Toni'

Furthermore, there is evidence that vocatives, typically incompatible with definite articles (**the waiter!*), appear with DP2 (*l-*) articles, but not with DP1 (*s-*) articles in Balearic Catalan:

- (9) a. *s'avi! (Balearic Catalan)
 b. l'avi!
 'the-grandfather'

In order to better understand the syntax of the synchronic forms (*s-/ch-* vs. *l-*), we also study their historical development, which provides further evidence for the analysis in (5). Specifically, we trace the grammaticalization of the Latin emphatic (*ipse, ecce*) and demonstrative (*ille*) forms. The multi-layered DP, schematized in (5) above for the present-day varieties, started to develop in the transition from Latin to early Romance for Balearic Catalan, and arguably at a later period for Picard. In this early period we already see the split DP system emerging: an emphatic form competes with a demonstrative form for DP1. At a later period, DP2 is filled with a grammaticalized demonstrative form:

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

In fact, the overlapping and differentiated semantic nature in Latin of *ipse-ecce* (emphatic and anaphoric) versus *ille* (anaphoric or unique) provides a clue about how they will eventually map onto the split DP. Medieval Latin (11) shows that these forms did not have the same use (Vincent 1997, Ledgeway 2012, Trager 1932).

- (11) ergo quarta pervenimus in **summitatem illam** montis Dei sancti Syna, ubi data est lex
 thus fourth we.arrived in summit **that** mount God saint Sinai, where given is law
 ... iubente Deo persubissemus in **ipsa summitate** (Medieval Latin)
 was.given God reach.we in **that-very** summit
 (*Peregrinatio Silviae ad loca sancta*; 4th/5th c.)

Latin *illam* refers to a unique referent; *ipsa* is discourse-oriented and anaphoric to that referent. *Illam* corresponds to Ortmann's 'semantic uniqueness', *ipsa* to his 'pragmatic uniqueness'. These same patterns are found in Medieval Latin texts in Catalonia, where we observe *ipse* and *ille* modifying identical nouns in the same sentence. The two determiners evolved and they started to be used with the same anaphoric value as illustrated in the following example:

- (12) consecro **ipsa ecclesia** et trado illi parrochiam per illam strata de Cardona...
 et per **illas ecclesias** usque ad illa Guardia ... (Medieval Latin in Catalonia)
 'I consecrate **the/that** church and constitute the parish through Cardona's road ...
 and through **the/these** churches and Guardia ...
 (*Bisbat d'Urgell*; 9th c.; from Moran 1984: 42-43)

Ortmann (p. 297) predicts that the use of the definite article can spread over time from the more anaphoric and deictic functions (higher on his scale) to the more redundant functions (lower on his scale): generics, uniques, proper names. Applied to our analysis, this suggests that DP1 is instantiated earlier than DP2. Data from Medieval Catalan bear out the prediction. That is, DP2 (but not DP1) contexts admit bare nouns, no longer possible in the modern language:

- (13) a. del remedi d' **infern** e de la glòria de **paradis**. (Medieval Catalan)
 of-the hell's avoidance of the glory of paradise. (*Homilies d'Organyà*; 13th c.)
 b. Lo philosoph dix que **lamp** es cremament soptós de ...
 the philosopher said that the lightning is a sudden burning of ...

Although the scant literature on Old Picard states that *ch-* developed from Latin *ecce* in the 17th century (Eloy 1986), many of the patterns of Latin and Medieval Catalan are already displayed in the 13th-14th c. in *Chronique Artésienne*. Moreover, Picard is interesting because *ecce* (rather than *ipse*) competed with *ille*, eventually winning out as the occupant of DP1. DP2 came to be occupied by *l-* in both present-day Picard and Catalan. Despite probable historical differences, the end result is remarkably similar in the modern varieties, supporting the idea of general underlying processes.