SYNCRETISM AND ORDERING IN THE EVOLUTION OF CATALAN PRONOMINAL CLITIC CLUSTERS[*]

Teresa Cabré
Universitat Autònoma de Barcelona

ABSTRACT. This paper examines the 3rd person clitic combinations found in a digital corpus of Catalan texts dating from the 11th century to the first half of the 18th (the CICA) and attempts to clarify the origin of the current clitic system of colloquial non-Valencian Catalan. Scrutiny of the database shows that the locative *HI* (i.e., *hi* or its variants *i/y/hic*) replaced the canonical dative clitic of 3rd person clusters in the 14th-15th centuries in both singular and plural forms, contrary to what has previously been claimed. The medieval patterns of usage that the data reveal are very close to those occurring in colloquial non-Valencian Catalan as it is spoken nowadays, as opposed to those seen in Valencian Catalan, where a locative clitic is no longer present. On the basis of the CICA' data, we argue that the incompatibility of plural morpheme combinations in Old Catalan—among other reasons—forced the generalization of the morpheme /i/ as a dative marker, thus converting it into the true 'elsewhere' item of the Catalan clitic system. The similarity between medieval and modern colloquial non-Valencian Catalan clitic forms allows us to analyze them in the same way. Specifically, we suggest that there is only one clitic area for these clusters in which the *HI* works as a place nominal located structurally in the nominal layer.

Keywords: dative; locative; clitic clusters; elsewhere clitic; Old Catalan

RESUMEN. Este artículo examina las combinaciones de clíticos de tercera persona que se encuentran en un corpus digital de textos catalanes (CICA), datados desde el siglo XI hasta la primera mitad del XVIII, e intenta esclarecer el origen del actual sistema de clíticos del catalán coloquial no valenciano. El escrutinio de la base de datos muestra que el locativo *HI* (i.e., *hi* o sus variantes *i/y/hic*) reemplazó al clítico dativo canónico en los grupos de tercera persona en los siglos XIV y XV tanto en singular como en plural, contrariamente a lo que se ha dicho hasta ahora. Los patrones medievales que revelan los datos obtenidos están muy cerca de los que se dan en el catalán no valenciano coloquial actual, frente a los del catalán valenciano, donde ya no está presente el clítico locativo. Sobre la base de los datos del CICA, argumentamos que la incompatibilidad de las combinaciones de morfemas plurales en catalán antiguo, entre otras razones, obligó a generalizar el morfema /i/ como marcador de dativo, convirtiéndolo así en el verdadero elemento 'por defecto' del sistema clítico catalán. La similitud entre las formas clíticas del catalán coloquial no valenciano medieval y las del moderno nos permite analizarlas de la misma manera. En particular, sugerimos que solo existe un área clítica para estos grupos de clíticos en los que *HI* funciona como un lugar nominal ubicado estructuralmente dentro del dominio del nombre.

Palabras clave: dativo; locativo; grupos clíticos; clítico por defecto; catalán antiguo

1. Introduction

It has been argued that the 3rd person dative clitics and their combinations in standard Catalan, the variety which obeys the official prescriptive grammar and is largely limited to formal registers (henceforth SC), are faithful to the clitics that were in common use in the 15th century, the so-called 'Golden Age' of Catalan literature. What is now the standard Catalan clitic system, shown in (1), was established as the official norm by Pompeu Fabra in the first prescriptive grammar of Catalan, published in 1918 by the Institut d'Estudis Catalans (the Academy of the Catalan language), though in fact this proposal had first appeared in one of Fabra's early grammars in 1898.

(1)

	DAT.SG: li	DAT.PL: els	
ACC.M.SG: el	l'hi donaré	els el donaré	*give.*1SG.FUT
ACC.F.SG: la	la hi donaré	els la donaré	
ACC.M.PL: els	els hi donaré	els els donaré	
ACC.F.PL: les	les hi donaré	els les donaré	
ACC.N: ho	li ho donaré	els ho donaré	

This alleged faithfulness to medieval Catalan, however, has not been properly tested against the historical data. Specifically, the internal ordering and the shape adopted by the clitic groups of 3rd person DAT + ACC in Old Catalan has never been subjected to systematic scrutiny.[1] So Fabra's claims justifying the system in (1) have been merely repeated without undergoing revision.

The remainder of the paper is as follows: §2 accounts for the main dialectal differences observed in the modern Catalan clitic system; §3 reviews what has been written to date about the evolution of medieval clitics; §4 describes the results of a search of the Corpus Informatitzat de Català Antic (Digital Corpus of Old Catalan) database (henceforth CICA); §4 analyzes the data with special attention paid to the examples of plural clusters found and proposes that *HI* serves as the 'elsewhere' clitic, a non-specific item that can be used by default; and finally §6 compares medieval Catalan with the modern non-VC and VC varieties and concludes the paper.

2. Variation in the modern clitic system

With regard to the varieties of Catalan in current usage, SC differs from the colloquial Catalan as commonly spoken by the population, which can be very broadly grouped into two main varieties, the variety (or family of varieties) spoken in Catalonia proper and the Balearic Islands (what we will here call 'non-Valencian Catalan', henceforth *non-VC*) and the variety spoken within Valencia (Valencian Catalan, henceforth *VC*), as referred to in Bonet (2002).[2] When the clitic paradigms for non-VC (2) and VC (3) below are checked against SC (1) above, it can be seen that all varieties share the same singular form [li] but diverge in their plural forms. While the plural is [elz] in both SC and VC, non-VC [elzi] retains the final *-i* from the singular after the plural morph /z/, which shows that it is an independent morpheme in this variety.

---

1 In her 2002 monograph, the most extensive work to date devoted to the Old Catalan clitics, S. Fischer focuses mainly on their position within the sentence and the verbs on which they cliticize. For their part, Batllori et al. (2005) focus on enclitic vs. proclitic pronouns in Old Catalan in comparison with Old French, Spanish and Portuguese.

2 The single forms of the dative clitic in the three varieties are as follows:

non-VC:	[li]	donaré	el regal /	[əlzi]	donaré	el regal
VC:	[li]	donaré	el regal /	[elz]	donaré	el regal
SC:	li	donaré	el regal /	els	donaré	el regal
	to him/her	give.1s.Fut	the gift /	to them	give.1s.Fut	the gift

(2) non-VC   DAT.SG: [li]   DAT.PL: [əlzi]
     ACC.M.SG: el   [li]   donaré   [əlzi]   donaré   *give.*1SG.FUT
     ACC.F.SG: la   [li]   donaré   [əlzi]   donaré
     ACC.M.PL: els [əlzi]   donaré   [əlzi]   donaré
     ACC.F.PL: les [əlzi]   donaré   [əlzi]   donaré
     ACC.N: ho   [li]   donaré   [əlzi]   donaré

(3) VC   DAT.SG: [li]   DAT.PL: [e(l)s]
     ACC.M.SG: el   [lil]   donaré   [e(l)zel]   donaré
     ACC.F.SG: la   [lila]   donaré   [e(l)zla]   donaré
     ACC.M.PL: els [li(l)s]   donaré   [e(l)ze(l)s]   donaré
     ACC.F.PL: les [liles]   donaré   [e(l)zles]   donaré
     ACC.N: ho   [liw]   donaré   [e(l)zo]   donaré

The coincidence between VC and SC suggests that the two systems work similarly. However, a comparison of the 3rd person clusters in the three varieties shows that the SC pattern in (1) does not correspond to either of the two modern colloquial varieties. The non-VC system seen in (2) was already thoroughly described by Fabra in various early works (1891, 1898, 1912). The VC data in (3), from Todolí (1992), show phonetic variants with or without [l] depending on the formality of the speech. Note that the accusative forms are alike in all dialects and the syncretism seen in non-VC, which is partially preserved in SC singular forms, contrasts strongly with its absence in VC. For ease of comparison, the pronouns are shown in their proclitic forms. Though strictly speaking the phonetic forms in brackets in (2) exemplify the variety of non-VC known as Eastern Catalan, these morphological forms are in general colloquial use throughout Catalonia.[3]

A cursory glance at the clitic system of SC as set out originally by Fabra suggests that it is in effect a blend of the patterns seen in VC and non-VC, with the balance slightly favoring VC.[4] The combinations with dative singular in SC in (1) are closer to non-VC and, consistent with that dialect, the SC system also includes the locative *HI* (as in *Hi aniré demà* 'I will go there tomorrow'), which is completely absent from VC. It is also important to bear in mind that the locative *HI* and the dative clitic are mutually incompatible, though the combination is allowed in written SC (e.g., *Li hi deixaré el llibre* "I'll leave the book there for him/her').

From the data in (2) and (3) it may be concluded that the 3rd person clusters in VC are two ordered clitics of dative plus accusative without any trace of syncretism, whereas those in non-VC arise as a complete syncretic system in which it is no longer possible to identify the clitics separately, what we see instead being only a sequence of items that carry the morphosyntax features. In addition, the VC dative clitic shows allomorphy in its paradigm (*li/els*), in contrast with non-VC, where the dative clitic is made up of /l/ and /i/ with the plural /z/ located in between. These two very contrasting systems are just one example of the great variation exhibited by the dative clitic in Romance.

As is well known, in most Romance varieties dative clitics—unlike accusatives—are subject to different types of incompatibilities, substitutions and syncretism. This is

---

3 As Fabra (1891) says "to whoever writes in Catalan" in the introduction to his grammar, "I study the grammatical system of Barcelona Catalan, which is spoken with minor variations throughout Catalonia, except for some regions of Lleida and Tarragona provinces" [author's translation].

4 Fabra wanted a transparent clitic system for SC in which each clitic corresponded to just one syntactic function; indeed, he expressed regret that the Valencian combinations of the singular dative *li-l*, *li-ls*, *li-la* and *li-les* had not been preserved in non-VC (1898: 100). In fact, these groupings sound completely unacceptable in non-VC, just as the clusters with *hi* are ungrammatical in VC.

because languages can incorporate a variety of traits into the dative entity (goal, affectation, experiencer, locative, etc.) and can represent them in different ways. In fact, languages can differ not only in how they categorize this set of semantic concepts but also in how they store them in the lexicon. Therefore, grammatical categories can be expressed lexically (or not) in a variety of ways (Wiltschko 2014). Following this idea, the morphosyntactic structure of datives in the two varieties of colloquial Catalan as represented in (7) is the instantiation of different items (Cabré & Fábregas 2019). While in VC the dative clitic is a single lexical piece with allomorphy (li/<e>l+z), in non-VC the dative is made up of two items, with discrete (but related) grammatical meanings. Since they are two morphemes, the plural can appear in between (i.e., l+i/<ə>l+z+i).[5] The structural difference between the two varieties lies in the fact that in VC the dative marker /li/ occupies the DP position, while in non-VC the dative marker /i/ works as a place nominal located in the nominal layer. This syntactic structure is not relevant to the historical evolution of the dative clitic in Catalan, but it does reflect the low position the dative clitic held in Old Catalan, a position which has persisted over the centuries in non-VC.

(7)   a) Dative clitic structure in non-VC: [$_{DP}$ l [$_{NumP}$ Ø/z [$_{GenP}$ Ø [$_{NP}$ i]]]]
        b) Dative clitic structure in VC: [$_{DP}$ li/l [$_{NumP}$ Ø/z [$_{GenP}$ Ø [$_{NP}$ Ø]

Given how different these structures are in the two main varieties of Catalan, it is worth exploring how this difference came about.

3. Traditional descriptions of the Catalan clitic system and problems therewith

In his grammatical works on the contemporary language, Fabra described the clitics that occurred in Old Catalan and speculated about how they had evolved. For example, in Fabra (1898: 98), he writes: "Old Catalan has *lo li, los li, la li, les li; lo ls, los los, la ls, les los* (…). Nowadays, changing the order of pronouns, these combinations should be: *li-l, li-ls, li-la, li-les; els-el, els-els, els-la, els-les*. (...); but *li-l* has been replaced by *l'hi*, *li-ls* by *els-hi*, *li-la* by *l'hi* and *li-les* by *les-hi*, thus confusing the accusative combinations of the third-person pronoun and the dative *li* with the combinations of the accusative and the particle *hi*" [here and elsewhere, the author's translation]. As we see, Fabra refers only sporadically to combinations with plural dative, such as when he reminds us that "the combination *els-hi* often replaces the plural *els* in the spoken language" (p. 99) and adds "the combination *els-hi* tends to replace not only the dative *els* but also the four combinations *els-el, els-els, els-la, els-les*" (p. 100). Needless to say, Fabra considers the grouping *els-hi* completely unacceptable in contemporary literary language because, among other reasons, he assumes that this change took place in modern times.

As stated by Fabra (1912: 167), "the medieval clusters *loy, lay, los hi* and *les hi* have the drawback of being easily confused with clusters involving adverbial *hi*, but they are the only ones that have remained in Central Catalan." Note that these forms, exemplified in (4) and (5) from a Valencian text,[6] only refer to combinations involving dative singular.

---

5 Other proposals have been made to explain the structure of dative clitics (Rigau 1983, Mascaró 1986, Bonet 1993, 1995; Martín 2012, among others), but the split structure found in non-VC is common in all of them.
6 There seem to have been no dialectal differences in the clitic system at that time.

(4)   E Tirant   la suplicà que li   donàs la mà, que **la y**   volia besar.
And Tirant her begged that him.DAT *gave the hand, that it*.F *her*.DAT *wanted kiss*
'And Tirant begged her to give him her hand, [because] he wanted to kiss it'
Tirant lo Blanch (15th century)

(5)   Fes   -li   tu gràcies per ta part, que yo **les   hi**   sabré   fer
Do him.DAT *you thanks for your part, that I them*.F *him*.DAT *know*.FUT   *do*
per la mia.
for the mine
'Give him thanks on your own behalf, [so that] I will know how to give him thanks on mine'
Tirant lo Blach (15th century)

According to Casanova (1989: 75), "all dialects show the same order ACC + DAT and all, except for Majorcan, reversed it over the same period, to wit, from the second half of the 16th century to the middle of the 17th century, (…) except for the 3rd person cluster when the dative singular was *hi*". These claims suggest that, for both Fabra and Casanova, clitic groups with a plural dative (ACC + *els*) reversed the order without any intermediate stage.[7] This change in order is attested in clusters involving datives of 1st, 2nd and *SE* person clitics but not with 3rd person plural datives. Indeed, in these groupings *HI* (i.e., *hi* or its variants *í/y/hic*) replaced the plural dative even before the singular dative did, as is shown in (6) from a non-VC text.

(6)   E   si per aventura eyl serà   en loch que pogués   retre   aquels diners
*And i*f *by   chance he be*.FUT *in place that could return*.INF   *that money*.PL
a aquels qui comanats   **los   li**   hauran e eyl no ·**ls**
to those who entrusted them.ACC *him*.DAT *have*.Fut *and he   not them*.ACC
hi   retrà, ...
them.DAT *return*.Fut
'If by chance he is in a place where he could return that money to those who entrusted it to him but he does not do so, ...'
Llibre del Consolat de Mar (14th century).

Both Fabra's and Casanova's assertions can be justified in part by the impossibility of finding enough combinations of clitics in a wide range of texts without a database.

4. CICA database search results

The CICA is an unlabelled corpus comprised of 414 texts of various written genres, ranging from literary, administrative and legal texts to collections of letters, court books and historiographical works, representing all the dialects of Catalan and constituting in all about nine million words. The texts date from the 11th century to the first half of the 18th century and are grouped by half-century, with the first fifty years of each century labeled "A" and the second fifty labeled "B". The few texts in Catalan to have been preserved from the 11th and 12th centuries have all been incorporated into the CICA, whereas of the much more abundant textual material dating from the 18th century only one text has been included. This last text and those from of the second half of the 17th century ('17thB') are all from Valencian country. The period with the most data

---

7 Batlle et al. (2016: 243) are explicit in this regard: "Starting in the 16th and 17th centuries, all Catalan varieties made third-person binary combinations in the same way: la hi/ els la". Our data clearly contradict this assertion.

corresponds to 14thA–15thB, with more than 3,000,000 words, followed by the preceding 13thB–14thA period.

Because the CICA database is limited in its time span, we will analyze the clitic forms that appear between the 13th and 16th centuries (the limited data the CICA affords on VC from the 17th and 18th centuries do not allow any conclusions to be drawn). Therefore, the aim of this work is to show in some detail the evolution of the 3rd person clitic system in Old Catalan and to clarify the origin of the current system in non-VC. The specific contribution of this paper is to sort out what factors lie behind the dative substitution process and the spread of *HI* in medieval clitic clusters and why this morpheme remained in the isolated dative clitic. This process occurred in clusters with accusative when the dative occupied the second position, the unmarked place of the locative *HI* in combinations. Over time, replacement became generalized and spread to the dative in isolation in non-VC. Conversely, the partial replacement of the dative and the loss of the clitic *HI* in VC allowed the reordering of the clitic cluster and the generalization of the unmarked DAT + ACC order in this variety.

Before we focus on the ordering in medieval clitic clusters, several observations regarding other clitics and clitic combinations need to be made.

a) In addition to the plural dative forms included in the table (*los, lus, ·ls*), the corpus also contains a few cases of the old *lur* or *lurs*,[8] as shown in (8) below, taken from a non-VC text. No combinations with the dative *lur* are found in the database.

(8)   e   aquí ajustà   totes les osts d'Aragó   e   de   les   altres terres,
*and here gathered.*3SG *all the hosts of Aragon and of the other lands,*
e   dix   **-lurs**   que   vulia   entrar en Navarra.
*and told.*3SG *them.*DAT *that wanted.*3SG *enter in Navarre*
'And here he gathered all the troops of Aragon and the other lands, and he told them that he wanted to enter Navarre'
Crònica [B. Desclot] (13thB)

b) Instances of the plural dative *lis* are found in Valencian and Tortosan texts as well as in 16th century texts from Alghero (where it is the current plural dative), all of them as a single clitic and not in combination with an accusative clitic. The example in (9) comes from a letter from King Ferdinand II to the Bishop of Girona, signed in Tortosa in February 1496.

(9)   Pregam   e   encarregam   -vos, que (…) sens dilació,   **lis**   respongau
*ask.*PRS1PL *and instruct.*PRS1PL *you, that without delay, them.*DAT *answer.*SBJ
donant   **-lis**   plen poder (...) pasar   a   dits   apuntaments.
*giving them.*DAT *full power pass.*INF *to said commitments*
'We ask and instruct you to answer them without delay by giving them full power to fulfill these commitments'
Lletres reials a la Ciutat de Girona (15thB)

c) There are no examples of the dative plus locative combination (**LI+HI, *LOS+HI*). The incompatibility of this clitic combination persists in contemporary colloquial Catalan varieties, despite what is set out in the prescriptive norms for SC, as noted above.

---

8 For a study on the different forms of the plural dative clitic in Old Catalan, see J. Ribera (2019).

d) There are examples of *ELS+HI* pronominalizing a plural dative from the latter half of the 13th century. Most of the verbs used in these constructions correspond to intransitives or verbal forms that govern an indirect object. This is illustrated in (10) and (11) below, taken from non-VC texts.

(10)   ...si perden   diners o altres coses a joch,   lo pare o el senyor, (...),
          ...if lose.PRS3PL *money or other tings in game,* the father or the lord,
          poden e deven   tot ço que perdut y   hauran,   demanar e cobrar
          can.3PL *and should*.3PL *all this that lost there have*.3PL.FUT *ask*.INF *and cash*.INF
          sens   tot contrast e   excepció que no ·**ls y**   pot   ésser posada.
          without all contrast and exception that not them.DAT *can*.3SG *be*.INF *posed*.F.SG
          'If they lose money or other things in gambling, the father or the lord, (...) can and should ask for and get back all that they have lost, and no ruling against this or exemption may be applied to them'
          Costums de Tortosa (13[th]B)

(11)   ... e   segons   aquella tatxació   fossen   fetes   obligacions
          …and according that appraisal were.3PL *made*.F.PL *obligations*
          als   dits   mercaders de ço   que   ·**ls hi**   pertanyia.
          to-the said merchants of this that them.DAT *belonged*.3SG
          'And in accordance with that appraisal, taxes were levied on the possessions of said merchants'
          Lletres reials a la ciutat de Girona (14[th]B)

e) The clusters with *HI* include singular and plural datives as well as locatives, as exemplified in (12), taken from a non-VC text.

(12)   Déus   no   ha   posats   los reys   ne els prínceps sobre ·ls pobles ne
          God not has placed.PL *the kings nor the princes over the towns nor*
          sobre les gens (...), ans   **los**   **hi**   á   posats   axí com lo pastor
          over the people, but them.ACC *there has placed*.PL *thus like the shepherd*
          és   posat   sobre les oveles.
          *is placed.*SG *over the sheep*
          'God has not placed kings or princes over towns or over people (...), but has placed them there as the shepherd is placed over the sheep'
          Crònica [B. Desclot] (13[th]B)

f) The combination of singular dative plus neuter accusative *LI+HO* is frequently represented by *LO+HI*, as the example in (13) illustrates.[9] This data is relevant because it reflects the extent to which the *HI* form had spread in combinations with dative, since the author of the text is Valencian.

(13)   E   manà   sa senyoria a aquella virtuosa Anna que   anàs
          And commanded.3SG *his lordship to that virtuous Anna that went*.3SG.SBJ
          al seu costat, (...); e, sa altesa   tornant   **-lo**   **-y**   a   manar
          to his side and his highness return.GER *it*.N *her*.DAT PRP *command*.INF
          ella **lo**   **hi**   obeý   ab   molta   vergonya.
          she it.N *him*.DAT *obeyed*.3SG *with great shame*

---

9 As Ribera (2018) explains, the neuter accusative clitic could be represented by either *HO* or *LO*, which was also singular masculine.

'And his lordship commanded that virtuous Anna should go to his side (...) and, His Highness commanding her again, she obeyed him with great shame'
Vita Christi (15thB)

With regard to the relative ordering of clitics, the following facts deserve to be highlighted:

a) The ordering of 3rd person clitic combinations is always ACC + DAT in Old Catalan, just as in the combinations of 3rd accusative plus 1st or 2nd dative and *SE*. For this group, the shift to the unmarked reversed order began in the 16th century, while for the previous group the shift did not begin until a century later. Although the CICA database does not provide enough data for a firm conclusion, we must assume that only VC changed the order of the 3rd person combinations.

b) There are no examples in the corpus of DAT + ACC (*LI+LA*) combinations, except for three cases of *·LS+LA* in an early 17th century Valencian text, where DAT + ACC alternates with the traditional ACC + DAT order. This alternation can be seen in (14a) and (14b), both from the same work.

(14)   a. ... a 10 de febrer   1627, tornaren   ab llibertat y   ab   molta salut.
      ...on 10 of February 1627, returned.3PL *with freedom and with great health.*
      Deu **la**   **·ls**   conserve.
      God it.F *them*.DAT *keep*.3SG.SBJ
      'On February 10, 1627, they returned free and in good health. May God keep them so'

      b. ...lo dia enans   no   ·s feu   la crida (...) perquè   los comissaris
       the day before   not was made   the call   because   the officers
      de la Santa Creuada   no   **·ls**   **la**   dexaren fer.
      of the Holy Crusade not them.DAT *it*.F.SG *let*.PF *do*.INF
      '...the day before, the call was not made (...) because the officers of the Holy Crusade did not let them do it'
      Dietari [P.J. Porcar] (17thA)

5. Incompatibilities, ordering and syncretism

The results of a search of the corpus for 3rd person clitic clusters are shown in Table 1, which includes the combinations with the neuter *HO* and the partitive *NE*. Unlike with other combinations, the order within clusters of these two clitics from the very earliest texts is DAT+ACC. The data tallied here includes all proclitic and enclitic instances in all their graphical forms (e.g., LO = *lo/·l;* LOS = *los/lus/·ls;* HI = *hi/í/y/hic;* HO = *ho/ó/u/hu;* NE = *ne/·n/n'*).

Table 1. Incidence of Catalan clitic combinations with direct object and indirect object pronouns in the CICA

↓ ACC. ⟍ DAT →	LI	LOS	HI
LA	LA+LI: 251	LA+LOS^{DAT}: 37	LA+HI: 374
LO	LO+LI: 247	LO+LOS^{DAT}: 25	LO+HI: 795
LES	LES+LI: 69	LES+LOS^{DAT}: 6	LES+HI: 82
LOS	LOS+LI: 72	LOS+LOS^{DAT}: 11	LOS+HI: 242
HO	LI+HO: 241	LOS^{DAT}+HO: 211	
NE	LI+NE: 940	LOS^{DAT}+NE: 235	

One of the most surprising things about the data in Table 1 is the very small number of plural+plural combinations (*LES+LOS*: 6 / *LOS+LOS*: 11) over more than four centuries, in sharp contrast to the abundant singular + singular combinations (*LA+LI*: 251 / *LO+LI*: 247). This low incidence of plural+plural combinations cannot be accidental.

On the other hand, if we compare combinations involving just one plural element, we find a significant difference between those combinations with plural accusative and those with plural dative: there are about half as many of the latter as there are of the former (*LA+LOS*: 37/*LO+LOS*: 25 vs. *LES+LI*: 69/*LOS+LI*: 72). In addition, these combinations in plural dative contrast greatly with the large number of combinations with neuter or partitive clitics (*LOS+HO*: 211 / *LOS+NE*: 235) in which the dative occupies the first position.

It must therefore be assumed that the plural dative in those clusters is expressed in other ways. As the example in (6) above illustrates, the plural dative *LOS* changed into *HI* even before the singular *LI* did so, which leads us to suggest that the origin of this change is the incompatibility of having two plurals in the same cluster and, perhaps to a lesser extent, a resistance to having the plural dative in second place. The fact that the replacement of the plural dative occurred chronologically earlier than the replacement of the singular may explain the difference in the number of examples found of plural accusative and plural dative combinations.[10]

It is clear, on the other hand, that this replacement not only prevented the shift to DAT + ACC ordering but also produced a syncretic cluster because several combinations yielded a single result. There are therefore two separate issues to deal with here, one being the change of order within 3rd person clusters, and the other being the solutions or incompatibilities presented by the language.

The clash in the expression of the number feature seems to be the main cause of the incompatibilities we find within 3rd person clusters in other Romance languages. This is supported by the fact that we also find it in languages such as Italian that do not have specific plural morphemes, as nominal endings involve gender and number. Indeed, Italian data from 13th century texts show that "it is difficult to establish order within the clitic group either because the two pronouns have the same form (*lili, li li, gli gli, le le*) or because the invariable clitic group *li le, lile, glile, gliele* is used regardless of the gender and number of the accusative pronoun and the dative pronoun" (Cardinaletti & Egerland 2010: 444).[11] These forms are empirical evidence that this phenomenon is due not to a problem of morphophonological OCP (Obligatory Contour Principle) but rather to a problem of fitting within the morphosyntactic structure of the clitic group.

---

10 Other Romance languages, such as Romanian (Nevins & Savescu 2008), show incompatibilities in plural clusters and, in this case, it cannot be due to the plural signs since gender and number are manifested in a single morpheme.
11 See examples (180–188) in Cardinaletti and Egerland (2010: 444–45).

A closer chronological analysis will facilitate our understanding of how the replacement of the 3rd person dative by *HI* occurred. What we see in Table 2, where the incidence of clitic clusters is broken down by half centuries —following the classification of the CICA database— confirms that the plural dative changed to *HI* before the equivalent singular did so. The table omits the 11th and 12th centuries because of the scarcity of data, and likewise stops short of the 17th century because the few data in the corpus from the 17th and 18th centuries are mostly Valencian and show no change from the 16th. The letter (V) in the table indicates that all data for that combination come from Valencian texts. The table includes the dative plus neuter accusative groups in order to show the change from *LI+HO* to *LO+HI*, which persists today in most non-VC colloquial varieties.

As the CICA database shows, no dialectal differences are observed in written form for the 3rd person ACC+DAT clitic groups, and it is not until the 16th century that some divergences begin to be seen between VC and non-VC, as we saw in (14a).

Table 2. Distribution of Catalan clitic combinations in the CICA corpus, by half century.

	13th-A	13th-B	14th-A	14th-B	15th-A	15th-B	16th-A	16th-B
la li	2	42	56	61	80	**8**	0	0
la hi	0	51	5	21	25	**114**	83	36
les li	1	35	13	11	**3**	6	0	0
les hi	0	1	3	10	**16**	23	7	8
la los	0	3	8	7	7	4	1(V)	5(V)
les los	0	1	3	1	1	0	0	0
lo li	7	*40*	59	**64**	40	33	1	0
lo hi	0	*92*	15	**99**	109	288	138	43
los li	0	24	**13**	17	10	8	0	0
los hi	1	18	**29**	55	26	82	20	19
lo los	0	1	7	9	2(V)	3(V)	0	0
los los	2	1	8	0	0	0	0	0
li ho	5	42	62	89	34	2	7(V)	0
los ho	0	15	45	57	22	22	22	19
Total	18	366	326	501	375	593	279	130

In order to make it easier to identify the point in time when dative was generalized to *HI*, we have highlighted in bold the data in Table 2 that show the consolidation of the change. These data prove our earlier assertion that the replacement of the dative by *HI* occurred sooner in combinations involving a plural than in combinations involving only singular clitics. Indeed, combinations of the plural masculine and feminine accusative with *HI* outnumber combinations with *LI* half a century earlier than the analogous combinations with singular. Moreover, this change in trend begins a century earlier in combinations involving the masculine accusative than in those involving the feminine accusative. This is probably due to the fact that the dative clitic has no gender marking and takes on an unmarked masculine form from the beginning, just like the masculine accusative.

The weight of the absolute data exhibited in Table 2 is supported by Table 3 which shows the relative frequencies of each combination for each period. We have also highlighted in bold the corresponding values from the table above.

Table 3. Frequency of Catalan clitic combinations in the CICA corpus (number of each clitic combination in each period/number of total clitics found in each period).

	13th-A	13th-B	14th-A	14th-B	15th-A	15th-B	16th-A	16th-B
la li	11.11	11.48	17.18	12.18	21.33	**1.35**	0.00	0.00
la hi	0.00	13.93	1.53	4.19	6.67	**19.22**	29.75	27.69
les li	5.56	9.56	3.99	2.20	**0.80**	1.01	0.00	0.00
les hi	0.00	0.27	0.92	2.00	**4.27**	3.88	2.51	6.15
la los	0.00	0.82	2.45	1.40	1.87	0.67	0.36	3.85
les los	0.00	0.27	0.92	0.20	0.27	0.00	0.00	0.00
lo li	38.89	10.93	18.10	**12.77**	10.67	5.56	0.36	0.00
lo hi	0.00	25.14	4.60	**19.76**	29.07	48.57	49.46	33.08
los li	0.00	6.56	**3.99**	3.39	2.67	1.35	0.00	0.00
los hi	5.56	4.92	**8.90**	10.98	6.93	13.83	7.17	14.62
lo los	0.00	0.27	2.15	1.80	0.53	0.51	0.00	0.00
los los	11.11	0.27	2.45	0.00	0.00	0.00	0.00	0.00
li ho	27.78	11.48	19.02	17.76	9.07	0.34	2.51	0.00
los ho	0.00	4.10	13.80	11.38	5.87	3.71	7.89	14.62

Although Ribera (2018: 107) explicitly states that "the use of *HI* with dative value preceded by a third accusative pronoun becomes very common from the second half of the 13th century" and that "it occurs in all the dialects", this does not contradict what we are arguing here. If we pay attention to the data for *LO+LI* and *LO+HI* for the second half of the 13th century (highlighted in bold italics in Table 2) we see that the number of examples with *HI* is indeed much greater than the number of those with *LI* (and to a lesser extent greater those corresponding to feminine accusative), which may suggest that the change occurred earlier in the singular than in the plural, as has long been claimed in the literature (Fabra 1898, Casanova 1989, Batlle et al. 2016, and others). There are, however, two facts to keep in mind: first, this substitution is not consolidated until one century later in the case of the masculine and two centuries later in the case of the feminine; and second, Ribera does not compare the data for combinations in the singular with those in the plural.

As both tables show, a comparison with plural dative combinations is not possible due to the paucity of data in the corpus. However, the instances obtained show that the disappearance of the dative *LOS* in favor of *HI* became complete in the fifteenth century. This means that occurrences with *HI* include both the singular and the plural dative. The very few data from Valencian in Table 2 (5 *LO+LOS* from the 15th century and 6 *LA+LOS* from the 16th) only hint at a possible alternative path taken by this variety. It must be born in mind that replacement occurs in all the texts in the database and no dialectal differences can be detected in medieval Catalan.

The following four graphs show the evolution of the replacement of the dative clitic by *HI* in combination with the accusative clitics: *LA, LES, LO, LOS*.
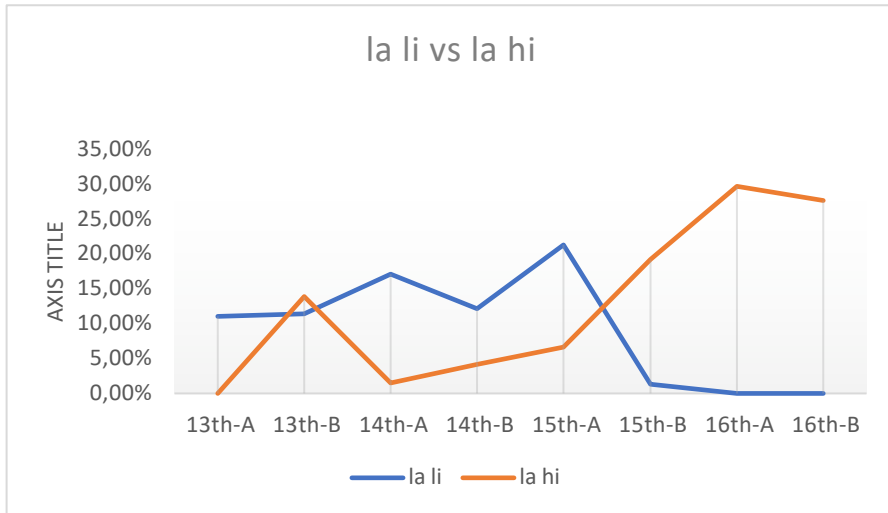
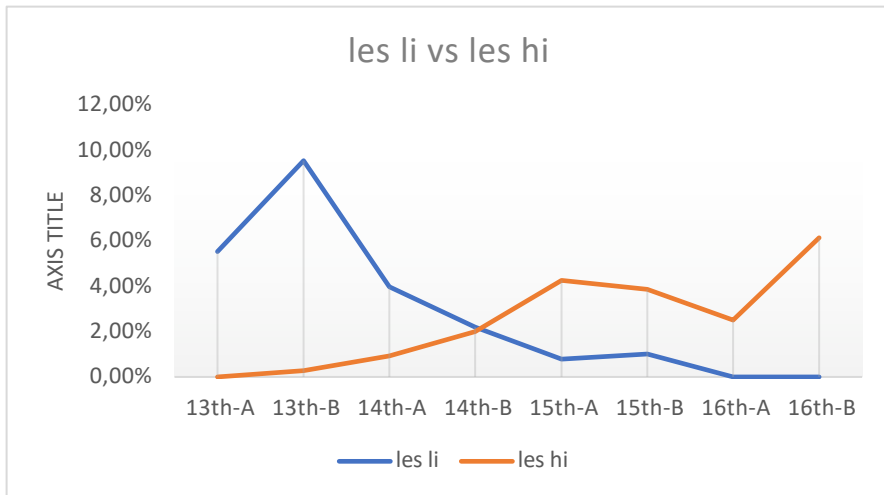*Figure 1. Evolution of LA LI compared to LA HI.*

Figure 2. Evolution of LES LI compared to LES HI.
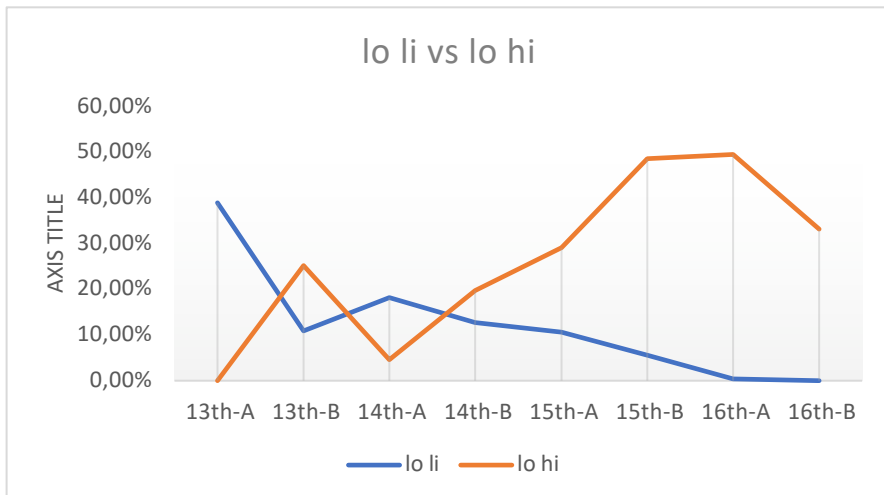

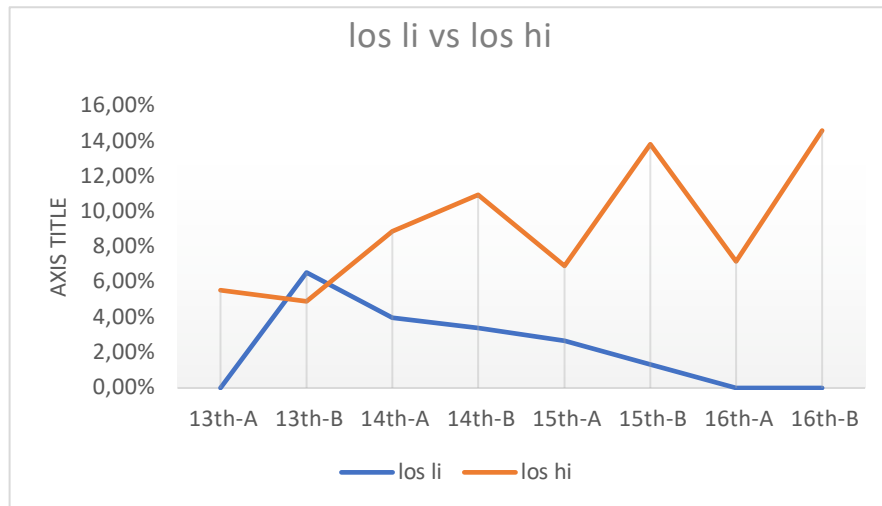*Figure 3. Evolution of LO LI compared to LO HI.*

Figure 4. evolution of LOS LI compared to LOS HI.

All this leads us to what we asserted above, to wit, that *HI* replaced the plural dative *LOS* chronologically earlier than the singular dative in 3rd person combinations, as we saw in (6). In addition, *LOS+HI* became the dative clitic governed by an intransitive verb, as shown in (10-11) above, which was generalized to all types of verbs, as can be seen in the example in (15). This is exactly the same form that is used in non-VC today.

(15) E   ell, per cascuna pedra que   gitàs   de sí,   d' aquí avant
      *And he, for   each   stone that came-out.*1SG.SBJ *from him   from here forth*
      dar   **-los-hi**   ha   una   bella   pedra   preciosa.
      give.INF *them*.DAT *have*.3SG.PRS *a   beautiful   stone   precious*
      'And henceforth, for every stone he took out of himself, he shall give them a beautiful gemstone'
      Dotzè del Crestià (15th A).

Another reason for the increase in the incidence of *LO+HI* in the 13th century is the fact that from very early on the neuter pronoun *HO* was often replaced by *LO*, as the example in (13) illustrated. As Ribera (2018: 125) notes, "the allomorph *LO* appears in Old Catalan from the beginning, in some examples without being combined with the dative, although its most frequent presence occurs in combination with the dative /i/". This means that a number of *LO+HI* examples result from the combination of a neuter accusative plus a singular dative, as shown by the example in (16), which also incorporates an isolated neuter *LO*.

(16) Emperò, puys   tant **ho** volets,   yo·**l** vos diré... Lo rey **lo y**   promés.
      *But, because so-much it want.*2SG.PRS *I it you tell*.FUT *The king it him promised*
      'But, because you want it so much, I will tell you. The king promised it to him.
      Curial e Guëlfa (15th B)[12]

Indeed, the data in Table 2 show that the *LI+HO* combination disappeared in favor of *LO+HI* in the second half of the 15th century, even before the *LO+LI* combination. It is important to emphasize that the seven examples in the corpus of *LI+HO* from the 16th century are all Valencian, the variety that has kept this combination alive. This means

---

[12] The CICA database does not ascribe this text to any specific dialect because its author is unknown.

that in Catalonia by the 15th century *LO HI* had already replaced *LI HO*, the combination that persisted in non-VC and continues in use today, as shown in (5).

Conversely, the data for the plural dative *LOS+HO* do not show any significant decline. This means that VC inherited the sub-paradigm DAT+NEUT from Old Catalan and regularized it from the plural without having lost that of the singular because, as shown in (3) above, this is the usual form in contemporary colloquial VC alongside the corresponding singular *LI+HO*, realized as [liw]. Instead, the replacement of *LOS HO* by *ELS HI* in non-VC must have taken place after the 16th century and occurred on the basis of the singular form *L'HI*.[13]

On the other hand, the neuter *HO* cannot be replaced by *LO* ACC in the cluster *LOS HO* —there is not a single example of *LOS+LO* in the whole corpus— but we do find this combination in reverse order, as shown in (17), another example from the same text.

(17) E   lo senyor de Vergues demanà en gràcia al   rey que son companyó
      And the lord   of Vergues   asked   in grace   to-the king that his   friend
      e   ell (...). Lo rey **lo**   **·ls**   atorgà
      *and he (...). The king it.*ACC *them.*DAT *granted*
      'And the lord of Vergues graciously asked the king that his friend and he (...).
      The king granted this to them'
      Curial e Güelfa (15thB)

Also striking in Table 2 is the fact that the incidence of *LOS+HI* overtook that of *LOS+LI* already in the first half of the 14th century, one century before the analogous shift with the feminine accusative (*LES+HI* vs *LES+LI*). In this regard, we must bear in mind that plural datives with *HI* had already been introduced by then, as shown in (10-11). Therefore, this piece of data corroborates what we have already argued repeatedly, that the plural feature was the trigger for the change and that the plural dative changed before the singular.

Finally, the total absence even in the earliest texts of a dative plus locative clitic combination (*LI/LOS+HI*) must be interpreted as reflecting a rejection of any combination in which a semantic feature is shared by both clitics. It seems, then, that this incompatibility allowed *HI* to be used as the exponent of dative. This phenomenon, which is present in many Romance languages, was maintained over time in non-VC. As for VC, the clitic *HI* was lost as both locative and dative by the 17th century (Casanova 1989). The CICA data do not afford much information in this regard, though they reveal some very early instances of this loss, such as in (18), taken from a Valencian text, where the verb *anar* appears without the clitic *HI*.

(18) "Vols   vengança (...) o   vols   anar   a paraís?"
      *Want.*2SG.PRS *revenge   or want.*2SG.PRS *go.*INF *to paradise*
      "Vull   anar [Ø], no vull   vengança", e llavors Jesuchrist dix…
      *want.*1SG.PRS *go.*INF *no want.*1SG.PRS *revenge   and then   JesusChrist said*
      "Do you want revenge or do you want to go to paradise?"
      "I want to go [there], I don't want revenge," and then Jesus Christ said…
      Sermons [St Vicent Ferrer] (15thA)

It seems clear that the generalization of *HI* as dative in 3rd person clitic groups during the medieval period prevented the shift from ACC+DAT to the unmarked order DAT+ACC

---

13 The phonetic evolution of *LO* and *LOS* to *EL* and *ELS* does not affect this replacement.

that occurred in many Romance varieties. However, this change in order did arise in the 16th century (except for in Majorcan Catalan) but only in groups with 1st, 2nd and *SE* pronouns plus 3rdACC, almost three centuries later than it occurred in the Italo-Romance varieties in which the shift is attested (Pescarini 2014: 156).

Last but not least, something should be said about the unmarked clitic. As revealed by the CICA data, over time the locative clitic *HI* replaced all datives of the 3rd person combinations, turning them into synthetic forms that are impossible to separate, and in the process becoming itself what Pescarini calls the true 'elsewhere' clitic, that is, "a non-specific exponent that is inserted by default" (2007: 285). This non-specificity comes from the fact that it can have different syntactic functions in different contexts. In Catalan, *HI* can act as destination, locative, goal, experiencer, affectation, etc., all features that can be expressed by means of the dative category. According to Pescarini (2007: 288), an elsewhere clitic cannot be replaced by another clitic when it collides; here we see that *HI* locative collides with LI dative (in actual fact *L+HI*) and needless to say two *HI* cannot coincide. This explains why a dative and a locative clitic never coappear, as seen in (19), although the prescriptive norms of SC in theory allow this.

(19) a) Li (l'hi)   portaré   el llibre   a la biblioteca
          him/her.DAT   bring.1SG.FUT the book   to the library

      b) Hi   portaré   el llibre
          There   bring.1SG.FUT the book

      c) *Li (l'hi)   hi   portaré
          him/her. DAT   there   bring.1SG.FUT

6. Discussion and conclusions

In this article we have tried to answer three questions: i) what factors conditioned the replacement of the dative *LI/LOS* by *HI* in clitic groups of 3rd person; ii) when and how the clitic *HI* spread and generalized in all these combinations, and iii) how and why *HI* is consolidated in the isolated dative clitic.

As we have shown, the historical shift from the old dative clitic (*LI/LOS*) to *HI* in 3rd person clusters was driven not by the dative singular, as has hitherto been claimed, but rather by the plural, since the disappearance of clusters with plural dative *LOS* took place before the disappearance of singular *LI*. Moreover, the severe scarcity of combinations with two plurals can only be interpreted as a blocking of the realization of two plural features inside the clitic area. In addition, we must take into account the semantic aspect, which is shared by the dative and the locative. Without question, these two factors are at the root of the replacement.

The spread and generalization of *HI* as DAT in clitic groups is due to its semantic specificity as opposed to the ACC. However, the ACC clitic brings the exponent of definiteness /l/ that it shares with the dative and which allows the inflectional markings to be affixed. This is also the reason why isolated datives adopt the exponent of definiteness with the clitic *HI* as an inseparable grammatical cluster. Obviously, this is only shown in the written language in the isolated plural datives since in singular the traditional written form *LI* did not make the separation *L+HI* necessary. The generalization of HI with dative value led to a reanalysis of the dative clitic, which became a complex structure.

It seems clear, then, that this is a phenomenon triggered not by a morphophonological blocking between the dative and accusative clitics—due to their similarity—but rather by

the loss of structure, which began with the loss of a number projection. The locative *HI* replaced the dative in the position it already had by default and this prevented the change of order to DAT+ACC that occurred in the combinations with 1st and 2nd person and reflective *SE* starting in the 16th century. In fact, it would be more appropriate to talk about loss of structure, given that the greater the cohesion within the cluster, the greater the degree of integration among its elements. In sum, the replacement of the dative clitic by *HI* in clusters created a morphosyntactic structure in which it was no longer possible to differentiate the two clitics. This change, consolidated in the 15th century, yielded in its turn a dative clitic which has survived to the present day in non-VC as well as in all 3rd person DAT+ACC combinations, the only observable morphosyntactic difference being the ACC gender marking in the Old Catalan form. In this fashion the dative clitic acquired the syncretic structure obtaining in current non-VC dialects, in which /i/ is situated at the nominal level (Cabré & Fábregas 2019), as shown in (7a) above and repeated in (20).

(20) $[_{DP} \, l \, [_{NumP} \, \emptyset/z \, [_{GenP} \, \emptyset \, [_{NP} \, i \,]]]]$

As for VC, the CICA data do not allow us to properly track the evolution of this phenomenon because there is a gap in the data precisely at the time of greatest change. However, we have clues as to how the current structure of the Valencian clitic paradigm has been arrived at. On the one hand, it seems that the plural dative *LOS+HI* did not spread in this dialect and, on the other, the combination *LOS+HO* persisted beyond the 16th century. In combination, this suggests that *HI* never had the same weight within the clitic paradigm in VC that it had in non-VC and therefore the two varieties gradually diverged, to the extent that *HI* disappeared in VC but spread to other combinations in non-VC.

From the CICA data, we can say, then, that VC/non-VC dialectal differences begin to be present starting in the 16th century and become increasingly widespread from that time. According to Casanova (1989:75), "since the middle of the 17th century, Valencian, having lost the use and awareness of the values of the personal adverbial pronoun *HI*, recreates a new combination of pronouns (…), forming the pair *LI LA*." That said, we should bear in mind that textual evidence of linguistic phenomena usually appears well after they begin to appear in speech and may persist even after their disappearance from the spoken idiom. However, it is clear that the contemporary colloquial non-VC has preserved the structure of the medieval clitic system in 3rd person combinations, in contrast to VC, which has retained relatively few features of Old Catalan in this regard.

The evolution of the dative clitic in Catalan clearly proves the close relationship that exists between dative and locative and shows how the locative *HI* became the "elsewhere" clitic in medieval Catalan, a fact that has been maintained and even extended in current non-VC usage. This relationship is also exhibited in clitic clusters of other Romance languages such as Logudorese Sardinian, Italian or French. The lower pressure of *HI* in the dative clusters probably led to the disappearance of the locative clitic in Valencian, as Casanova (1989) points out. Finally, it is worth noting that the evolution of the Valencian system makes it possible to predict that *HI* may also eventually disappear from colloquial Catalan under pressure from the Standard Catalan paradigm set out in (1).

Teresa Cabré
Departament of Catalan Philology
Universitat Autònoma de Barcelona
08193 Bellaterra, Spain
Teresa.cabre@uab.cat

References

Batlle, Mar; Martí Castell, Joan; Moran, Josep; Rabella, Joan Anton. 2016. *Gramàtica històrica de la llengua catalana*. Barcelona: Publicacions de l'Abadia de Montserrat.

Batllori, Montserrat, Narcís Iglésias & Ana Maria Martins. 2005. Sintaxi dels clítics *pronominals* en català medieval. Caplletra, 38: 137-177.

Bonet, Eulàlia. 2002. Cliticització. In Joan Solà et al. (eds.) *Gramàtica del català contemporani*. Vol. I: 933–989. Barcelona: Empúries.

Bonet, Eulàlia. 1995. Feature structure of Romance clitics. *Natural Language and Linguistic Theory,* 3(4): 607–647. https://doi.org/10.1007/BF00992853

Bonet, Eulàlia.1993. 3rd person pronominal clitics in dialects of Catalan. *Catalan Working Papers in Linguistics,* 3: 85–111.

Cabré, Teresa; Fábregas, Antonio. 2019. 3rd person clitic combinations across Catalan varieties: Consequences of the nature of the dative clitic. *The Linguistic Review*, 36(2): 151–190. https://doi.org/10.1515/tlr-2018-2010

Cardinaletti, Anna; Egerland, Verner. 2010. I pronomi personali e riflessivi. In Gianpaolo Salvi & Lorenzo Renzi (eds). *Grammatica dell'italiano antico*. Vol. 1. Bologna: Il Mulino.

Casanova, Emili. 1989. Aproximació a la gramàtica contrastiva dels dialectes catalans al segle XVIII: la combinació de pronoms de 3a persona. *Actes del VIII Col·lqui Internacional de Llengua i literatura catalanes*. Vol.2:57–80. Tolosa de Llenguadoc.

Corpus Informatitzat de Català Antic (CICA). 2009. Joan Torruella (dir). http://www.cica.cat

Fabra, Pompeu. 1891. *Ensayo de gramática de catalán moderno*. Barceona: L'Avenç.

Fabra, Pompeu. 1898. *Contribució a la gramàtica de la llengua catalana*. Barcelona: L'Avenç.

Fabra, Pompeu. 1912. *Gramática catalana*. Barcelona: L'Avenç.

Fabra, Pompeu. 1918. *Gramàtica catalana*. Barcelona: IEC

Fischer, Susann. 2002. *The Catalan Clitic System. A Diachronic perspective on its Syntax and Phonology*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110892505

Institut d'Estudis Catalans. 2016. *Gramàtica de la llengua catalana*. Barcelona: IEC.

Martín, Txuss. 2012. The internal structure of dative cltics. *Proceedings of the ConSOLE XVII*: 231–252. http://www.sole.leidenuniv.nl.

Mascaró, Joan. 1986. *Morfologia*. Barcelona: Enciclopèdia Catalana.

Nevins, Andrew; Ioana Savescu. 2008. An apparent 'number case constraint' in Romanian: the role of syncretism. *Selected papers from the 38 LSRL*

Pescarini, Diego. 2007. Types of syncretism in the clitic systems of Romance. *Anuario del Seminario Julio Urquijo.* ASJU, XLI.2: 285–300.

Pescarini, Diego. 2014. The evolution of Italo-Romance clitic clusters. Prosodic restructuring and morphological opacity. In P. Benincà, A. Ledgeway & N. Vincent (eds.) *Diachrony and dialects. Grammatical change in dialects of Italy*. pp. 155–176. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198701781.003.0008

Regione Autònoma de Sardigna. 2006. *Limba sarda comuna: Normas linguìsticas de referèntzia a caràtere isperimentale pro sa limba sarda iscrita de s'Amministratzione regionale*. http://www.sardegnacultura.it/documenti/7_108_20090205122432.pdf.

Ribera, Josep E. 2019. Els pronoms clítics de datiu plural en català antic. Evolució diacrònica i estat de la qüestió entre els segles XV i XVI. *Estudis Romànics*, 41: 117-145. https://doi.org/10.2436/20.2500.01.260

Ribera, Josep E. 2018. "Déu lo í perdó tot y·l tinga en la sua santa glòria". Al·lomorfia i reanàlisi en el pronom de datiu singular *li* i en el neutre *ho* en català antic. Anuari de Filologia. Estudis de Lingüística, 8: 103–135. https://doi.org/10.1344/AFEL2018.8.5

Rigau, Gemma. 1983. Inanimate indirect object in Catalan. *Linguistic Inquiry* 13(1): 146–150.

Todolí, Júlia. 1992. Variants dels pronoms febles de tercera persona al País Valencià: regles fonosintàctiques i morfològiques subjacents. *Zeitschrift für Katalanistik: Revista d'Estudis Catalans*, 5: 137–160. Frankfurt am Main.

Wiltschko, Martina. 2014. *The universal structure of categories*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139833899