Two guiding questions:
1. What is specific in Corsican DOM?
2. How can we syntactically model the à-marker in Corsican (and maybe identify its function)?

1. Introduction: Descriptive overview over DOM in Corsican

- Generally, Corsican DOM is morphologically realized by the prenominal marker à, which is homophonous with the definite feminine article singular a, and with the preposition à used to mark datives and locatives. In general, Corsican à can be considered a poly-functional marker, just like the DOM-marker in Spanish, Sicilian or Rumanian.

DOM in Corsican operates as a privative opposition, i.e. à- versus zero-marking. The category of nominals in DO-function is thus split up into two subclasses, whose members are either marked by à or not (cf. Neuburger/Stark 2014: 271f., Bossong 1982: 24).

  - animacy (ontological category, but only for pronouns, also interrogative pronouns, universal quantifiers as pronouns, negative pronouns, Giancarli 2014:200f.), kinship terms (ontological category);
  - proper names, personal pronouns, certain noun phrases (?) (grammatical category);
  - less valid: definiteness (cf. Giancarli 2014: 200);
  - But: à incompatible with (definite) articles (first systematic observation done by Marcellesi 1986, see also Giancarli 2014: 202-204, bidirectional implication, at least for animate referents, with exceptions in the domain of kinship terms, see below):

(1) “L’article exclut la préposition […]” (Damiani 1993: 28); “[…] l’article bloque d’ordinaire l’apparition de la préposition” (Chiorboli 1994: 77), cf.
What does à mark in Corsican?

2. Distribution of à

2.1 Results from Neuburger/Stark 2014

- Small corpus (about 33000 tokens) of written original Corsican texts, mainly narrative prose, but also newspaper articles and one scientific text about Corsican toponymes, all texts written between 1992 and 2005; manual analysis of 1139 DOs.

<table>
<thead>
<tr>
<th>Strong pers. pronouns</th>
<th>Proper names</th>
<th>Definite pronouns</th>
<th>Definite nominalss</th>
<th>Indefinite pronouns</th>
<th>In-defin. nominals</th>
<th>Quan.-defin. nominals</th>
<th>Bare nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+human]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>[+anim.]</td>
<td>+</td>
<td>ø</td>
<td>ø</td>
<td>–</td>
<td>–</td>
<td>ø</td>
<td>–</td>
</tr>
<tr>
<td>[+inanim.]</td>
<td>±</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 1: Overview over distribution of à with different DP categories, written data

¹ Kinship terms with a definite article are counted under definite nominals, definite kinship terms without article as proper names, indefinite kinship terms (non-specific use) as bare nouns.
² Most probably a typing error: à instead of definite article, feminine singular: a.
What does à mark in Corsican?  

<table>
<thead>
<tr>
<th>DOM with…</th>
<th>Described in the literature(^3):</th>
<th>Checked in the corpus:</th>
</tr>
</thead>
<tbody>
<tr>
<td>… full lexical nominals and proper nouns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>determiners + nominals [+ / - human/animated]</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>proper nouns [+human]</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>kinship terms</td>
<td>yes</td>
<td>yes/no</td>
</tr>
<tr>
<td>‘God’, saints</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>proper nouns [+animated] (pets)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>proper nouns [-animate] (brands, months,…)</td>
<td>yes</td>
<td>?</td>
</tr>
<tr>
<td>toponymes</td>
<td>yes</td>
<td>yes/no(^4)</td>
</tr>
<tr>
<td>metonymic/metaphorical use of nominals</td>
<td>yes</td>
<td>yes/no</td>
</tr>
<tr>
<td>bare nouns (indefinite)</td>
<td>?</td>
<td>no</td>
</tr>
<tr>
<td>… pronouns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>personal pronouns</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>- clitic personal pronouns</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>- demonstratives [+human]</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>- demonstr. [+human] + relat.clause</td>
<td>?</td>
<td>no</td>
</tr>
<tr>
<td>- interrogatives [+human]</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>adnominal relatives [+ / - human]</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Free relatives [+ human]</td>
<td>?</td>
<td>yes</td>
</tr>
<tr>
<td>Quantifiers [+ human]</td>
<td>yes</td>
<td>yes/no</td>
</tr>
<tr>
<td>- Universal ~</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>- Negative ~</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>- free-choice ~</td>
<td>?</td>
<td>no</td>
</tr>
</tbody>
</table>

yes = DOM-marked nominals | no = nominals not DOM-marked | yes/no= marked and non-marked nominals | ? = not found in the literature/in the corpus | grey marked = divergence between literature and corpus

Table 2: Distribution of à in the written corpus in comparison to previous literature

- Regular DOM-marking with strong personal pronouns, never with clitics: pan-Romance regularity:

(3)  
\[ \text{S’è tû inganni a terra, da dopu a terra inganna à te.} \]  
(Travisagna, 4)  
‘If you betray Earth, afterwards Earth will betray you’

- Almost consistent DOM-marking with proper names, independently of animacy (only two exceptions with toponyms in the written corpus, cf. Neuburger/Stark 2014: 376f.):

(4)  
\[ \text{Vinz, cù i so 600 omi, più un centu di a furtezza di Calvi, più un centu di Corsi di l’Algaiola, decide d’occupà à Calinzana.} \]  
(Cronache, 5)  
‘Vinz, with his 600 men, more than hundred from the fortress of Calvi, more than hundred Corsicans from Algaiola, decides to occupy Calinzana’

\(^3\) Albertini 1972; Bottiglioni 1933-42, 1957; Marcellesi 1986; Chiorboli 1987; Damiani 1993; Loi Corvetto/Nesi 1993; Dalbera-Stefanaggi 2002; Casta 2003; Giancarli 2014.

\(^4\) Not when used as song titles or with lexicalized definite articles.
• Almost consistent marking with definitely referring kinships terms without determiners (see above, examples under (2) from Giancarli 2014).

• Animacy relevant for pronouns: DOM-marking with universal quantifiers and negative pronouns, but not with demonstratives followed by a relative clause (see also Giancarli 2014, 206, note 13):

(5)  
*I dibattiti nant’à l’origine di i nomi di lochi sò spessu passiunati è interessanu à tutti.* (Nazione, ‘Lingua’, Nr.13, 04/08)  
‘The discussions about the origin of toponyms are often hold impulsively and interest everybody’

(6)  
*Oedippe ghjura di casticà quellu chi hà occisu à Laios.* (Mitulugia, 40)  
‘Ôdipus swears to punish the one who killed Laios’

• Differently from the data observed in Giancarli 2014 and Neuburger 2013 (see below), *unu* (‘one’, ‘someone’) and *calchissia* (‘whoever’), even when denoting human referents, are not à-marked in our written data (cf. Neuburger/Stark 2014: 379).

• No noun phrases with articles, quantifiers, numerals are DOM-marked, nor are indefinite bare nouns (Neuburger/Stark 2014: 380f.).

2.2 Field work in Corsica (cf. Neuburger 2013)

• Based on two written questionaries, *Qlab* (grammaticality judgements) and *QII* (translation Italian – Corsican):
  – Fieldwork with *QIa* was conducted in June 2010 in Corsica;
  - Fieldwork with *QIb* and *QII* was conducted in September 2011 in Corsica.

Attention: It was possible to not give an answer or propose alternatives, numbers of respondents/informants diverge between different tasks!
<table>
<thead>
<tr>
<th>Strong pers. pronouns</th>
<th>Proper names</th>
<th>Definite pronouns</th>
<th>Definite nominals(^5)</th>
<th>Indefinite pronouns</th>
<th>Indefinite nominals</th>
<th>Quan- tiﬁed nomin- als</th>
<th>Bare nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>anthro-pon., topon., kin-ship term</td>
<td>(\forall)</td>
<td>dem.</td>
<td>art.+N, art.+poss.(^6), dem.+N, (\forall)+N</td>
<td>neg. pronoun</td>
<td>others</td>
<td>indef. art.+N</td>
<td>nisciu (n) + N</td>
</tr>
</tbody>
</table>

| [+human] | \(+\) | ± | ± | \(+\) | ± | ± | ± | ± |
| Consistently marked | 49(56) | 33(55) | 11(12) | 12(12) | Consistently unmarked | 11(1) | 11(1) | Consistently unmarked | 12(3) | 10(44) |

| [+anim.] | ± | − | − | − | \(\emptyset\) | \(\emptyset\) | − | \(\emptyset\) | \(\emptyset\) |
| Consistently unmarked | 20(43) | 0(0) | 0(0) | Consistently unmarked | 0(0) | 0(0) |

| [+inanim.] | ± | − | − | − | \(\emptyset\) | \(\emptyset\) | − | \(\emptyset\) | \(\emptyset\) |
| Consistently unmarked | 65(120) | 0(0) | 0(0) | Consistently unmarked | 0(0) | 0(0) |

Table 3: Overview over distribution of à (positive grammaticality judgements and in translation task) with different DP categories, results of two fieldwork data collections: absolute numbers (yellow = main divergences from table 1)

\(^5\) Including proper names with article such as La Marana (region in Corsica).
\(^6\) Differently to Giancarli’s results, one nominal with a possessive is preferably not marked, maybe the possessive reanalyzed as D°.
\(^7\) Despite the definite article in the horse name U Sciroccu.
\(^8\) Without brand names and names for dates: years, months, days.
What does à mark in Corsican?

<table>
<thead>
<tr>
<th>Strong pers. pronouns</th>
<th>Proper names</th>
<th>Definite pronouns</th>
<th>Definite nominals</th>
<th>Indefinite pronouns</th>
<th>In-defin. nominals</th>
<th>Quan-ified nominals</th>
<th>Bare nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+human]</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>+</td>
<td>−</td>
<td>±</td>
<td>−</td>
</tr>
<tr>
<td>100%</td>
<td>87.5%</td>
<td>60%</td>
<td>+</td>
<td>19.57 % a madon na</td>
<td>91.6%</td>
<td>7%</td>
<td>100%</td>
</tr>
<tr>
<td>[+anim.]</td>
<td>±</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0</td>
<td>0</td>
<td>−</td>
</tr>
<tr>
<td>46.5%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0(0)</td>
<td>0(0)</td>
<td>0%</td>
<td>0(0)</td>
</tr>
<tr>
<td>[+inanim.]</td>
<td>±</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>0</td>
<td>0</td>
<td>−</td>
</tr>
<tr>
<td>54.16%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0(0)</td>
<td>0(0)</td>
<td>0%</td>
<td>0(0)</td>
</tr>
</tbody>
</table>

Table 4: Overview over distribution of à (positive grammatical judgements and in translation task) with different DP categories, results of two fieldwork data collections: percentages (yellow = main divergences from table 1)

- Some remarks:
  - Only some divergences to table 1, maybe due to methodology / sample examples;
  - consistent à-marking for strong personal pronouns and absence of à-marking with complex DPs (indefinite and definite) confirmed, irrespective of animacy.
  - less consistent à-marking for proper names, with anthroponyms and definite kinship terms being usually marked. Unclear picture for toponyms (7) and animal names (8):

  (7) a. *Eri ci hê statu un grande focu. U fume cupria Portivechju.* ‘Yesterday there was a big fire. The smoke covered Portivechju.’ (Neuburger 2013: 5) **11:1**
  b. *Bastia, in l’aghju mai vista prima. Devu anda ci.* ‘Bastia, (I) haven’t seen it before. (I) have to go there.’ (Neuburger 2013: 5)

---

9 Kinship terms.
I dui vanu à cavalcà: Petru s’hè coltu **U Sciroccu** è Maria s’hè colta **Bella**.

‘The two go to horseride. Petru has taken **U Sciroccu** and Maria has taken **Bella**.’

4 both with DOM; 1 both without; 2 à with **Bella**, but zero with **U Sciroccu**, 5 with à, but without **U**;

interesting comments: ‘**U** of **U Sciroccu** must be unseparable part of the name’ (1 informant); ‘no such thing as names with articles’ (1 informant) (Neuburger 2013: 4f.);

that reflects clearly a conflict of contraints: à with proper names, but not with (def.) articles!

- animacy plays a role with pronouns: universal quantifiers and demonstratives with human referents are almost always marked; the same holds for negative pronouns, also indefinite pronouns like *qualchidunu* (‘somebody’).
- Some quantified nominals are à-marked as well as some indefinite bare nominals (only human referents), in contrast to our written data:

Petru si vole marità. Ma u so amicu Marcu hè assai indecisitu: Ùn hè micca difficile à tene moglie fidele in issi tempi?

‘Petru wants to marry. But his friend Marcu is very sceptic: Isn’t it difficult to have a faithful wife in these times?’ (Neuburger 2013: 3) 10:44, maybe due to the lexical category (kinship terms when definite take à)

3. Generalization

- Application of Klein/de Swart’s (2011) model for multidimensional DOM-systems to our Corsican data (cf. Neuburger 2013: 14):

```
(10) Split I: category of DO-nominal
    [+ pro]             [- pro]
    Split II: Animacy
    [+ human]           [- human]  
    à                   *à
    Split III: case-marking
    [+ case-marked]     [- case-marked]  
    à                   *à
```

- - -
The system is based on two properties of Corsican DO-nominals: animacy (lexical element) and syntactic definiteness or DP-type plus case-marking (Neuburger in prep; i.e. not: definite reference, cf. Klein/de Swart 2011: 4 and Danon 2006 for Hebrew).

Pan-Romance comparison: Both triggers are active in Spanish and Romanian, though to a different degree/in different multiple split configurations:

Spanish: animacy > syntactic definiteness (cf. Klein/de Swart 2011: 7f.);
(Mongolian: syntactic definiteness > animacy (Klein/de Swart 2011: 9)).

Romanian: syntactic definiteness – animacy – syntactic definiteness (Klein/de Swart 2011: 10):

Common points with Romanian (Corsican being based on a Tuscan dialect belongs typologically to 'oriental Romance', cf. originally von Wartburg 1936): three splits, two based on syntactic definiteness / DP-type, same order!

Additionally: incompatibility of DOM-Marker and definite article in Romanian (cf. Stark 2008: 59):

Looks similar to Corsican:

(11) a. In America, daca încîrîiezi un apartament și ai vreo problema, trebuie să contactezi proprietarul.

b. In America, daca încîrîiezi un apartament și ai vreo problema, trebuie să contactezi pe proprietar.

Common points with Romanian (Corsican being based on a Tuscan dialect belongs typologically to 'oriental Romance', cf. originally von Wartburg 1936): three splits, two based on syntactic definiteness / DP-type, same order!

Additionally: incompatibility of DOM-Marker and definite article in Romanian (cf. Stark 2008: 59):

Looks similar to Corsican:

(12) a. In America, daca încîrîiezi un apartament și ai vreo problema, trebuie să contactezi proprietarul.

b. In America, daca încîrîiezi un apartament și ai vreo problema, trebuie să contactezi pe proprietar.

Looks similar to Corsican:

(13)

(Giancarli 2014: 209)
4. Sketch of a syntactic analysis

- Some terminology:
  1. Opposition between functional and lexical morphemes:
     functional elements convey grammatical information (e.g. gender, number, case, person, definiteness) about the function of particular types of expressions in the sentence (cf Radford 2004: 66), whereas lexical elements have descriptive content (concepts, 'world-knowledge'): the vs. dog, his, these. Additionally functional heads constitute closed classes, are generally phonologically and morphologically dependent, select only one category, are usually inseparable from their complements (= not mobile) and lack substantive content (cf. Giusti 2002: 55, following Abney 1987).
  2. Heads vs. Phrases:
     heads are minimal (often monomorphemic) expressions that may project (= take complements or adjuncts) into larger constituents, but still determine the properties of the whole phrase: help (verbal head, V°) vs. help you (VP; cf. Radford 2004: 66f.)
     Lexical heads: Adj°, N°, V° etc.
     Functional heads: D° (= articles), T° (= tense morphemes)…


(14)

Proper names may sometimes be in a SpecDP position with a non-realized definite article: Giusti 2002: 77 (parallel behavior of kinship terms); Corsican possessives and demonstratives behave like modifiers (adjectives).
• Case is hosted in the highest functional projection of a nominal and linked to def. articles in many languages (cf. German \( \text{der}_{\text{NOM}}/\text{des}_{\text{GEN}}/\text{dem}_{\text{DAT}}/\text{den}_{\text{ACC}} \text{ Mann(es)} \)):

\[
\begin{array}{c}
\text{KP} \\
\text{K}^\circ \\
\text{DP} \\
\text{D}^\circ
\end{array}
\]

• Many indefinite quantifiers like \textit{many/few/none}, numerals and the indefinite article assign \textit{partitive} case in Romance (cf. Giusti 1991: 446; see also Giancarli 2014: 200, ex. (11)), as opposed to def. articles (nominative/accusative):

\[
\begin{array}{c}
\text{(15)} \\
\text{(16)}
\end{array}
\]

\[
\begin{array}{c}
\text{QP} \\
\text{Q}^\circ \\
\text{NP} \\
\text{AdjP} \\
\text{NP} \\
\text{N}^\circ
\end{array}
\]

\[
\begin{array}{c}
\text{molli} \\
\text{antichi} \\
\text{monumenti}
\end{array}
\]

• Strong personal and interrogative pronouns are considered to be (noun) phrases (see Cardinaletti 1994 in Giusti 2002 and Giusti 2015) without accusative or dative case-marking; the same holds for pronominally used quantifiers (\textit{tutti, nimu, unu}) or demonstratives (in contrast to elitics! Cf. Giancarli 2014: 198 en bas).

• Regularity (in syntactic terms): \( \dot{\text{à}} \) incompatible with definite articles (= overt accusative case-markers) or partitive case (= indefinite nominals with \textit{many, few, numerals, indefinite article etc.}), also with incorporated bare nominals (= no verbal arguments, as such not case-marked, no KP-projection).
What does à mark in Corsican?

Stark / Neuburger

Corsican DOM-marker is an accusative case-marker.

Application to our data:

(2) \([KP \dot{a} [DP \textit{mamma} _{j} [NP t_{j}]]] (\text{=} \text{Giancarli 2014, (24)})

(3) \([KP \dot{a} [DP te_{j} [NP t_{j}]]]

(4) \([KP \dot{a} [DP \textit{Calinaza} _{j} [NP t_{j}]]]

(5) \ast \([KP \dot{a} [DP \textit{quellu} _{j} [CP \textit{chi} \ldots]]] (\text{see also discussion above around (7c) and Giancarli’s example (23); unclear selectional relation between determiners of antecedent and restrictive relative clauses: Kayne 1994, chap. 8})

(8) \([KP \dot{a} [DP U \textit{Sciroccu} _{j} [NP t_{j}]]] (U \text{ is part of the proper name and moved with it})

\ast \([KP \dot{a} [DP U [NP \textit{Sciroccu}]]] (U \text{ is in D° position taking Sciroccu as its NP complement})

(9) \ast \([KP \dot{a} [DP [NP \textit{moglie}]]], \text{if moglie incorporated: } \textit{tener moglie fidele} = \text{‘to be happily married’}

Further examples:

(17) \textit{Custi scontra trè Corsi chì avianu cunnusciatu à Teodoru, rè di Corsica.} (Cronache, 8) ‘There he meets three Corsicans, who have known Teodoru, king of Corsica.’

\ast \([KP \dot{a} [QP/\textit{NP tre} [\textit{NP Corsi}]]]

(18) \textit{Sò quistione chì cuncernanu l’umanità sana.} (Nazione, ‘Interviste’ Nr.14, 05/08) ‘These are questions which concern the whole mankind’

\ast \([KP \dot{a} [DP l’ [NP \textit{umanità}]]]

Counterexamples:

(19) \textit{Un vegu nisciun’omu.} ‘(I do) not see any man.’ (Neuburger 2013: 9; 12:38)

(20)

(23) \textit{Fala incù u fusili è tira à a mamma annant’à u chjarasgiu. Il descend avec le fusil et il tire sur la mère qui était dans le cerisier.} (CS)

(Giancarli 2014: 203)

5. Conclusion and open questions

- Distribution of Corsican DOM-marker à a puzzle at first sight, e.g. the inherent semantic feature [human] or [animate] does not permit a clear cut distinction of DOM-
marked and unmarked DO-nominals (see e.g Giancarli 2014: only 55% of human DO-referents take â), nor does semantic definiteness (cf. Neuburger in prep.).

- Main trigger seems to be syntactic definiteness in the sense of Klein/de Swart (2011), which reminds of the Romanian DOM-system.

- Analysis: â is an overt accusative marker in complementary distribution with other accusative case markers (clitic pronouns, definite articles, cf. Giusti 2015, Neuburger in prep.) and incompatible with elements assigning partitive case (indefinite quantifiers, numerals etc.).

- Open questions:
  1. Where and how does animacy come in? Or can â be found with all non-case marked pronominals?
  2. What is the role of restrictive relative clauses?

References

Albertini, Jean (1972), Precis de grammaire corse. Corte: Edt. CERC.
Bottiglioni, Gino (1933-1942), Atlante linguistico etnografico italiano della Corsica. (ALEIC), vol.I-X, Pisa: L’Italia Dialettale
Casta, Santu (2003), La syntaxe du corse, Ajaccio: CRDP Corse.


Klein, Udo / de Swart, Peter (2011), „Case and referential properties” in: Lingua 121, 3-19.


Loi Corvetto, Ines / Nesi, Annalisa (1993), La Sardegna e la Corsica, Torino: UTET.


Written Corpus (see Neuburger/Stark 2014):


