ON THE INTERNAL STRUCTURE OF PRONOUNS:
FROM THE PERSPECTIVE OF NOUN ELLIPSIS[*]

ION GIURGEA[1], RODICA IVAN[2]

Abstract. The widespread pattern of "noun-less" variants of DPs (e.g. *two/these [$_{nP/NumP}$Ø]* vs. *two/these cars*) might lead to the expectation that the definite determiner should be able to head DPs with a null complement. The best candidate for representing the structure [D_{def} [$_{nP/NumP}$Ø]] are 3rd person personal pronouns (Elbourne, 2005; a.o), an analysis supported by evidence for NP-ellipsis (or N-anaphora) in pronouns. However, equating pronouns with THE+[$_{nP/NumP}$Ø] faces a number of issues: the different behavior of pronouns and other definite descriptions with respect to binding; the differences in syntactic and semantic features between pronouns and DPs headed by THE (e.g. gender); the one-to-many structure-form relationship between the single structure THE+[$_{nP/NumP}$Ø] and different series of pronominal forms (strong and weak forms); and, a gap in the attested combinations of THE and [$_N$Ø]. Our account addresses these issues. We propose that the D used in pronouns, which we label D_{pron}, has the semantics of THE, but THE and D_{pron} differ in their formal features. We posit that a null NP complement needs to be licensed by features on the determiner the NP merges with, and that D_{pron} bears such features. Thus, we capture the fact that typical THE does not surface in cases where there is no overt element in the complement of D (i.e. what we call *total emptiness*) via the absence of these features on THE. Additionally, to further explore the differences between strong and weak forms in Romanian, we present the results of a corpus study on these forms.

Keywords pronouns, definites, ellipsis, anaphora, determiners

1. INTRODUCTION. THE PROBLEM OF EMPTY NOUNS IN DEFINITES

DPs have systematic "noun-less" variants, which often – but not always – rely on the recovery of an NP-property (nominal description) from the context ("N(P)-ellipsis" or "N(ominal)-anaphora"). Examples (1) illustrate noun ellipsis in various types of DPs, both

---

[1] The "Iorgu Iordan – Alexandru Rosetti" Institute of Linguistics of the Romanian Academy, Bucharest, giurgeaion@yahoo.com.
[2] The "Iorgu Iordan – Alexandru Rosetti" Institute of Linguistics of the Romanian Academy, Bucharest, rudmila.rodica.ivan@gmail.com.

indefinite (see (1)a) and definite (see (1)b), in Romanian. For similar English[3] examples, see the translations of the Romanian examples[4]:

(1)   a. Este nevoie de spitale. S-a   decis   să   se   construiască **unul/două/**
          is   need of hospitals REFL-has decided SBJV REFL build.3   one /two/
          câteva/multe /altele [ₙØ] în anii   următori.
          some/many / others   in years-the following
          'Hospitals are needed. It has been decided to build **one/two/some/a lot (of them)/**
           several/others [ₙØ] in the following years.'

                                                                           [ₙØ] = *spitale/ hospital(s)*
       b. Aceste şerveţele sunt mai potrivite   decât **acelea/celelalte [ₙØ]**.
          These   napkins are   more appropriate than those/ the-others
          'These napkins are more appropriate than **those/the others [ₙØ].**'

                                                                           [ₙØ] = *şerveţele/napkins*

    Nominal ellipsis may also occur in the case of bare nouns, see (2). As observed in the English version of (2)a, some languages may use an overt pro-N element, such as *ones*. In some languages, clitics also function as pro-N elements: see Fr., Cat. *en*, It. *ne*, Dutch *er*. While Romanian allows for ellipsis of the entire complement in examples like (2)b, its English version shows that some languages disallow ellipsis of the entire complement of D if this determiner is a null D. In this case, English must use *some*, which has the same semantic import as the null D which presumably underlies Romanian argumental bare nouns (see Longobardi 1994) as well as the English example *I have beer in the fridge*.

(2)   a. Se   pare că îi   plac   doar trandafirii   roşii, iar eu i-am
          REFL seems that 3S.DAT like.3PL only roses(M)-the red   and I 3S.DAT-have.1
          adus   **[ₙØ] galbeni**.
          brought   yellow.MPL
          'It seems she only likes red roses, and I brought her **yellow ones**'.

                                                                           [ₙØ]/*ones = trandafiri/roses*
       b. N-a   adus   nimeni bere. Noroc că am   eu **[ₙØ]** în frigider.
          not-has brought nobody beer fortune that have.1 I   in fridge
          'Nobody brought beer. Fortunately, I have **some [ₙØ]** in the fridge.'

                                                                           [ₙØ] = *bere/beer*

---

   [3] Other than English, our examples mostly consist of Romanian data. When a language other than English or Romanian is used in the examples, this will be flagged (e.g. (Sp.) for Spanish data).
   [4] Regarding (1)b, where a DP without an overt N is used, note that the demonstrative *acelea* 'those' obligatorily takes the 'augmented' form (with a final -*a* vs. regular *acele*). In DPs where the N is overt as in the subject of the sentence below, this augmented form is only found for post-determiner demonstratives (demonstratives that obligatorily follow N+definite inflection, see *şerveţelele*):

(i)   {Aceste şerveţele/Şerveţelele acestea}   sunt mai potrivite   decât {acelea /*acele}
            these   napkins/ napkins-the these-AUGM are   more appropriate than   those-AUGM those
The obligatory use of the augmented form with ellipsis can be analyzed either as a PF-phenomenon – a vocabulary insertion rule requiring an augmented form before [ₙØ] – or as a syntactic phenomenon. If the latter were true, it would reflect the inability of the prenominal Dem to license an empty N. See §4 below for more evidence in favor of the view that null nominal constituents need a licensing feature on the selecting functional head.

In the examples above, where ellipsis is employed, the null N has an antecedent in the linguistic context. As shown in the examples below, however, the null N can also lack an antecedent. This use is much more common in Romanian than in English, as shown by the English versions of the examples (3)a-d[5]. The descriptive content of the missing N is established based on gender features (e.g. *una* 'one.FSG') but also depending on the determiner[6]. The most widespread interpretation is +human (with the feminine explicitly indicating +female), as in (3)a-c and e. Note, however, that inanimate interpretations are also possible, as in (3)d, where the feminine occurs because it is the default value for gender in the plural (see Farkas 1980).

(3)   a. Am   văzut   **doi/ mulți**   [$_N$Ø]   care se   sărutau.
           have.1 seen   two.MPL/many.M   which REFL   were-kissing
           'I saw **two/many people** kissing.'

                                                                                         [$_N$Ø] = *people*

         b. **Fiecare [$_N$Ø]** vrea   să   fie   iubit.
            each   wants SBJV be.3 loved
            '**Every person/Everybody** wants to be loved'

                                                                                         [$_N$Ø] = *individual/person*

         c. A   intrat   [**una [[$_N$ Ø]**   foarte   beată]] .
            has   entered one.FSG   very   drunk.FSG
            'A very drunk **woman/girl** came in.'

                                                                                         [$_N$Ø] = *female person*

         d. Știe   **multe   /destule   / altele   [$_N$Ø]**.
            knows many.FPL enough.FPL/other.FPL
            '(S)he told me **many/enough/other things**.'

                                                                                         [$_N$Ø] = *things*

         e. Trebuie să   avem   mai multă încredere în **ceilalți [$_N$Ø]**
            must   SUB have.1PL more much trust   in the-others
            'We should trust **the others [$_N$Ø]** more'

                                                                                         [$_N$Ø] = *people*

Henceforth, we will use [$_N$Ø]$_{anaph}$ for null nouns which have a linguistic antecedent, like in (1)-(2), and [$_N$Ø]$_{non-anaph}$ for null nouns which do not have a linguistic antecedent in the context, like those in (3).

This systematic pattern leads to the expectation that definite Ds, in particular the definite article and the demonstrative, should occur in such "noun-less" DPs. We have already seen some examples, with demonstratives ((1)b) or with the definite article + the

---

[5] A possible account for (3)e in English could be to assume, instead of a null N, an incorporated grammatical N – or n – as reflected in the number feature of the functional item *other*.

[6] Some of these forms might be analyzed as complex D+N heads resulting from the incorporation of a grammatical N or n (cf. English *somebody*). However, the availability of this use in Romanian is too systematic, extending to whole classes such as quantitatives (cardinals as well as scalar quantitatives such as *mult, mulți* 'much, many', *destul, destui* 'enough' etc.), rendering an analysis in terms of ambiguity between D and D+N suspicious (see Giurgea 2010, 2013a).

alternative (see (1)b and (3)e;[7]). However, for the definite article (THE), we find a difference between *partial NP-emptiness* and *total NP-emptiness* in many languages.

  With respect to *partial-NP emptiness*, which we use to refer to ellipsis phenomena in which only a part of the NP is elided or null, as in (4), THE behaves by and large like other Ds. However, special forms are sometimes used because THE in Romanian is weak (affixal or clitic-like) and weak forms are not allowed before [$_N$Ø] (see Giurgea 2010). One such case is that of the use of Romanian *cea* instead of *-a* in (4)a[8] and French *celle* instead of *la* in (5). As shown by the English translation of (5), strong forms sometimes have the same form as the (distal) demonstrative, e.g. *that*. The fact that this determiner is however not identical to the demonstrative is indicated by the change in meaning between (5) and the alternative with the overt N *translation* instead of [$_N$Ø][9].

(4) a. Maşin**a** verde e mai frumoasă decât [**cea** [[$_N$Ø]$_{anaph}$ roşie]
    car(F)-the green is more beautiful than the.FSG   red.FSG
    b. **El** coche verde es más bonito que [**el** [[$_N$Ø]$_{anaph}$ rojo]   (Sp.)
    La voiture verte est plus jolie que [**la** [[$_N$Ø]$_{anaph}$ rouge]   (Fr.)
    the car green is more nice than the   red
    c. **The** green car is nicer than **the** red **one**.

                [$_N$Ø]/[*one*] = *maşină/coche/voiture/car*

(5) La traduction de l'interview   m'a pris plus de temps que
   the translation of the interview me-has taken more of time than
   [**celle [[$_N$Ø]$_{anaph}$** de l'article]]   (Fr.)
   celle   of the article
   'The translation of the interview took me longer than [**that [[$_N$Ø]** of the article]]'
               [$_N$Ø] = *traduction/translation*

A non-anaphoric null N is also possible:

(6) [context: no animate noun serving as a potential antecedent]
   Nu este indicat pentru [**cei** [[$_N$Ø]$_{non-anaph}$ cu frică de înălţime]] (Ro.)
   not is suitable for the.MPL   with fear of height
   'It is not suitable for [**those** [[$_N$Ø]$_{non-anaph}$ who are afraid of heights]].'
                [$_N$Ø] = *oameni/people*

---

  [7] In Romanian, the definite article and the alternative are fused in a single word, but they are still recognizable insofar as they both inflect, see e.g. *ce-i-l-alţ-i* 'the-MPL-*l*-other-MPL' vs. *cel-e-l-alt-e* 'the-FPL-*l*-other-FPL'. The -*l*- element intervening between the two components has a historical explanation – it comes from the former strong definite article *al*, which underlies the OldRo. definite alternative *alalt* 'the other' (see Giurgea 2012, 2013b).

  [8] Note that *cel* is also used in certain DPs with overt N, e.g. *cei doi oameni* 'the two persons'.

  [9] *The traslation of the interview took me longer than **that translation** of the article* is felicitous only if there are multiple events of translating the article, and *that translation* is contrasted to other events of translating the article. This requirement is not present in the [$_N$Ø]-variant, which has the same felicity conditions as *The traslation of the interview took me longer than **the** translation of the article*. On the use of *that* as equivalent to *the* before [$_N$Ø], see Sommerstein (1972), Schütze (2001).

However, with respect to *total emptiness*, a term we use to refer to ellipsis phenomena in which there is no overt constituent in the complement of D, as in (7), the strategies illustrated in (4)-(6) above are not available in many languages. The closest counterparts to THE+[NPØ] seem to be 3rd person personal pronouns (henceforth PRON):

(7)   A   venit cu o maşină nouă. Mi-a   spus că a cumpărat**{-o   /*cea}**
       has come with a car(F) new   me.DAT-has told that has bought-3FSG.ACC/the.FSG
       în iunie.
       in June
       'He came in a new car. He told me he bought {**it/*the one**} in June.'

A widespread account of this apparent gap in the distribution of definite article forms is that PRON (e.g. *o* and *it* in (7)) spell out structures of the type [THE [NPØ]] (Postal (1966), Panagiotidis (2002), Elbourne (2001, 2005, 2013), Sauerland (2000, 2008), Kratzer (2009), Patel-Grosz & Grosz (2017), a.o.). Independent evidence in favor of this analysis has been adduced, which we summarize in section 2. However, the assumption that the difference between PRON and THE is merely a matter of PF faces a number of problems, which we discuss in section 3.

The solution we propose in section 4 is that 3rd person pronouns do indeed contain a D with the interpretation of THE and an NP-component, but that this determiner, which we call D_pron, is distinct from the definite article (THE). The different behavior of typical definite determiners and pronouns with respect to *partial* and *total emptiness* suggests that a total identity between THE and D_pron is impossible. We propose that determiners whose entire complement is empty, i.e. *total emptiness*, must bear a feature licensing this emptiness. Specifically, we propose these determiners bear Merchant's [E] feature for an elided complement and a [+Ø] feature for the non-anaphoric null N (section 4). Under this system, we can capture the inability of typical THE determiners to license *total emptiness* via the features THE bears. Within this analysis, we can also account for the existence of series of strong and weak pronominal forms (section 5, which concentrates on Romanian). Section 6 concludes.

2. EVIDENCE FOR N-ANAPHORA IN PERSONAL PRONOUNS

2.1. Neontological pronouns

Kartunen's classical example in (8)a teaches us that there are pronouns that are neither co-referent with their antecedent nor bound by it: *it* does not refer to the same object that *his paycheque* does. Instead such pronouns are interpreted as THE+NP, where the NP is taken from the antecedent – e.g. 'the paycheck of x'[10]

(8)   a. The man who gave **his paycheque** to his wife was wiser than the man who gave **it**
          to his mistress.   (Karttunen 1969)
       b. John gave **his paycheck** to his mistress. Everybody else put **it** in the bank.
                                                                        (Cooper 1979:77)

---

[10] The antecedent, *his paycheque*, is interpreted as [THE [paycheque of him]], the possessor being part of the description that constitutes the argument of THE. For evidence that argumental DPs with possessive determiners are always definite, see Cheng et al. (2017) and references therein.

A thorough overview of this type of pronouns can be found in Elbourne (2005), who proposes the term 'neontological', alluding to the fact that the referent introduced by these pronouns is new. These pronouns are also known as 'paycheck pronouns', 'pronouns of laziness', and 'descriptive pronouns' (Partee, 1978; Cooper 1979; Evans, 1980; Haik, 1986, a.o.). In (8), the new reference ('neontological') effect is caused by the existence of a variable inside the description ('the paycheque of x'). As shown in (8)b, this inner variable can be bound by a local antecedent. The possessor that triggers disjoint reference can also be implicit. For instance, in (9), tables of contents are understood as being related to the books introduced by the quantificational DP *some* (interpreted as *some of those books*)[11].

(9)   Most books contain **a table of contents**. In some, **it** is at the end. (Heim 1990: 39)
        SOME (x a book, s a situation containing x) [in x, ιy.table-of-contents)(x)(y)(s) is at the end]

As this example shows, the situation *s* at which the description is evaluated may also differ from the situation of the antecedent[12]. There are cases where the situation variable is the only element that differs between the description inside the pronoun and the antecedent. A case in point is (10), under the assumption that temporal interpretation follows from the temporal parameter of the situation. Both situations are located in America, but the situation of the antecedent is specified as occurring this year, and that of the pronoun, next year. This example also allows a co-referential interpretation in which the president in office changes his party allegiance from one year to the next, but the most salient interpretation is the neontological one, where a different president is in office next year:

(10)   This year **the president** is a Republican. Next year **he** will be a Democrat
                                                          (Cooper 1979)

        he = ιx.presidentUSA(x)(s),   where time(s) $\subseteq$ next year

The same situation variability can explain the possibility of disjoint reference in (11). In this case, there are two different situations, each with a unique doctor: the first situation also includes Marie, the second one includes the speaker:

(11)   **Le médecin** a   interdit   à Marie de fumer. À moi, **il** ne m'a   rien   dit.
        the doctor   has   forbidden   to Marie to smoke to me   he not me-has nothing   said
                                                          (Fr., Corblin 2006: 7)
        'The doctor forbade Mary to smoke. To me, **he** didn't say anything.'

Corblin (2006) shows that pronouns can also stand for definite DPs that are parts of idioms. In this case, they completely lack reference. This is illustrated in (12)a, with the

---

[11] As an aside, note that the interpretation of *some* as *some of those books* follows from a structure [some [N∅anaph]] and a covert partitive interpretation of *some*, but could also follow from eliding the entire partitive construction (*of them/of the books*). For possible evidence for the latter alternative with certain partitive determiners, see Giurgea & Nedelcu (2009).
[12] For a comprehensive overview of the evidence for using situation variables (bound at the D-level) for accounting for domain restrictions and for the temporal and modal interpretation of definite descriptions, see Schwarz (2009).

French idiom 'catch the fly' meaning 'get upset (suddenly and unduly)'; in the second sentence in (12)a, *la* '3FS.ACC' stands for 'the fly', the definite object inside the idiom. A similar Romanian example is (12)b, with the subject DP *ţandăra* 'the splinter' resumed by *pro* as part of the idiom *a sări ţandăra*, and a similar English example is in (12)c.

(12)   a. Pierre a   pris   **la mouche**. Il **la** prend souvent pour un rien. (Corblin 2006 :8)
          Pierre has caught the fly   he it catches often   for   a nothing
          'Pierre got ticked off. He often gets ticked off for a trifle.'
    b. Lui Petru i-a   sărit   **ţandăra**.   Îi   sare *pro* adesea pentru un fleac.
        DAT Peter CL.DAT-has jumped splinter-the 3S.DAT jumps often   for   a trifle
        'Petru got ticked off. He often gets ticked off for a trifle.'
    c. Peter's **heart** skipped a bit. **It** often does so.

Another case where the referent of the pronoun differs from that of the antecedent is when the pronoun is interpreted generically and the antecedent refers to a specific individual, as in (13). These situations may be analyzed as involving a generic D + $[_N\text{Ø}]_{anaph}$ (*they* = $[D_{gen}$ $[_N\text{Ø}]_{anaph}]$ with $[_N\text{Ø}]$ = *babies*), as in Giurgea (2010); see Heim (2011) for the idea that the generic D is a sub-type of THE[13]:

(13)   A: How's baby? B: Oh, she's crying now. A: Yes, **they** do tend to cry.
                               (Huddleston & Pullum 2002: 17.2.3.[25])

2.2. The interpretation of grammatical (non-semantic) gender on pronouns

Another argument for N-anaphora in pronouns comes from the interpretation of the gender feature on these pronouns. In various languages, grammatical gender can be arbitrarily assigned to nouns, without reflecting a property of the referent. For instance, in Romanian, *cămaşă* 'shirt' is feminine, *pantof* 'shoe' is masculine and *palton* 'coat' is 'neuter' or ambigeneric (triggering masculine agreement in the singular and feminine agreement in the plural). We will call this type of gender 'non-semantic'. Pronouns that have a nominal antecedent with non-semantic gender typically show the grammatical gender of their antecedent, see (14). In Romanian, this is obligatory for inanimates:

(14)   a. Am   pus **pantoful**ᵢ   pe scaun. Peste **el**ᵢ   am   pus umbrela.
        have.1 put shoe(M)-the on chair   over 3MS.ACC have.1 put umbrella-the
        'I put the shoe on the chair. I put the umbrella over **it**.'

---

[13] Elbourne (2005) also treats donkey pronouns (e.g. *Every farmer who owns a donkey beats **it***) as neontological. However, Schwarz (2009) presents evidence that indexical (referential) anaphora is in fact involved with donkey pronouns. If so, this would require a dynamic framework in which referents introduced in the antecedent of material implication are accessible in the consequent.
    An additional issue for Elbourne's approach is posed by indistinguishable participants, illustrated by so-called 'bishop sentences' in (i). In the referential anaphor approach, the two donkey-pronouns take their reference from the list of discourse referents introduced by the antecedent:
(i)   If **a bishop** meets **another bishop**, **he** usually greets **him**.
    For an overview of the various problems for an 'e-type' analysis of donkey-anaphoric expressions, see Mandelkern & Rotschild (2020).

    b. Am   pus **cămaşaᵢ**   pe scaun. Peste **eaᵢ**   am   pus umbrela.
       have.1 put shirt(F)-the on chair over 3FS.ACC have.1 put umbrella-the
       'I put the shirt(f) on the chair. I put the umbrella over **it**.'

Note that the gender feature on the pronouns in (14) cannot be the result of agreement because the pronoun occurs in a different utterance. In fact, pronouns with grammatical gender may occur even when there is no linguistic antecedent at all. In the exophoric use, where pronouns refer to entities salient in the context, the gender reflects the nominal concept that characterizes the referent. This is exemplified in (15): the salient referent is categorized as a bill. As the words for 'bill' (*hârtie*, *bancnotă*) are feminine in Romanian, the pronoun will surface as feminine, whereas in French, where *billet* 'bill' is masculine, the pronoun will be masculine.

(15)   [Context: a bill is at the hearer's feet]
    a. Ia-**o**,   ce   mai aştepţi?
       take.IMPV.2SG-3FS.ACC what more wait.2SG
    b. Prends-**le**,   tu   hésites encore?   (Fr.)
       take.IMPV.2SG-3MS.ACC you hesitate still
       'Take it, what are you waiting for?'

As the gender feature cannot result from agreement with the antecedent and does not encode a feature of the referent, its presence can only be explained as resulting from the [NØ]anaph inside the pronominal DP. This implies that the antecedent of nominal anaphora can be provided by a salient concept in the extralinguistic context. This is a well-known property of NP-ellipsis, see Hankamer & Sag's (1976) example in (16)a, whose Romanian version in (16)b shows that gender comes from the nominal antecedent:

(16)   a. [Sag produces an apple]
        Hankamer : Did you bring **one [NØ]** for me ?
                         [NØ] =*apple*   (Hankamer & Sag : 1976 : 34)
    b. Ai   adus   [**unul** [NØ]] şi   pentru mine?
       have.2SG brought one.MSG   also for   me
                         [NØ] =*măr* 'apple': NEUT (i.e. SG → M; PL → F)

In the rare cases when number is non-semantic, being an idiosyncratic property of nouns (see lexical plurals – *pluralia tantum*), it behaves like gender, occurring on anaphoric pronouns as the result of N-anaphora:

(17)   Have you seen **my scissors**? I'm afraid I forgot **them** at home.

In the case of animates, the gender on the pronouns can be interpretable – the so-called 'natural' gender. When grammatical gender and natural gender are in conflict, a coreferent pronoun can take either of them. For instance, in (18), where the feminine *gardă* 'guard' is used to refer to male guards (as is typically the case), either the feminine *ele* 'they' or the masculine plural pronoun *ei* 'they' can be used.

(18)   A văzut **gărzile**   dar nu se   teme de **ele/ei**.
        has seen guards(F)-the but not REFL fears of they.F/they.M
        '(S)he saw the guards but is not afraid of them.'

This example also shows that pronouns do not necessarily involve N-anaphora. Under the D+[$_N$Ø] analysis, the use of natural gender indicates a structure [D+[$_N$Ø]$_{non-anaph}$].

3. ISSUES FOR EQUATING *PRON* WITH *THE*+[$_{NP}$Ø]

Despite this evidence regarding the availability of an N-component in pronoun structure, a broad assumption even in recent research on pronouns is that they generally lack an N or NP component (see Wolter 2006, Neeleman & Szendrői 2007, and Johnson 2013, Bruening 2014, who follow assumptions laid out in Schlenker (2005) based on Elbourne (2001)). One approach is to assume that there is a distinction between pronouns which may contain an N-component and pronouns that do not. This is also the stance of Jenks & Konate (2022), who propose that only neontological pronouns contain an N-component, the other ones being determiner phrases with no N (D/idxP in their system).

Indeed, a number of issues can arise from equating all instances of 3rd person pronouns (PRON) with THE+[$_{NP}$Ø]. The best-known problem concerns the different behavior of pronouns (e.g., *her*) and longer definite descriptions (e.g., *the woman*) with respect to binding (see subsection 3.1 below). In subsections 3.2-3.4, we will list additional issues, discussed in Giurgea (2010). As a preview to the following sections, in section 4 we will propose a general solution which avoids the stipulations made in Giurgea 2010. We further develop this proposal in section 5.

3.1. Binding

One immediate issue that arises from the claim that PRON have a THE+[$_{NP}$Ø] structure is the fact that pronouns differ wrt. binding/coreference from other definite descriptions which are typically associated with a THE + NP structure. While personal pronouns are subject to Principle B, DPs like *the doctor* (and even DPs including elided material, like *the other's*) are subject to Principle C of Chomsky's (1981) Binding Theory:

(19)   a. [The doctor]$_I$ said that {he$_i$ /*[the doctor]$_i$} is right.
        b. [The other]$_i$ invited a friend of {his$_i$/*[the other's]$_i$}.

Principle C, in its various instantiations (Chomsky 1981; Chomsky 1986; Büring, 2005), states that a non-pronominal referring expression cannot be locally bound. As observed in (19), *the doctor* and *the other's* are ungrammatical in the same position where a pronoun (*he, his*) can corefer with the local subject. Since principle C applies to whole DPs, if pronouns had THE+[$_{NP}$Ø] structure, additional assumptions would be required to explain the different behavior of pronouns and other DPs.

3.2. Difference in syntactic and semantic features

Under the assumption that pronouns spell-out THE+[$_{NP}$Ø], one way of capturing differences between pronouns and typical THE determiners is by appealing to morpho-

phonological constraints, i.e. at the PF level. This could be implemented, for instance, in Distributed Morphology, by using special vocabulary insertion rules for THE in the context [DP THE [Ø]], or by assuming rules of phrasal spell-out for DPs made available by the null complement of THE. However, pronouns may differ from THE in features that are relevant for syntax and semantics, which cannot be relegated to PF.

One such difference in semantic features between pronouns and typical definite determiners can be found in English and Scandinavian. English does not mark gender on THE, or on any other determiner, except for personal pronouns, where *she*, for instance, can be used to refer to female persons. However, despite the fact that gender can be marked on English pronouns, we cannot assume that pronouns like *she* necessarily presuppose gender (see Kratzer, 2009) or that they necessarily encode it syntactically, because animate pronouns also have neontological uses (as we have seen in (10) above)[14].

In Mainland Scandinavian, pronouns have singular forms reserved for animates; in Swedish and Danish, where nouns have two genders (common and neuter), pronouns are the only forms with a semantic gender distinction. This results in a system with several forms in Swedish: *han* 'he', *hon* 'she', *den* 'it.COMMON', *det* 'it.NEUTER' (Holmes & Hinchliffe, 2013). One possibility could be that interpretative differences between animate and inanimate forms reflect differences in the determiners of these pronouns. For instance, animate pronouns might include determiners with the semantics of THE but which carry an additional feature, whereas inanimate forms might use typical THE. In this case, the distribution of the two determiner types and their specialized use could be explained via competition (see Wolter 2006 for an extensive analysis of this type, relying on feature markedness).

Additionally, pronouns may also differ from other definite DPs in their distribution. Such cases include the special positions occupied by clitic forms in various languages (Cardinaletti & Starke, 1999) or the positions only available to weak pronouns (e.g. in Scandinavian-type object shift, see Holmberg 1986). For instance, in (20), the DP *la théorie* 'the theory' follows the verb, whereas the clitic *la* precedes the verb.

(20)   Je connais **la théorie**. / Je **la**   connais.
        I   know   the theory(F) I   3FS.ACC know
        'I know the theory / I know it.'

The features responsible for the special syntax of clitics must be present in narrow syntax. Consequently, at least some pronouns must be different from THE in syntax, which would be problematic for the assumption that pronouns and other definite descriptions are headed by the same THE determiner.

---

[14] Suppose one assumes that gender features originate in the null complement of pronouns, and that pronominal forms are allomorphs of THE used in the context of variously gender-marked [NØ]. In this case, one would also have to assume that these features are present on [NØ]anaph (where the is a linguistic antecedent) given the fact that animate pronouns allow neontological uses. This additional assumption is not unconceivable, but is stipulative, as there is no independent evidence that we know of for the presence of gender features in N-ellipsis contexts in English.

3.3. The weak/strong contrast: one-to-many correspondence

The existence of series of strong and clitic forms raises another issue for the identification of PRON with THE+[NPØ]. In many languages clitics and strong forms occupy different syntactic positions, as discussed in the subsection above and shown in (20). If this syntactic contrast is to be captured in the structure of weak and strong forms, then an approach which analyzes pronouns as THE+[NPØ] should make additional assumptions about how the structure of weak and strong forms differ.

The fact that strong pronominal forms have the same syntactic distribution as run-of-the-mill DPs makes strong pronouns more likely to have a THE+[NPØ] structure. But the neontological interpretations, which are indicative of a structure containing null [NØ], are sometimes restricted to *weak* forms (see Kurafuji 1998 for Japanese, Runić 2014, Bošković 2018 for Serbo-Croatian, Bi & Jenks 2019 for Mandarin)[15]. In Romanian, such a restriction surfaces in the case of objects, see (21), where the object pronoun realized as a clitic has a paycheck pronoun interpretation, namely that Dana made her own son a doctor, but the one realized as the strong pronoun *pe el* cannot refer to Dana's son[16].

(21)   Ioana l-a   făcut pe fiul   ei muzician. Dana {l-a   făcut
        Ioana CL.ACC-has made DOM son-the her musician Dana CL.3MS.ACC-has made
        doctor. /* l-a   făcut **pe el** doctor.}
        doctor   CL.ACC-has made DOM him doctor
        'Ioana made her son a musician. Dana made him (= her own son) a doctor.'

With respect to the one-to-many correspondence issue between the assumed shared structure THE+[NPØ] and the variety of pronominal forms, different solutions could be considered for different languages. For instance, in the case of languages where strong forms must be clitic-doubled whenever a clitic is available for the relevant configuration (such as Romanian and Spanish), one could think of the following solution to the one-to-many correspondence problem: clitics are agreeement markers or functional heads and the object is always a THE+[NPØ] constituent that can have an overt or null spell-out depending on constraints operating at PF; if PF well-formedness requires a constituent with a lexical accent (a prosodic word), an overt form is inserted, otherwise the form is null. It is indeed well-known that strong forms must be used in order to allow modification by focal particles, coordination, occurrence as the complement of P, or prosodic marking of information-structural features (focus, contrast; see e.g. Cardinaletti & Starke 1999). This is illustrated in (22) for the direct object position in Romanian: no matter whether the antecedent was mentioned in the same sentence or was mentioned before (thus remaining an active discourse referent), the use of a strong form is ruled out, unless it is required by the structure (e.g. modification by focal particles) or by the necessity of marking contrast. We use capitals to indicate

---

[15] In Japanese and Mandarin, weak forms are null pronouns, in Serbo-Croatian, clitics.

[16] A restriction also appears sometimes in the case of strong (i.e. overt) subject pronouns, but not always – see §5.2 below for discussion.

contrastive focus (i..e HIM), which is the most likely interpretation with this word order. In non-final positions, the contrast introduced by a strong pronoun may be weaker:

(22)   [Discourse topic: Dan$_d$]
        Ion$_i$ se   bucură fiindcă l$_{i/d}$-am   invitat {#pe el$_{i/d}$/ doar pe   el$_{i/d}$
        Ion REFL rejoices because CL.3MS.ACC-have.1 invited DOM him / only DOM him
        / şi   pe   el$_{i/d}$/pe   EL$_{i/d}$}.
        also DOM him DOM him
        'Ion is glad that I invited him/only him/him too/HIM.'

However, such an account still requires an additional featural difference between THE-DPs and pronominal DPs, to explain why clitic doubling is obligatory only for pronouns (as is the case of standard Spanish[17]). In addition, this account would not extend to languages where strong pronouns are not obligatorily clitic-doubled, e.g., Italian and French. Moreover, the difference between strong and weak forms cannot be reduced in this way not even in Romanian.

First, the sentence in (22) also allows the stressed strong form *pe el* 'DOM him' to be interpreted deictically, similar to demonstratives. In this use, illustrated in (23), the referent does not need to be contrasted with other referents, but it is new and introduced via deixis. The pointing gesture, indicated by the sign ☞, is necessary here to identify the referent:

(23)   [Extralinguistic context: a man$_m$, not previously mentioned + other persons...]
        Ion$_i$ se   bucură fiindcă l-am   invitat pe el[☞]$_m$.
        Ion REFL rejoices because CL.3MS.ACC-have.1 invited DOM him

One might claim that contrastive stress is necessary to introduce a new referent via deixis (see Grosz 2019). However, there is evidence that this type of strong forms differs from weak ones in more than stress patterns. First, notice that a paraphrase by THE+NP is infelicitous: if the intended referent is not already the most salient in the current situation and pointing is necessary to direct the hearer's attention to a salient sub-situation containing the intended referent, one would typically use the demonstrative in a DP with an overt N (see Wolter 2006):

(24)   [Extralinguistic context: a boy$_b$, not previously mentioned + other persons...]
        Ion$_i$ se   bucură fiindcă l-am   invitat {pe [băiatul acela [☞]]$_b$
        Ion REFL rejoices because CL.3MS.ACC-have.1 invited DOM boy-the that

---

[17] In Romanian, DOM-marked definites by and large require clitic doubling, but another syntactic difference between THE and pronouns pops up: inanimates with overt Ns do not take DOM. It is true that they allow DOM with partial ellipsis, but in this case DOM is sometimes optional, unlike for pronouns:
        (i)   Am   luat(-*o   pe) ultima prăjitură.
               have.1 taken(-3FS.ACC DOM) last-the cake(F)
        (ii)   Am   luat(-o   pe) ultima.   (with [$_N$Ø] = *prăjitură* 'cake')
               have.1 taken(-3FS.ACC DOM) last-the
        (iii) Am luat*(-o)   doar *(pe) ea.   (with *ea* = *prăjitura* 'the cake')
               have.1 taken(-3FS.ACC) only DOM 3FS.STRONG

/ # pe   [băiat [☞]]ᵦ }.
   DOM boy
'Ion is glad that I invited that boy/#the boy.'

Secondly, the use in (23) is only possible with animate referents. Wolter (2006) also makes this observation for English, where *he* and *she*, but not *it*, allow this use.

A more general preference for strong forms to refer to animates has been noticed for a number of languages: see Perlmutter et Oresnik (1973) for Slovenian, Jaeggli (1982) and Schroten (1992) for Spanish, and Cardinaletti & Starke (1999) for Italian, German, Slovak, Hungarian, Hebrew, Gun. According to Cardinaletti & Starke (1999), forms restricted to animates are strong not just by being non-clitic, but also by virtue of other features: stress, modification, coordination. Therefore, in addition to null and clitic pronouns, Cardinaletti & Starke (1999) introduce a category called 'weak pronouns' for word-level unstressed pronouns in argument positions (they use the label 'deficient' for what we refer here as 'weak'). The restriction of strong forms to animates is illustrated in (25):

(25)   a. Elles   sont trop grandes.   (Fr.)   (+/- human)
       they.FPL are   too big
       'They are too big.'
     b. Elles   et celles [ₙØ] d'à côté   sont trop grandes. (+human; *-human)
       they.FPL and the-ones.FPL of besides are   too big
       'They and those besides are too big.'   (Cardinaletti & Starke 1999:145, ex. 2)

Following Giurgea (2010), we suggest that the contrast in (25) is of the same type as in (24): forms restricted to animates are used to mark a lower degree of accessibility (or salience) of the antecedent (cf. Ariel 1990, Gundel et al. 1993). In (25)b the pronoun is coordinated with *celles d'à côté* 'those besides', a DP whose null N is interpreted as the null N of the pronoun (by virtue of the usual parallelism in coordination). This coordination implies that the referent of *elles* does not exhaust the entities of type N in the context. Given that there are additional possible sets of referents (expressed via *celles d'à côté* 'those besides') in the context, the referent of *elles* is less accessible. Giurgea (2010) proposes that personal pronouns are normally contrasted with demonstratives in terms of degrees of accessibility, but for animates there is an intermediate degree between weak forms (highest accessibility) and demonstratives (lowest accessibility). If a strong form is required for modification by a focal particle, without implying a lower degree of accessibility, inanimates appear to be allowed in French, as shown in the example below:

(26)   [Context: On n'analysera pas ici les multiples initiatives de niveau national, européen et international... 'We will not analyze here the multiple initiatives...']
       Elles aussi démontrent l'attention   politique consacrée depuis des   années
       they too   demonstrate the-attention political devoted since   INDEF years
       au   secteur.
       to-the sector
       'They also demonstrate the political attention devoted for years to the sector.'
       (https://www.vie-publique.fr/sites/default/files/rapport/pdf/134000641.pdf)

Admittedly, not all Cardinaletti & Starke examples can be explained in this way, and it may well be that in certain languages strong forms came to be restricted to animates (e.g., It. *lui* 'he/him', *lei* 'she/her', *loro* 'they'). In Romanian, inanimate interpretations of strong object pronouns are by and large allowed in the contexts which force the use of a strong form. This is so especially in the case of PPs (where *pro* is ungrammatical and clitics are not available), but also in other syntactic contexts, such as the contrastive topic object in (27)a and the object modified by a focal particle in (27)b. However, as we will see in §5.4, there is a preference for animates in the case of strong objects.

(27) a. Rochiile   GYA. **Pe ele**   le-am   inclus   și la „8 branduri care..."
          gowns(F)-the GYA DOM them.F CL.ACC-have.1 included also at 8 brands that....
          'The GYA gowns. Those I have also included under "8 brands that" '
                                                                        (https://finesociety.ro/...)
      b. Și cu [cărțile   lui Tudor]ᵢ ce-ai   făcut? Le-am   pierdut și **pe ele**ᵢ.
          and with books-the GEN Tudor what have.2SG done CL-have.1 lost also DOM them
          'And what did you do with Tudor's books? I lost them too.'

Moreover, the difference between strong and weak forms is also affected by whether weak forms are represented by null pronouns. For instance, in Romanian, the degree of accessibility/salience required for the null subject *pro* seems to be higher than for clitic objects. For a referent first mentioned in the preceding sentence, in a non-prominent role (e.g. not as the subject of the clause), like Kant in (28), the use of an overt subject is preferred to the use of *pro*. Example (29) shows that this preference does not appear in the choice between a clitic and a strong object (but rather goes into the opposite direction: the clitic alone is preferred). The French counterpart of (28), namely example (28)′, shows that in a non-null-subject language that has a contrast between weak and strong subject pronouns, it is the weak pronoun (*il*) that will be used in this context, given that there is no special contrast that would require the strong form *lui*.

(28) Vom   discuta acum categoriile   lui Kantᵢ. {**El**ᵢ / **?pro**ᵢ} le   obține
        will.1PL discuss now categories-the GEN Kant   he   them.F.ACC obtains
        pornind de la tipurile   de judecăți.
        starting from types-the of propositions
        'We will now discuss Kant's categories. He obtains them based on types of
          propositions.'
(29) Vom   discuta acum categoriile   lui Kantᵢ. Problema   lor   **l**-a
        will.1PL discuss now categories-the GEN Kant   problem-the their CL.ACC-has
        preocupat **(?pe el)** mult timp.
        preoccupied DOM him much time
        'We will now discuss Kant's categories. Their problem preoccupied him for a long
          time.''
(28)′ On discutera maintenant les catégories de Kant. {**Il**ᵢ /**#Lui**ᵢ} les obtient en partant des
          types de jugements.   (Fr.)

This might be correlated with the fact that the featural makeup of *pro* is less complex than that of clitics. In languages with *pro* that also have rich agreement marking, such as

Indo-European null subject languages, a typical assumption is that the features of *pro* are identified by verbal agreement (Rizzi 1986). But in most such languages, verbal agreement does not express one of the φ-features: gender[18]. The assumption that features of *pro* are identified by verbal agreement leads to the conclusion that *pro*, contrary to clitics and THE, lacks gender in these languages. Note however that the gender of *pro* can be recovered in sentences which contain a predicative adjective (e.g. Ro. *E frumoasă* 'is beautiful.FSG' = 'She is beautiful'), but such cases could be analyzed by adopting a more liberal theory of agreement, which (*contra* Chomsky 2000) allows uninterpretable features to be valued (cf. Pesetsky & Torrego 2007) and conceives agreement as feature unification. Under such an account, the assumption is that the φ-features of the predicative AP (*frumoasă* 'beautiful.FSG') must unify with those of the subject, on which they are interpreted. In this sense, we may describe the gender agreement between *pro* and a predicative adjective as a situation in which AP has valued gender and this feature comes to be assigned to the subject as a result of feature unification. A formal implementation of this idea can be found in Wechsler (2004, 2011). Alternatively, we may assume that there is a +gender variant of *pro* which is used only when it enters agreement with a predicative adjective. No matter how we analyze *pro* in these cases, we will see, in §5.3 below, that there is evidence that *pro* lacks gender in Romanian, particularly in cases where it lacks a nominal antecedent and the entity it refers to is inanimate.

To sum up, in languages with clitics and *pro*, such as Romance, there is evidence that the difference between weak and strong forms goes beyond spell-out. For a theory that equates pronouns with THE+[NPØ], this evidence raises the issue of the one-to-many correspondence between structure and forms.

3.4. A gap in the possible THE+[NØ] combinations

As discussed in §2.1, anaphoric relations between a pronoun and its antecedent can be at two different levels. One is the referential level, where coreference and bound variable relationships are established; these kinds of relationships may be subsumed under *referential anaphora*. Another type of anaphoric relationship takes place at the NP-description level, yielding neontological pronouns interpreted via *N-anaphora*: in this case, the pronoun and the antecedent need not refer to the same entity. Granting that non-semantic gender on pronouns is indicative of N-anaphora, we have a combination of referential and nominal anaphora in examples such as (14) in §2.2 above – see (14)b, repeated under (30):

(30)   Am   pus **cămaşa**ᵢ   pe scaun. Peste **ea**ᵢ   am   pus umbrela.
         have.1 put shirt(F)-the on chair   over 3FS.ACC have.1 put umbrella-the
         'I put the shirt(f) on the chair. I put the umbrella over **it**.'

We have also seen that gender on pronouns can reflect semantic gender, rather than the grammatical gender of their antecedent. One such case is that of (18), repeated below in (31), where a grammatically feminine DP *gărzile* 'guards(F)-the' can be anaphorically resumed by a masculine pronoun, *ei* 'they.M', when the referent or referents are male.

---

[18] A partial exception are languages where agreeing participles have been integrated into verbal paradigms, see Indo-Aryan. Moreover, this is irrelevant for those null-subject Indo-European languages that lost grammatical gender (Armenian, part of the Iranic languages).

(31)   A văzut **gărzile**   dar nu se   teme de **ele**/**ei**.
       has seen guards(F)-the but not REFL fears of they.F/they.M
       '(S)he saw the guards but is not afraid of them.'

Examples like (31) illustrate that pronouns which express referential anaphoric relationships may also include N-anaphoric relationships, as is the case for *ele* 'they.F' which resumes the grammatical gender of the linguistic antecedent above. At the same time, referential anaphoric pronouns need not express N-anaphoric relationships, as is the case for *ei* 'they.M', which reflects the natural gender of the referents instead. This latter type also includes pronouns whose antecedent is not categorized by a nominal concept ('anominal pronouns', see Giurgea 2010), i.e. pronouns referring to objects of the event or proposition type, introduced by CPs, VPs or other non-nominal constituents (see Wolter 2006:§5.2.2-5.2.3 for a detailed overview), illustrated in (32). Another instance is that of pronouns which refer to objects accessible via perception but not yet categorized, as in (33).

(32)   Crede [**că vom**   **câştiga**]$_i$. ***pro**$_i$* este imposibil.
       believes that will.1PL win   is   impossible
       '(S)he thinks [**we will win**]$_i$. **That**$_i$ is impossible.'
(33)   [context: an object *i* is visible in the direction of pointing]
        Ia uite-acolo! Mă   întreb   ce-o   fi ***pro**$_i$*.
        behold-there REFL ask.1SG what-may.3SG be
       'Look over there. I'm wondering what **that** is.'

We have thus far established three types of pronouns, which can be represented as in (34)a-c, where we use the following notations. Referential anaphora is represented via an *anaph* subscript holding of the entire DP (we use this label descriptively here; we discuss how referential anaphora is syntactically encoded in §5.1). The presence of N-anaphora is marked via $[_N\emptyset]_{anaph}$ and the absence of N-anaphora is notated as $[_N\emptyset]_{non\text{-}anaph}$. If the definite determiner marked by D_{def} in the combinations below is THE, then we expect a fourth combination (34)d, with both types of anaphoric relationships being absent:

(34)   a. $[D_{def} [_N\emptyset]_{anaph}]_{anaph}$: ex. (30)
       b. $[D_{def} [_N\emptyset]_{non\text{-}anaph}]_{anaph}$: ex. (31), (32), (33)
       c. $[D_{def} [_N\emptyset]_{anaph}]$: neontological pronouns (see §2.1; ex. (8)-(12), (21))
       d. $[D_{def} [_N\emptyset]_{non\text{-}anaph}]$: ??

The combination in (34)d is expected because (i) THE is not restricted to anaphoric DPs, it can introduce new referents, and (ii) the empty N can in principle be non-anaphoric, as we have seen for pronouns in this section, and for the empty N with partial emptiness in §1. In particular, for [THE $[_N\emptyset]_{anaph}$] combination, see (6), resumed under (35) below:

(35)   [context: no animate noun serving as a potential antecedent]
       Nu este indicat pentru [**cei** [[$_N\emptyset]_{non\text{-}anaph}$ cu   frică de înălţime]]
       not is   suitable for   the.MPL   with fear of height
       'It is not suitable for [those [[$_N\emptyset]_{non\text{-}anaph}$ who are afraid of heights]].'

However, the combination in (34)d seems to be missing. Possible interpretations for a pronoun reflecting this combination are: (i) 'unique person/maximal sum of persons in situation *s*' and (ii) 'unique (inanimate) entity/maximal sum of (inanimate) entities in situation *s*', where *s* can be variously specified, yielding specific or generic readings. It is plausible that the interpretations under (ii) may be ruled out for conceptual reasons – the descriptive content is too vague to identify a referent. But the +human interpretation, at least in the plural, gives perfectly reasonable readings: 'people in general', or 'the people in a contextually restricted situation', e.g. in a town, at an event etc. However, the generic interpretation is clearly ruled out, as illustrated in (36), where *pro* can only refer to some relevant and salient previously introduced set of discourse referents.

(36)   a. Acum vorbim   despre **ei**.
            now   talk.1PL about them
            'Now we're talking about **them**.'
            Impossible interpretation: 'we're talking about humans in general'
        b. ***pro*** sunt fiinţe sfâşiate de contradicţii.
               are beings torn   by contradictions
            '**They** are beings torn apart by contradictions.'
             Impossible interpretation: 'humans in general are beings torn apart by contradictions.'

When it comes to restricted situations, there is an impersonal use of the third plural which may, at first glance, correspond to the combination in (34)d, namely with the interpretation of the the maximal sum of people in situation *s*[19]:

(37)   [context: no antecedent for *pro/they*]
        Aici/În oraşul ăsta, ***pro*** nu-şi   lasă   maşinile în stradă.
        here in city-the this   not-3REFL.DAT leave.3PL cars-the   in street
        'Here/In this city, **they** (people) don't leave their cars on the street.'

However, this type of *pro* differs from a run-of-the-mill definite DP in that it has restricted anaphoric antecedent potential for subsequent pronouns. As shown below, resuming *pro* with a pronoun like *lor* in the context in (38)a is less acceptable than if *lor* were to resume a definite description like *orăşenii* 'townspeople' as in (38)b.

(38)   a. [context: no antecedent for *pro/they*]
        Aici, ***pro*i** fac   curăţenie duminica. ?? Admir   comportarea **lor**i / ??Un oraş
        here   do.3PL cleaning Sunday-the admire.1SG behavior-the their   a city
        ca al **lor**i ar   trebui să fie ţinta   noastră .
        like GEN their would must SBJV be.3 goal-the our
        'Here, they*i* clean on Sundays. ??I admire their*i* behavior / ??A city like theirs*i* should be our goal.'

---

[19] In Romanian this impersonal use is only possible with *pro*, not with strong subjects:
(i)   [context: no antecedent for the pronoun]
          # Aici/În oraşul ăsta, **ei** nu-şi   lasă   maşinile în stradă.
          here in city-the this they not-3REFL.DAT leave.3PL cars-the   in street
          Intended: 'Here/In this city, people don't leave their cars on the street.'

b. [context: no antecedent for *oamenii/people*]

Aici, ***oamenii*_i** fac   curăţenie duminica.   Admir   comportarea **lor_i** /
here people-the do.3PL cleaning Sunday-the admire.1SG behavior-the their
Un oraş ca al **lor_i** ar   trebui să fie ţinta   noastră.
A   city like GEN their would must SBJV be.3 goal-the our
'Here, people_i clean on Sundays. I admire their_i behavior / A city like theirs_i should be our goal.'

Since definite DPs with THE have non-restricted anaphoric potential, as shown above, it is unlikely that the impersonal *pro*_{pl} or *they* corresponds to the combination [THE [_NØ]_{non-anaph}]. It is possible that impersonal *pro* might differ from [THE [_NØ]_{non-anaph}] in lacking a referential index, i.e. the feature responsible for assigning a discourse referent to the DP which would allow it to function as an antecedent for referential anaphora.

4. TOWARDS AN ACCOUNT

4.1. A complex structure with a D different from THE

The arguments presented in the preceding sections point to the following conclusions: PRON may contain a [D_{def}+N] structure, where N can be anaphoric, but the D_{def} used in pronouns – which we will call D_{pron} – is formally distinct from THE. This is also the conclusion reached by Giurgea (2010). This would account for the functional equivalence between PRON and THE+[_NØ], but it nevertheless leaves the issue raised in §1 open: why, after all, is THE ruled out with *total emptiness*? That is, why does the spell-out of THE require the existence of an overt complement?

Giurgea (2010:245) proposes a principle which requires that a definite D without an explicit restriction must carry a specification of the degree of accessibility of its antecedent, (where the anaphoric relationship with the antecedent can be either can be referential or nominal). Under this approach, D_{pron} differs from THE by bearing this feature. This principle can account for the fact that the difference between strong and weak pronominal forms is sometimes correlated with the degree of accessibility (see §3.3 above) of the referent. Additionally, this assumption may also explain the absence of the combination in (34)d ([D_{def} [_NØ]_{non-anaph}]), where there is no antecedent at all. However, this principle is stipulative, merely encoding the facts, and it relies on a special feature 'degree of accessibility' which is not clearly defined.

In what follows, we propose a different solution for the formal distinction between D_{pron} and THE, relying on features for heads whose complement is empty (§4.2). Then, we sketch a way in which this analysis can account for other issues raised in §3 (see §4.3) and, in particular, for the contrast between strong and weak forms (see §5).

4.2. Features for emptiness

The expectation that a typical THE is always employed in the combinations in (34) is only justified if we take $[_N\emptyset]_{anaph}$ and $[_N\emptyset]_{non\text{-}anaph}$ to be ordinary Ns. But there already are a number of restrictions with respect to their distribution across languages which indicate that this assumption is unwarranted.

$[_N\emptyset]_{anaph}$ is a sub-case of ellipsis, which requires syntactic licensing and is associated to dedicated heads. Following Merchant (2001)'s assumptions regarding ellipsis more broadly (applied by Giurgea (2010) and Saab (2019) to nominal ellipsis), we assume that the heads F whose complement is elided carry an [E] feature – which comes with the interpretative requirement of an antecedent for the complement of F. Given that the *entire* complement of D is elided only in the case of *total emptiness*, it is only in these configurations that a definite D would bear [E]. The [E] feature is thus already a plausible point of difference between D_{pron} and THE (both with an overt N and *in partial emptiness*):

(39) a. When combining with $[_{NP}\emptyset]_{anaph}$, as in (34)a,c, D_{pron} differs from THE by bearing [E].
      b. In partial emptiness, [E] is on a lower head (Num or n)[20], so the D may be THE.

The distribution of $[_N\emptyset]_{non\text{-}anaph}$ is also constrained, which indicates that it would also require licensing. Instances of $[_N\emptyset]_{non\text{-}anaph}$ in English can be found in indefinites like *few* (e.g., *Few would disagree*), *not much* (e.g., *Not much happened*; see Huddleston & Pullum 2002:414), *some* (e.g., *Be not afraid of greatness: some are born great, some achieve greatness and some have greatness thrust upon them*, Elbourne 2013: 207). In definites with adjectives, the English $[_N\emptyset]_{non\text{-}anaph}$ is restricted to generic environments, where it has a human interpretation in the plural (e.g. *The very poor $[_N\emptyset]$ envy the rich*) and an inanimate/abstract interpretation in the singular (*This is verging on the immoral $[_N\emptyset]$*, see Huddleston & Pullum 2002: 417). An inanimate $[_N\emptyset]_{non\text{-}anaph}$ is also found with the singular demonstratives (*this, that*), for which the N-anaphoric interpretation requires the pro-N form *one* (*this one*, *that one*). However, $[_N\emptyset]_{non\text{-}anaph}$ appears to be less frequent in English than it is in Romanian.

The examples in (3), some of which are repeated below in (40), illustrate that the $[_N\emptyset]_{non\text{-}anaph}$ found in Romanian indefinites is often ruled out in English. As can be seen by comparing the Romanian examples with their English translations, English requires an overt

---

[20] Complements of N can be remnants of ellipsis, genitives as well as subcategorized PPs:
(i)   Examinarea   ipotezelor   a   fost mai   rapidă decât [cea [[$_N\emptyset$] a surselor]].
         examination-the hypotheses-the.GEN has been more fast   than the   GEN sources-the.GEN
         'The examination of the hypotheses was faster than [that [[$_N\emptyset$] of the sources]].'
(ii)   E importantă referirea   constantă la comentatori consacraţi, dar şi [cea **[[$_N\emptyset$]** la surse]].
         is important   reference-the constant to commentators established but also the   to sources
         'It is important to constantly refer to established commentators, but also to sources.'
The assumption that [E] sits on n requires that these complements may attach above n. Alternatively, they move to a position above n and do not reconstruct. In English, the overt pro-N *one* does not allow complement remnants, but such remnants may appear with the strong article form *that*, see the English version of (i). *One* may be analyzed as a Num head or a n[+Number] head, being restricted to count nouns (see Llombart-Huesca 2002, Giurgea 2010).

noun used instead, e.g. *two people* for *doi [ₙØ]*ₙₒₙ₋ₐₙₐₚₕ, *a woman/girl* for *una [ₙØ]*ₙₒₙ₋ₐₙₐₚₕ 'one.FSG', *many things* for *multe [ₙØ]*ₙₒₙ₋ₐₙₐₚₕ 'many.FPL'. Nevertheless, [ₙØ]ₙₒₙ₋ₐₙₐₚₕ is also restricted in some contexts in Romanian. For instance, it is ruled out after the indefinite determiner *nişte* 'some' (used with plural and mass nouns), as can be seen when comparing (40)c with (40)d:

(40)   a. Am   văzut   **doi/mulţi**   **[ₙØ]ₙₒₙ₋ₐₙₐₚₕ** care   se   sărutau.
         have.1 seen   two.MPL/many.M   which   REFL   were-kissing
         'I saw **two/many people** kissing.'
       b. A   intrat   [**una** [[ₙ Ø]**ₙₒₙ₋ₐₙₐₚₕ**   foarte   beată]]   .
          has   entered one.FSG   very   drunk.FSG
          'A very drunk **woman/girl** came in.'
       c. Ştie   **multe   /destule   / altele**   **[ₙØ]**ₙₒₙ₋ₐₙₐₚₕ
          knows many.FPL enough.FPL/other.FPL
          '(S)he told me **many/enough/other things**.'
       d. * Ştie   **nişte [ₙØ]ₙₒₙ₋ₐₙₐₚₕ**
            knows some
            Intended interpretation: 'She know some.'

The constraints on the distribution of [ₙØ]ₙₒₙ₋ₐₙₐₚₕ can be accounted for via a mechanism similar to the [E] feature needed for [ₙØ]ₐₙₐₚₕ. Thus, we argue that the licensing of [ₙØ]ₙₒₙ₋ₐₙₐₚₕ also requires a feature on a nominal functional head. We label this feature [+Ø].

One may wonder why the determiner should bear [+Ø] as opposed to a feature selecting an n head (possibly with no N) with a null realization, or an n incorporated in the determiner. The problem with this alternative is that it is not clear how it can account for the distinction between *partial* and *total emptiness*. Recall that THE allows [ₙØ]ₙₒₙ₋ₐₙₐₚₕ for *partial emptiness*, as in (41)a, where there is a null noun but an overt modifier. However, the [D_def [ₙØ]ₙₒₙ₋ₐₙₐₚₕ] structure is not realized for *total emptiness*, as shown in (41)b, and as we have discussed in §3.4 above. If (41)a had relied on a Num or D selecting or incorporating a specific n, we would expect this feature to occur irrespective of the presence of restrictive modifiers, which would have predicted (41)b to be grammatical.

(41)   a. Îmi   plac   **cei [ₙØ]ₙₒₙ₋ₐₙₐₚₕ** {modeşti / care nu se   tem de nimic.}
          me.DAT like.3PL the.MPL   modest   who not REFL fear of nothing
          'I like the modest / those who fear nothing.'
       b. * Îmi plac   **cei [ₙØ]ₙₒₙ₋ₐₙₐₚₕ**   /   [ei [ₙØ]ₙₒₙ₋ₐₙₐₚₕ]ₙₒₙ₋ₐₙₐₚₕ
          me.DAT like.3PL the.MPL   they.MPL
          Intended: 'I like people in general.'

Thus, we conclude that [ₙØ]ₙₒₙ₋ₐₙₐₚₕ is similar to [ₙØ]ₐₙₐₚₕ in that in involves a feature licensing emptiness of the entire complement of a head. Under this account, the difference between (41)a and b, and, more generally, the fact that THE does not occur in a DP with [ₙØ]ₙₒₙ₋ₐₙₐₚₕ with *total emptiness*, follows from the fact the determiners in the two cases cannot be identical. We argue that in the case of *total emptiness*, e.g. (41)a , the [+Ø] feature is on D, whereas in the case of *partial emptiness*, [+Ø] is on a lower head, such as Num or n.

In this sense, the two examples above differ in the choice of determiners: *total emptiness* with [$_N$Ø]$_{non-anaph}$ as in (41)a involves a D$_{+def}$ [+Ø], whereas for *partial emptiness* as in (41)b, the determiner is typical THE.

This analysis provides a formal implementation of the gap in the interpretations of personal pronouns described in §3.4 above. We capture the lack of realization of the predicted fourth combination in (34)d, [D$_{def}$ [$_N$Ø]$_{non-anaph}$], in pronouns without referential anaphoric relationships as a gap in the grammatical lexicon:

(42)   There is no D$_{def}$ bearing [+Ø] and lacking an anaphoric index[21].

In other words, we argue that any definite determiner D$_{def}$ with a [+Ø] feature also contains whatever features or hidden structure are responsible for referential anaphora (see section 5 below). This formalizes the intuition that pronouns must involve an anaphoric link, which may be referential, descriptive (nominal-anaphora) or both. That is, it is not possible for a pronoun to include *both* a non-anaphoric D$_{pron}$ *and* a non-anaphoric NP[22].

Our account based on features for emptiness, combined with the assumption that the number feature is introduced below D, makes a prediction for N-anaphoric pronouns. According to (39)a, the [E] feature is on D$_{pron}$, which requires that the entire NP complement must have an (identical) antecedent. Assuming that Num (or a Number-bearing n) are inside this complement, this predicts that number should also be identical between the antecedent and the pronoun. This prediction appears to be borne out – sloppy readings (of the paycheck-type) require the same number feature in the antecedent and the pronoun. To illustrate, see the contrast between (43)a where the antecedent and pronoun have the same number, and (43)b-c, where the number varies[23]. We tested the English sentences out with 6 native

---

[21] See §5.1 below for more detail on anaphoric indices.

[22] It should be noted that in the analysis of pronouns proposed by Elbourne (2008), the absence of pronouns lacking both N-anaphora and referential anaphora follows directly. However, this analysis has other issues, which led Elbourne (2013) to abandon it, turning back to the analysis in Elbourne (2005). In a nutshell, following insights in Nunberg (1993), Elbourne (2008) analyzes pronouns, in all their uses, as constituents of the type [D$_{def}$ [R i]], where R and i are variables that receive a value in the context: i is an individual concept (type $<s,e>$) called *index* and R is a relation mapping individual concepts to the restriction of D$_{def}$ (of type <se,<se,st>>). Direct referential uses obtain when *i* is the individual concept associated to a (salient) entity and R is identity. For neontological uses, *i* is a nominal linguistic expression and R is the interpretation function associated to NPs. In this system, the presence of the index explains why pronouns always involve an anaphoric link. Some of the issues such an account faces include: (i) R is not sufficiently constrained (so we expect a large variety of interpretations of pronouns), and (ii) in the absence of an N-component, the interpretation of grammatical gender in pronouns (see §2.2) cannot be captured. (iii) Moreover, N-anaphora is an indepedently attested phenomenon, found in DPs of all sorts (see §1 and Elbourne 2013: chapter 10). Therefore, a specialized account for N-anaphora in pronouns and definites might require further motivation.

[23] Note that forms of THE typically used with partial emptiness (i.e. *the one* in English, *cel* in Romanian) are also ungrammatical / infelicitous in these contexts. (e.g. *Mary forgot the books she had to translate. I also forgot the one*). This suggests that a [THE [$_{NumP}$ [Num+[E]] [$_{NP}$ Ø]]] structure is not allowed. For Romanian, which does not have *one*-support, as well as for the Engl. demonstrative forms, the account is straightforward: given that Num is null, D would have to bear [+Ø] in this case, but there is no D$_{def}$ bearing [+Ø] but lacking indexical anaphora (see (42)). It is not clear why *the one(s)* is ruled out in this context. Also note that it is not the case that *the one* is always blocked: it is acceptable in postcopular position (Schütze 2001).

speakers of English. While the acceptability of sloppy readings for the (43)a example varied between speakers, all six speakers agreed that (43)b-c were ungrammatical or marginal at best. For Romanian, we report our own judgments.

(43)   a. Mary forgot [the **book** she had to translate]. I also forgot **it.**
             Ro.: Maria a uitat [**cartea** pe care trebuia s-o traducă]. Şi eu am uitat-**o**.
             it/o = the book I had to translate = D_{def} [*book x had to translate*], where *x* is
             bound by the subject: Engl.: %, Ro.: ✓
        b. Mary forgot [the **book** she had to translate]. **??** I also forgot **them**.
             Ro.: Maria a uitat [**cartea** pe care trebuia s-o traducă]. ?? Şi eu **le**-am uitat.
             them/le = the books I had to translate = D_{def} [*books x had to translate*], where *x*
             is bound by the subject: Engl. *, Ro.: ??
        c. Mary forgot [the **books** she had to translate]. ?? I also forgot **it**.
             Maria a uitat [**cărţile** pe care trebuia să le traducă]. ?? Şi eu am uitat-**o**.
             it/o = the book I had to translate = D_{def} [*book x had to translate*], where *x*
             is bound by the subject: Engl.: */%??, Ro.:??

4.3. Interim conclusion

The differences between typical THE determiners and pronominal determiners with respect to ellipsis and null nouns more broadly led us to conclude that THE and D_{pron} are formally different. We proposed that one such formal difference lies in the features that D_{pron} bears to license emptiness (i.e. features which typical THE does not bear).

Under this account, the presence of *additional* semantic or formal features on D_{pron} which would further distinguish it from THE no longer poses an issue. The features discussed in §3.2 (e.g. syntactic features responsible for clitic placement, additional gender features) can be additional properties by which D_{pron} differs from THE in the lexicon.

The different behavior between pronominal and non-pronominal DPs with respect to binding (see §3.1) can also be encoded as a distinction between D_{pron} and THE. In essence, binding principles restrict the reference of DPs headed by D_{pron}, treating THE on a par with the other determiners (as an elsewhere case). A complication arises for accounts of principle C which make reference to the existence of a restriction of D, such as Schlenker (2005). We intend to address this issue in a further study.

An issue which requires further elaboration of the analysis concerns the differences between strong and weak forms. As discussed in §3.3, weak and strong forms differ with respect to their distribution (e.g. syntactic position), the anaphoric relations they hold (e.g. neontological readings are sometimes excluded for strong forms), as well as the degree of accessibility of the referent and animacy. An analysis which only relies on the [+Ø] and [E] features for D_{def} is insufficient to account for these differences. We address this in the following section. For reasons of space, we will limit our discussion to Romanian.

5. ON THE STRUCTURAL DIFFERENCES BETWEEN WEAK AND STRONG FORMS IN ROMANIAN AND THE STRUCTURES FOR REFERENTIAL ANAPHORIC PRONOUNS

5.1. Note on the representation of referential anaphora

We have seen in §3 that sometimes the difference between strong and weak forms concerns the type of anaphoric relation (i.e. referential anaphora and N-anaphora, see §3.4). As discussed in §4.2, we represent N-anaphora, i.e. N(P)-ellipsis, via the feature [E]. With respect to referential anaphora, we follow Elbourne (2005) and Schwarz (2009) in using syntactically projected *index arguments*, inside the anaphoric DP. These indices, which we call *anaphoric indices*, differ from indices informally used to mark coreference or a bound variable relations, which occur on both the antecedent and the anaphoric element. Unlike these traditional indices, they occupy a distinct position in the tree and are interpretable by being mapped to an entity via the assignment function. Furthermore, anaphoric indices resemble Reinhart's (1983) and Roelofsen (2010)'s 'binding indices' insofar they only appear on anaphoric elements (not on their antecedents) but differ from them in not being restricted to bound variable readings: anaphoric indices are also used in cases of coreference.

An explicit syntactic representation for anaphoric definites in general, using indices, was proposed by Schwarz (2009) in his analysis of West Germanic varieties that distinguish strong definite articles, which are anaphoric, from weak definite article forms, which merely encode uniqueness/maximality relative to a situation (e.g. *the sun*, *the king*)[24]. Schwarz (2009) argues that anaphoric determiners come with an additional index argument, as illustrated in (44)b, whereas weak definites, as in (44)a do not bear an index. As shown below, both weak and strong determiners take a situation argument *s* and an NP property.

(44)   a. [DP [*the*weak s] NP]

   $\llbracket \text{the}_{\text{weak}} \rrbracket = \lambda s\ \lambda P_{<e,st>}: \exists!x\ P(x)(s).\iota x.P(x)(s)$

   b. [DP i [[*the*strong s] NP]]

   $\llbracket \text{the}_{\text{strong}} \rrbracket = \lambda s\ \lambda P_{<e,st>} \lambda y: \exists!x(P(x)(s) \wedge x=y).\iota x[P(x)(s) \wedge x=y]$

The fact that situation-relativized uniqueness is sometimes insufficient to establish reference can be seen in examples such as (45). In a context where George and his colleagues are all boys, there would be no unique boy in the described situation. However, if the description 'boy in the situation x' is supplemented with 'co-referent with a salient individual', uniqueness is finally achieved:

(45)   I saw George with some colleagues at the mensa. **The boy** was exultant with joy.

Thus, anaphoric definites are not an *alternative* to unique/maximal definites, but rather are a special sub-type thereof. The difference between (44)a and (44)b, is that in the case of the latter, the description to which *iota* (i.e. uniqueness) applies is enriched via an

---

[24] Since then, evidence for special marking of anaphoric definites has been found in other languages: Akan, Korean, Mauritian Creole, Czech, Thai, Mandarin, Upper Sorbian, Ngamo, American Sign Language, Lithuanian, Icelandic, Hausa and Lakot; see Schwarz (2019) for an overview.

identity relation (with the entity picked out by the assignment function for the given index). Explicit representations in which the anaphoric relation is part of the restriction of THE were proposed by Simonenko (2014) and Hanink (2017, 2021), who include the index inside the complement of THE; in Hanink's implementation, rendered in (46) (see Hanink 2017:67), a predicate expressing identity with the index combines via Predicate Modification (intersection) with the descriptive part of the DP:

(46) $[_{DP}$ D $[_{idxP}$ idx$_{[index:i]}$ [NP]]]
       $[\![$idx$_{[index:i]}]\!]$ $^g = \lambda x.x = g(i)$
       $[\![_{idxP}$ idx$_{[index:i]}$ [NP] $]\!]$ $^g = \lambda x.$ $[\![NP]\!]$ (x) $\land$ x$=$g(i)

A different syntactic account is that of Patel-Grosz and Grosz (2017), who follow Schwarz (2009) in having two entries for THE, but instead of introducing the index as the specifier of THE, propose a functional layer above D$_{def}$, called D$_{deix}$P[25]. For other accounts where the index is above D, see Ahn (2019) and Jenks & Konate (2022).

Due to space contraints, we will not settle on a specific syntactic analysis in this paper, but we agree that anaphoric definites are syntactically different from non-anaphoric ones, involving additional structure which introduces an index. Henceforth, we employ the D[idx] notation for definite determiners with this additional structure. Recall that for pronouns we assumed a special variety of definite determiners labeled D$_{pron}$. Extending the analysis of anaphoric definites to pronouns, we will use D$_{pron}$[idx] for indexical-anaphoric pronouns and D$_{pron}$ for neontological pronouns.

5.2. Anaphoric vs. non-anaphoric pronouns: indexation & animacy

Patel-Grosz and Grosz (2017), discussing the difference between the personal pronoun *er* and the demonstrative pronoun *der* in German, propose that personal pronouns differ from demonstratives in that personal pronouns lack the additional structure encoding indexation, consisting only of D$_{def}$ and [$_{NP}$Ø]. Under this account, the reference of personal pronouns is restricted via the situation argument. We find this proposal too strong, because in contexts such as (45), where an index is needed to satisfy the uniqueness presupposition, personal pronouns can be used, including German *er* ((47)b) and Romanian *pro* ((47)c):

(47)   a. I saw George with some colleagues at the mensa. **He** was exultant with joy.
       b. Ich habe Georg mit einigen Kollegen in der Mensa gesehen. **Er** jubelte vor Freude
       c. L-am văzut pe George la cantină cu niște colegi. ***pro*** deborda de veselie.

This shows that weak pronominal forms cannot be assumed to always lack [idx].
Also recall that the assumption that [idx] is needed for *all* referential anaphoric pronouns can neatly account for the gap in the possible D$_{def}$+[$_N$Ø] combinations discussed in §3.4 – recall (42) in §4.2 above, which in this system can be rewritten as:

(48)   There is no D$_{def}$ bearing [+Ø] and lacking [idx].

---

[25] A functional projection for "strong definite articles" as a level above the DP-layer of "weak definite articles" was also proposed by Bernstein et al. (2018, 2021), who call the higher projection DP$_1$ and the lower one DP$_2$, but they do not offer semantic analyses.

Patel-Grosz and Grosz (2017) provide as evidence for their proposal some examples of pronouns with no linguistic antecedent, see (49)a,b,c, where demonstrative pronouns are not allowed. We added the Romanian translations under each of the German examples to compare the two languages. In all examples, weaker forms are listed first (e.g. *es*, *îl*) and strong forms second (e.g. *das*, *pe el*). The difference between (49)a and (49)a´ shows that if the antecedent is explicitly introduced, the demonstrative becomes possible. The Romanian versions of the examples below show that weak forms are acceptable at least for some speakers, whereas strong forms (where there is a choice) as well as demonstratives are rejected. Lastly, the Romanian version of (49)c shows that a strong form, *lui*, is allowed when no weak form is available.

(49) a. Wenn ich schwanger werde, werde ich {**es** / # **das**} auf jeden Fall behalten. (Ge.)
        if   I pregnant become will   I   it   DEM   on every case keep
        'If I get pregnant, I will definitely keep **it** (= the baby).'
        Ro.: Dacă sunt însărcinată, cu siguranţă %**îl** păstrez (***pe el/*pe acesta/*pe acela**)
     a´. Wenn ich **ein Kind** kriege, werde ich {**es** / **das**} auf jeden Fall behalten.
        if   I a child   get   will   I   it   DEM on every case keep
        'If I have a child, I will definitely keep **it**.' (Patel-Grosz & Grosz 2017: ex. 34)
     b. Hans hat so sehr geblutet, dass { **es** / * **das**} durch den Verband gedrungen ist
        Hans has so much bled   that   it   DEM   through the bandage soaked   is
        und   sein Hemd verschmutzt hat.
        and   his shirt stained   has
        'Hans bled so much that **it** (= the blood) soaked his bandages and stained his
        shirt.'   (Patel-Grosz & Grosz 2017: ex. 38a)
        Ro.: A sângerat atâta încât {%***pro*/*el/*acesta**} i-a trecut prin bandaj şi i-a pătat
            cămaşa.
     c. Manche Frauen sind schon   seit mehr als zwanzig Jahren verheiratet und
        many   women are   already for more than twenty years   married   and
        wissen noch immer nicht, was { **sein** / * **dessen**} Lieblingsbier ist.
        know   still always not   what his   DEM.GEN favorite-beer is
        'Some women have been married for more than twenty years and still do not
        know what **his** (= the husband's) favorite beer is.'
                                                    (Patel-Grosz & Grosz 2017: ex. 38b)
        Ro.: % Unele femei sunt căsătorite de mai bine de 20 de ani şi tot nu ştiu care-i
            berea **lui** preferată / *care-i berea preferată **a aceluia/ăluia/acestuia**

As these pronouns lack a linguistic antecedent, it may seem problematic to assume that there are no pronouns exempt of both indexical and nominal anaphora, as we claimed above (see (42), (48)). However, it is possible that such cases involve instances of N-anaphora, where the context provides a salient nominal concept as an antecedent. For instance, world knowledge was used for the first clause in the examples in (49) to make the concepts *child/baby* (Ge. *Kind*, Ro. *copil*), *blood* (Ge. *Blut*, Ro. *sânge*) and *husband* (Ge. *Mann*, Ro. *soţ*), respectively, prominent. We have observed that N-anaphora can take extralinguistic antecedents earlier (see (15)-(16) in §2.2). Evidence for this view comes from grammatical gender: the neuter *es* in German (49)a is probably due to the fact that the noun *Kind* 'child' is neuter, whereas Romanian uses the masculine *îl*, because *copil* 'child' is

masculine. If the only anaphoric link is indeed N-anaphora, then the restriction of these interpretations to weak forms (when there is a strong/weak choice) can be explained given the preference for neontological pronouns to be realized as weak forms (see §3.2). Thus, we conclude that an additional type of neontological pronouns should be added to those covered in §2.1: neontological pronouns with an N-antecedent not encoded by an NP, but made salient by the preceding clause.

Let us now further explore the correlation between neontological interpretations and weak forms that was noted in §3.2. First, despite (50)a and (50)b, where the use of *el* is odd or ungrammatical, it seems that neontological readings are not always ruled out for pronouns in subject position, as shown in (51):

(50) a. Anul   acesta preşedintele e un republican. La anul, {***pro/??el}*** va fi un democrat.
         year-the this   president-the is a Republican at year-the   he   will be a Democrat
         'This year the president is a Republican. Next year he will be a Democrat.'
      b. [Context: Maria and the speaker have different doctors, the speaker's doctor was
         not mentioned previously]
         Doctorul   i-a   interzis Mariei   să   fumeze. Mie, {***pro/*el***} nu
         doctor-the CL.DAT-has forbidden Maria.DAT SBJV smoke.3 me.DAT   he not
         mi-a   spus nimic.
         me.DAT-has told nothing
         'The doctor forbade Mary to smoke. To me, he didn't say anything'

(51) a. Unii   nu-şi   mai găseau   cărţile   /cartea   de identitate.
         some.PL not-3REFL.DAT more were-finding cards(F)-the card(F)-the of identity
         La mine, ***pro/ea*** stă   mereu în portofel.
         at me   3FS stays always in wallet
         'Some couldn't find their identity cards. I always keep it in the wallet.'
      b. Raportul   economic anual al firmei   noastre abia   dacă are două
         report(N)-the economic anual GEN company-the.GEN our   hardly if   has two
         pagini. La altele, ***pro/el*** poate să   ajungă până la zeci de pagini.
         pages at others   3MS can SBJV reach.3 up   to ten of pages
         'The annual economic report of our company is barely 2 pages long. In the case
         of other companies, it can reach tens of pages.'

We suggest that the difference between these sets of examples is in fact related to animacy: note that (50) involves animates and (51), inanimates. In the case of animates, gender-marked forms tend to be interpreted as involving natural gender, i.e. $[_N\emptyset]_{\text{non-anaph}}$, which excludes an N-anaphoric interpretation.

Above, we have discussed neontological pronouns in subject position, where neontological readings of strong pronouns tend to be excluded for animates. In object position, neontological readings are also excluded for inanimates, as shown in (52). This example uses focal particle modification (i.e. *şi* 'also') to ensure that the pronoun is not ruled out just because of its inanimate interpretation (for acceptable inanimate strong objects with focal particles, see (27)b):

(52) Eu am   rămas   măcar cu cartea de identitate. * Ceilalţi   au
      I   have.1 remained at-least with card-the of identity   others-the have

pierdut-o   **şi   pe   ea** / ??Ceilalţi   le-au   pierdut **şi   pe   ele** /
lost-CL.ACC also DOM 3FS /   others-the CL.ACC-have lost   also DOM 3FS
*Maria a   pierdut-o   **şi   pe   ea**.
Maria   has lost-CL.ACC alsos DOM FS
'At least I kept my identity card. The others lost them too/Maria lost it too.'

Lastly, in the case of PPs, where there is no competition between strong and weak forms (there are no weak forms) in Romanian, neontological readings are allowed:

(53)   Eu mi-am   pus **geanta** pe umăr.   Ceilalţi   au   venit **cu   ea**
        I   me.DAT-have.1 put bag(F)-the on shoulder the-others have come with 3SF.STRONG
        în mână.
         in hand
        'I put the bag on my shoulder. The others came with it in their hand.'
        ea 'it' = their bag: [$_N$Ø]$_{anaph}$ = *bag of x*, where *x* is bound by the subject

The data discussed in this subsection illustrate that neontological readings are available for strong pronouns in the following cases: (i) for strong pronouns in subject position referring to inanimate referents, e.g., (51); and (ii) when weak forms are not available, like in PP object positions, e.g., (53). The fact that strong pronouns do not have neontological readings in object position can be captured as:

(54)   Object strong forms are restricted to D$_{pron}$[idx]

5.3. *Pro* vs overt subject pronouns

(54) indicates that strong object forms cannot be treated completely on a par with strong subject forms. Another difference between the two was illustrated in section 3.3: in (28)-(29), repeated below, we see that overt pronouns are preferred in subject position if their antecedent is not clearly the topic (*Kant's categories* is the actual topic, not *Kant*). Note that this preference goes in the opposite direction for strong pronouns in object position.

(28)   Vom   discuta acum categoriile   lui Kant$_i$. {**El**$_i$ / *?pro*$_i$} le   obţine
        will.1PL discuss now   categories-the GEN Kant   he   them.F.ACC obtains
        pornind de la tipurile   de judecăţi.
        starting from types-the of propositions
        'We will now discuss Kant's categories. He obtains them based on types of
          propositions.'
(29)   Vom   discuta acum categoriile   lui Kant$_i$. Problema   lor   **l$_i$**-a
        will.1PL discuss now   categories-the GEN Kant   problem-the their CL.ACC-has
        preocupat **(?pe el$_i$)** mult timp.
        preoccupied   DOM him much time
        'We will now discuss Kant's categories. Their problem preoccupied him for a long
          time.''

As observed above, the requirement of using an overt pronoun in subject position appears to be related to the likelihood of that newly introduced entity to become a topic. For instance, in (55), the new referent has been introduced as the object of *see* and we expect an elaboration of the scene introduced by this perception verb, where the only entity is Daniel. In (28), the more prominent new referent was *categoriile* 'the categories'.

(55)   Tata   mi-a   spus că   l-a   văzut pe Daniel$_d$. ***pro*$_d$** era cu bicicleta.
        father-the me.DAT-has told that CL.ACC-has seen DOM Daniel   was with bike-the
        'Dad told me he had seen Daniel. He was riding his bike.'

We may characterize entities likely to be topics in a certain sentence by the concept of 'G(iven)-topic' proposed by Frascarelli & Hinterhölzl (2007) (see also Bianchi & Frascarelli 2011)[26]. According to the typology in Frascarelli & Hinterhölzl (2007), topic shift involves another sub-type of topics, Aboutness-Shift topics (AS-topics). Strictly speaking, both (28) and (55) involve topic shift, but let's assume that Given-topic does not mean topic continuity, but instead refers to the entities likely to be topics at that point in the discourse. After uttering the first sentence in (55), Daniel enters the list of suitable topics and therefore it does not need overt marking as an AS-topic. In (28), because Kant is not expected to function as a topic, its use as the subject of the sentence involves marking him as an AS-topic, which requires overt realization. This explanation offers a means of encoding the difference between overt subjects and *pro* in terms of what we intuitively called 'degree of accessibility' (cf. Ariel, 1990)[27]:

(56)   3rd person *pro* is +G(iven)-topic[28]

A potential issue arises for topic-based accounts when considering the use of *pro* in subordinate clauses, especially when it is a variable bound by a quantifier, as in (57)[29]:

(57)   N-am   vorbit cu   nimeni$_n$ cât timp ***pro*$_n$** era   în anchetă.
        not-have.1 talked with nobody while   was.3SG in investigation
        'I talked to nobody while that person was under investigation'

According to Bianchi & Frascarelli (2010), G-topics are allowed in "central" adverbial clauses (which include the temporal clause in (57)). It is not impossible to say that

---

[26] A similar notion is 'in Focus' in Gundel et al. (1993), the highest level of their 'Givenness hierarchy'.

[27] The hypothesis in (56) resembles Frascarelli (2007)'s proposal that *pro* must agree with a local Aboutness-topic, which can be overt or covert (the latter in cases of topic continuity). But her proposal does not account for the difference between (28) and in (57): in both cases the Aboutness-topic is different from that of the preceding sentence yet in (28) it is overt but in (57) it is null. The problem is merely displaced: instead of explaining why the subject is overt in the former example and null in the latter, we have to explain why the A-topic is overt in the former but null in the latter.

[28] We do not analyze +Participant pronouns here, as this paper focuses on 3rd person pronouns from the perspective of noun ellipsis.

[29] Frascarelli (2007:728) proposes that in such cases the A(boutness)-topic is the *type* of entities quantified over and *pro*, referring to tokens of this type, "matches with the referential features of the individuals it [the A(boutness)-topic] includes."

at the level where the subordinate clause attaches, the variable quantified over is likely to be resumed in the subordinate and therefore it is marked as a G-topic.

A further difference between *pro* and overt subjects concerns gender. As we have seen in §3.3, the absence of verbal agreement in gender suggests that *pro*, whose features are in principle made visible via verbal agreement, might not be marked for gender. If absence of gender is a feature distinguishing *pro* from overt subject pronouns, then this may explain the fact that for the combination D_{pron}[idx]+inanimate [$_N$Ø]$_{non-anaph}$ only *pro* is available, while overt subject pronouns are ruled out.

As explained in §3.4 above, an inanimate [$_N$Ø]$_{non-anaph}$ is used to refer to entities that are not animate and are not categorized by a nominal concept: pronouns referring to objects of the event or proposition type, introduced by CPs, VPs or other non-nominal constituents and pronouns referring to objects accessible via perception but not yet categorized. As illustrated in (58)-(59), overt pronouns cannot be used in these contexts. Besides *pro*, feminine forms of demonstrative pronouns (e.g. *asta, aia*) can be used, but there is evidence that these forms are not actually feminine. For instance, in (58), the predicative adjective *imposibil* should agree with the feminine demonstrative and result in feminine agreement on the adjective (i.e. *imposibilă*), but masculine singular *imposibil* is used instead.

(58) Crede [**că vom   câştiga**]$_i$. ***pro$_i$/*el/*ea/Asta***   este imposibil(*ă).
      believes that will.1PL win   3MS/3FS/this.F is   impossible.MSG/*FSG
      '(S)he thinks [**we will win**]$_i$. **That$_i$** is impossible.'
(59) [context: an object *i* is visible in the direction of pointing]
      Ia uite-acolo! Mă întreb   ce-o   fi ***pro$_i$/*el$_i$/*ea$_i$/aia$_i$***
      behold-there REFL ask.1SG what-may.3SG be   3MS/3FS/that.F
      'Look over there. I'm wondering what **that** is.'

These data can be explained as follows (see Farkas 1990, Giurgea 2008, 2010): the masculine singular is a default form (as also evinced by masculine singular forms being used as adverbs in Romanian). In (58)-(59), the default form of the adjective is used because the subject lacks gender. The fact that genderless forms are used for the inanimate [$_N$Ø]$_{non-anaph}$ in definites is due to the fact that the Romanian gender system only contains two gender features: masculine and feminine (the so-called 'neuter' represents a nominal class that triggers masculine agreement in the plural and feminine agreement in the singular). All languages that have a neuter gender use the neuter for the inanimate [$_N$Ø]$_{non-anaph}$. Therefore, if Romanian masculine singular forms had been ambiguous between masculine and neuter (as traditional grammars of Romanian claim), we should have found masculine singular subject forms in (58)-(59) (e.g. *ăsta, ăla*). This expectation is contrary to fact. Note that *el* '3MS' can be used for antecedents headed by neuter nouns, see e.g. (51)b, so under the traditional analysis it would be ambiguous between masculine and neuter.

Thus, data like (58)-(59) provide evidence that *pro* can lack gender. The fact that *pro* is the only personal pronoun form acceptable in anomial uses is explained by the obligatory presence of a gender feature on strong subjects and its (possible) absence on *pro*.

Summing up, we have proposed that the D_{pron} found in 3rd person *pro* differs from the D_{pron} found in overt subject pronouns in that *pro* may lack gender and *pro* carries a [G-Topic] feature. Under this view, *pro* is a special pronoun restricted to nominative subjects. This restriction may of course be correlated with the existence of rich verbal agreement, as is the

case for Romanian, via which *pro*'s person and number features can be recovered in most environments. However, the actual possibility of feature identification is not required for *pro* to be licensed. Romanian possesses two nominative environments where there is no verbal agreement: infinitive CPs and perfect subjunctives. In both environments, as illustrated below, *pro* is licensed and can have any φ-feature. (60)b, where the overt subject pronouns are nominative, provides evidence that (60)a is a nominative environment.

(60)   a. Înainte de a merge *pro* acolo, s-a   întâmplat o nenorocire
               before of to go   there REFL-has happened a misfortune
               'Before going there, a misfortune happened.'
           b. Înainte de a merge {eu/tu/el/noi/voi/ei}   acolo...
               before of to go   I/you.SG//he/we/you.PL/they there

(61)   Să   fi   fost acolo!
           SBJV PRF been there
           'If only **I/you/(s)he/it/we/they** were there!'

We thus conclude that *pro* is restricted to nominative case. Chomsky's (2000) analysis of structural case as unvalued case excludes a Nominative specification in the lexicon, but as proposed in Giurgea (2014) for independent reasons, this proposal can be relaxed so as to allow distinct uCase features for each structural case, corresponding to the categorial features of the case licensors (uv* for accusative, uT for nominative, un for genitive, perhaps uAppl for structural dative)[30]. Under this proposal, *pro* would be lexically marked as [uT]. Alternatively, the absence of non-nominative features would be a matter of morphology (the only vocabulary item for the D_{pron} realized as *pro* would be +Nominative).

5.4. Corpus Study: Strong object forms vs. clitics

As we have seen in (52), there is evidence that strong object forms cannot receive neontological readings, as opposed to strong pronouns in other positions. This led to the proposal in (54), that object strong forms are restricted to D_{pron}[idx]. But this cannot be the feature that distinguishes strong forms from clitics, because, obviously, clitics are very often [idx] (they are not restricted to neontological readings, which are in general infrequent). We have also seen that clitics are not banned in environments where *pro* is dispreferred because of the reduced topic potential of the antecedent (see (29) vs. (28) in the preceding sub-section), which also means that clitics are not distinguished from strong forms by bearing [G-Topic]. As we already suggested in §3.3, discussing ex. (22), what differentiates strong object forms appears to be a contrast feature.

We tested this intuition against a corpus of 11 books (listed at the end of the article) in readable .pdf format. We collected examples by searching for the strings *pe el*, *pe ea* and *pe ei* (which identify strong 3rd person direct objects; searching *pe el* also returns *pe ele* 'DOM 3.FP'). First of all, the corpus search showed a clear preference for reference to persons: out of 184 instances of strong objects, only 2 refered to inanimates (1,08%). The corpus included

---

[30] The idea that case may reflect the licensor's categorial feature can also be found in Pesetsky (2013), who treats morphological case in terms of the copying of part-of-speech information from heads to their dependents.

scientific texts, where inanimates are most likely to be found, but we found extremely few examples of strong objects in them overall[31], none of which was inanimate. Both examples of inanimates come from Mihail Sebastian, *Opere I*. Secondly, of the 184 examples of strong DOM pronouns, 76 are 'forced' uses, i.e. cases where the syntax requires a strong form: modification by focal particles, by appositions, and coordination. Thirdly, out of the remaining 108 'unforced' uses, a contrast is clearly employed in 82 cases (75,92%). Of these 82 cases where there is contrast, not all require contrastive stress.

      An example of a strong form where contrastive stress *is* required (due to the existence of a parallel clause) is provided in (62). In (63) we offer an example of what we consider *weak* contrast; in this case, there is no parallel antecedent and no other male referent in the context. Thus, no contrastive stress is placed on the pronoun *pe el*. In fact, the clitic (*-l*) alone would have been possible; but some contrast seems to be involved between the subject, who knows her own feelings, and the object, who is not in a position to understand these feelings. Another example where the clitic alone would have been possible is (64), which also illustrates a strong pronoun used to refer to an inanimate object in an unforced context. In this case, there is a contrast between the left and right skis, but the event of pushing the right ski occurred at a distance in the previous text, and in the preceding sentence the left ski occupies a more prominent position, so resuming it does not require contrastive stress.

(62)   ființele   pe care le   iubea   mama   sau care   o
        beings(F)-the DOM which CL.3FP.ACC loved.3SG mother-the or   which CL.3FS.ACC
        iubeau   **pe**   **ea.**(M. Sebastian, *Opere* I, 1116)
        loved.3PL DOM her
        'the beings that my mother loved or who loved her'
(63)   Surâse   de aceea nehotărâtă, lăsându-1   **pe el**   să   înțeleagă   ce
        smiled.3SG therefore undecided.FS letting-3MS.ACC DOM him SBJV understand.3 what
        voia.   (M. Sebastian, *Opere* I, 747)
        wanted.3SG
        'That's why she had smiled undecided, letting him understand whatever he wanted.'
(64)   Împingi întâi skiul drept înainte cu genunchiul îndoit și cu piciorul stâng bine întins. Așa! Acum, trage skiul stâng până în dreptul celuilalt... bun!... și împinge-1 **pe el** înainte. Perfect.   (M. Sebastian, *Opere* I, 1029)
        'You first push the right ski forward with your knee bent and keeping the left leg straight. That's it! Now, pull the left ski until it's next to the other....good!... and push **it** forward. Perfect.'

Among the remaining 26 examples, 2 examples clearly belong to a type already described in (23), the deictic type, which is restricted to animates and requires stress, and a third example may be included in this type, in free indirect discourse. This type is not restricted to objects, it can be represented by PPs and subjects as well (see (65)). We assume a special D_{pron}[+idx][+deixis] that selects an animate $[_N Ø]_{non\text{-}anaph}$.

---

[31] In the scientific texts we consulted, we found instances of strong objects as follows: no example in V. Drăguț, *Arta gotică în România*, C. D. Nenițescu, V. Ioan (eds.), *Manualul inginerului chimist*, and Theodor Hristea, *Probleme de etimologie*; 2 examples in Șerban Papacostea, *Evul Mediu românesc*, 3 examples in Ovid Crohmălniceanu, *Literatura română între cele două războaie mondiale* 3rd volume, 2 examples in citations in the 2nd volume of this book.

(65)   [Extralinguistic context: a female person_p, an object_o, other possible referents....]
       a. Ce   zici   de   ea [☞]_{p/*o} ?
          what say.2SG about 3FS
        'What do you think about HER/*IT?'
       b. Ea [☞]_{p/*o} e mereu aici
          3FS   is always here
          'She/*it is always here.'

Among the remaining 23 examples (representing 21,29% of the unforced cases), where a contrast is not clear, 11 involve accusative experiencers arguably functioning as quirky subjects; in 4 examples, the pronoun is preverbal and the referent is the current topic and occurred as a nominative subject in the previous sentence, see (66). Here, the strong preverbal accusative may be a form of overtly marking topic continuity.

(66)   El însuşi îşi scornise numele de Alexei Alexeevici, (...)
       'He himself had come up with the name of Alexey Alexeevich for himself. (....)'
       Pe el îl   chema   Alexandru Spovidău.   (Ş. Bănulescu, *Opere*, 642)
       DOM him 3MS.ACC called.3SG Alexandru Spovidău
       'His name was Alexandru Spovidău.'

In the remaining 6 examples of quirky subjects, the strong accusative is postverbal and the sentence has focus fronting of an XP (see (67)a) or verum focus (see (67)b). Here, the overtness of the quirky subject seems to be a means of highlighting the fact that the active probe in the preverbal domain is a focus probe, triggering postverbal subject placement (the pattern is documented for nominative strong pronoun with verum focus, see Giurgea & Mîrzea-Vasile 2017).

(67)   a. Asta îl   ardea   **pe el**! (M. Preda, *Opere* I, 57)
          this CL.3MS.ACC burned.IMPF.3SG DOM him
          'That's what he cared about! (lit. That's what was burning him!)'
       b. N-o   muncea   **pe ea** grija   de ruptura   ei, (...)
          not-3FS.ACC tormented.IMPF.3SG DOM her care-the of breaking-the her.GEN
          'It's (definitely) not that she cared of her getting broken...'

Among the remaining 12 examples, a contrast is conceivable for 9 cases, leaving us with 3 examples where contrast is unlikely. It is important to note that contrast is also present in the majority of the forced examples; contrast is obvious for the focal particle modification cases, but is also arguably present in coordination cases.
       The corpus data strongly suggests that [contrast] is a feature that distinguishes the strong pronoun D_{pron}[idx] from the clitic D_{pron}[idx]. This is compatible with hypotheses that less complex forms (e.g. clitics) are used unless more complex forms are needed to signal a further distinction (e.g. Ariel, 1990; Levinson, 1987, 2000; Ivan, 2020).[32]
       With respect to the predominance of animate referents in the corpus, a possible explanation is that the D_{pron}[+contrast] is lexicalized as taking a human $[_N\emptyset]_{non-anaph}$, like the

---

[32] In Romanian, where clitic doubling is obligatory for strong pronouns as well as for other types of DPs (see Irimia 2020, a.o.), clitics can be analyzed as the spell-out of an agreeing verbal functional head, rather than as moved pronouns (see Sportiche, 1999).

deictic D_{pron}. For inanimate referents, it might be the case that a PF-rescuing operation inserts a strong form for an inanimate D_{pron}+E [+contrast], so as to fulfill a requirement of overt realization of items marked with [contrast].

5.5. Summary of the lexical differences between the strong and weak series

Summing up, we proposed an account that can capture the competition between strong and weak series of (3rd person) pronouns in Romanian without resorting to degrees of accessibility. Here are our core claims and observations:

(i) *Pro* differs from overt subject pronouns by bearing [G-Topic] and lacking gender (at least in some of its uses). Moreover, it is specified in the lexicon as nominative (uT-Case).

(ii) Strong object forms differ from clitics (or rather from the null pronouns whose features are spelled-out via agreement on a verbal functional head) by bearing [contrast]. Moreover, the D_{pron} marked [contrast] is correlated with [idx] and the selection of an animate $[_N\emptyset]_{non\text{-}anaph}$.

(iii) Strong forms can also represent a D_{pron}[idx][+deixis] which selects an animate $[_N\emptyset]_{non\text{-}anaph}$. This pronoun differs from demonstratives by not being marked as proximal or distal.

6. CONCLUSIONS

We examined the distribution of empty N(P)s, both with an anaphoric ($[_N\emptyset]_{anaph}$) and a non-anaphoric ($[_N\emptyset]_{non\text{-}anaph}$) reading, and found that the combination D_{def}+$[_N\emptyset]$, in cases of 'total emptiness' (where there is no overt constituent inside the complement of D), is spelled-out as 3rd person personal pronouns. This hypothesis is supported by evidence for N-anaphora in pronouns, including 'paycheck' or 'neontological' uses of pronouns and non-semantic gender on pronouns. However, an analysis of pronouns as THE+$[_N\emptyset]$ faces various issues, which we addressed by proposing that the definite determiner used in pronouns (labeled D_{pron}) has the semantics of THE but differs from it in formal features.

We argued that THE could not be used in cases of *total emptiness* because it does not bear the required features. A determiner must bear features licensing emptiness in order to combine with an empty complement, in particular, [E] for an elliptical empty complement and [+∅] for a non-elliptical one. Thus, the impossibility of typical THE to surface with null complements is a matter of THE lacking [E] and [+∅]. In addition to these features, D_{pron} may differ from THE in other features, such as gender (e.g. *he, she, it* vs. *the* in English or the null subjects of Indo-European languages, which may lack gender) and the formal features responsible for clitic placement. Additionally, we argued that the differences in formal features between D_{pron} and typical THE can also account for the different behavior of pronouns and other definite DPs with respect to the binding principles.

This hypothesis allows for the existence of several different series of forms which functionally correspond to the THE+$[_N\emptyset]$ combination, as is the case of the series of strong and clitic or null forms found in various languages. To investigate the differences between strong and weak pronominal forms, we analyzed the particular situation of Romanian, proposing features by which *pro* differs from overt subjects and strong objects differ from clitics. One such feature is [contrast], as also evidenced by the data in our corpus study.

CORPUS

Ş. Bănulescu, *Opere*, I, ed. O. Soare, Bucharest, Fundaţia Naţională pentru Ştiinţă şi Artă and Univers Enciclopedic, 2005.
O. Crohmălniceanu*, Literatura română între cele două războaie mondiale*, II, Bucharest, Minerva, 1974; III, Bucharest, Minerva, 1975.
M. Daneliuc, *Marilena şi câteva voluptăţi*, Bucharest, Editura Univers, 1999.
V. Drăguţ, *Arta gotică în România*, Bucharest, Meridiane, 1979.
T. Hristea, *Probleme de etimologie: studii, articole, note*, Bucharest, Editura Ştiinţifică, 1968.
G. Liiceanu, *Despre limită*, Bucharest, Humanitas, 1994.
C. D. Neniţescu, V. Ioan (eds.), *Manualul inginerului chimist.*, I. Bucharest, Editura Tehnică, 1951.
Ş. Papacostea, *Evul Mediu românesc*, Bucharest, Corint, 2001.
M. Preda, *Opere*, I, ed. V. Crăciun, Bucharest, Fundaţia Naţională pentru Ştiinţă şi Artă, 2002.
M. Sebastian, *Opere*, I, ed. M. Constantinescu-Podocea, Bucharest, Fundaţia Naţională pentru Ştiinţă şi Artă, 2011.

REFERENCES

Ahn, D., 2019, *The Determinacy Scale: A competition Mechanism for Anaphoric Expressions*, doctoral dissertation, Harvard University. http://nrs.harvard.edu/urn-3:HUL.InstRepos:42029564.
Ariel, M., 1990, *Accessing noun-phrase antecedents*, New York, Routledge.
Bernstein, J., F. Ordóñez, F. Roca, 2018, "Sardinian descendants of *ipse* in comparative relief", in: A. Chilà, A. De Angelis (eds), *Capitoli di morfosintassi delle varietà romanze d'Italia: teoria e dati empirici*, Palermo, Centro di Studi Filologici e Linguistici Siciliani, 317–336.
Bernstein, J., F. Ordóñez, F. Roca, 2021, "Emphatic elements and the development of definite articles: evidence for a layered DP in Early Romance", *Journal of Historical Linguistics* 5, Article 22: 1-32, https://doi.org/10.18148/hs/2021.v5i16-25.63.
Bi, R. A., P. Jenks, 2019, "Pronouns, null arguments, and ellipsis in Mandarin Chinese", in: M.T. Espinal, E. Castroviejo, M. Leonetti, L. McNally, C. Real-Puigdollers, Cristina (eds.), *Proceedings of Sinn und Bedeutung 23*, Bellaterra, Universitat Autònoma de Barcelona, 127–142.
Bianchi, V., M. Frascarelli, 2010, "Is topic a root phenomenon?", *Iberia*, 2, 1, 43–88.
Bošković, Ž., 2018, "On pronouns, clitic doubling, and argument ellipsis: Argument ellipsis as predicate ellipsis", *English Linguistics*, 35, 1–37.
Bruening, B., 2014, "Precede-and-command revisited", *Language* 90, 342–388.
Cardinaletti, A., M. Starke, 1999, "The typology of structural deficiency: A case study of the three classes of pronouns", in: H. van Riemsdijk (ed.), *Clitics in the Languages of Europe*, Berlin, New York, Mouton De Gruyter, 145–234.
Cheng, L. L., C. Heycock, R. Zamparelli, 2017, "Two levels for definiteness", in: M. Yoshitaka Erlewine (ed.), *Proceedings of GLOW in Asia XI*, 79–93. https://doi.org/10.13140/ RG.2.2.25020.95364.
Chomsky, N., 1981, *Lectures on Government and Binding*, Dordrecht, Foris.
Chomsky, N., 1986, *Knowledge of language: Its nature, origin, and use*, Greenwood Publishing Group.
Chomsky, N., 2000, „Minimalist inquiries: The framework", in: R. Martin, D. Michaels, J. Uriagereka (eds), *Step by Step*: *Essays on Minimalist Syntax in Honor of Howard Lasnik*, Cambridge, MA, MIT Press, 89–155.
Corblin, F., 2006, "Pronouns and mentions", in: I. Korzen, L. Lundquist (eds), *Comparing Anaphors. Between Sentences, texts and Languages*, Copenhagen, Samfundslitteratur, 27–43.
Cooper, R., 1979, "The interpretation of pronouns", in: F. Heny, H. Schnelle (eds.), *Syntax and Semantics 10: Selections from the Third Gröningen Round Table*, New York, Academic Press, 61–92.

Elbourne, P., 2001, "E-Type Anaphora as NP-Deletion", *Natural Language Semantics*, 9, 241–288.

Elbourne, P., 2005, *Situations and Individuals*, Cambridge, MA, MIT Press.

Elbourne, P., 2008, "Demonstratives as individual concepts", *Linguistics and Philosophy*, 31, 4, 409–466.

Elbourne, P., 2013, *Definite descriptions*, Oxford, Oxford University Press.

Evans, G., 1980, "Pronouns", *Linguistic Inquiry*, 11, 337–361.

Farkas, D., 1990, "Two Cases of Underspecification in Morphology", *Liguistic Inquiry*, 21, 539–550.

Frascarelli, M., 2007, "Subjects, topics and the interpretation of referential *pro*: An interface approach to the linking of (null) pronouns", *Natural Language and Linguistic Theory*, 25, 691–734.

Frascarelli, M., R. Hinterhölzl, 2007, "Types of Topics in German and Italian", in: K. Schwabe, S. Winkler (eds), *On Information Structure Meaning and Form*, Amsterdam, Philadelphia, John Benjamins, 87–116.

Giannakidou, A., J. Merchant, 1996, "On the interpretation of null indefinite objects in Modern Greek", *Studies in Greek Linguistics¸* 17, 141–155.

Giurgea, I., 2008, "Gender on definite pronouns", *Bucharest Working Papers in Linguistics*, 10, 1, 97–120.

Giurgea, I., 2010, *Pronoms, Déterminants et Ellipse Nominale. Une approche minimaliste*, Bucureşti, Editura Universităţii din Bucureşti.

Giurgea, I., 2012, "The origin of the Romanian possessive-genitival article *al* and the development of the demonstrative system", *Revue roumaine de linguistique*, 57, 1, 35–65.

Giurgea, I. 2013a, "The Syntax of Determiners and Other Functional Categories", in: C. Dobrovie-Sorin, I. Giurgea (eds.), *A Reference Grammar of Romanian. Volume I: The Noun Phrase*, Amsterdam, Philadelphia, John Benjamins, 97–174.

Giurgea, I., 2013b, *Originea articolului posesiv-genitival* al *şi evoluţia sistemului demonstrativelor în română*, Bucharest, Editura Muzeului Naţional al Literaturii Române.

Giurgea, I., 2014, "Romanian *al* and the syntax of case heads", *Bucharest Working Papers in Linguistics*, 16, 2, 69–98.

Giurgea, I., C. Mîrzea Vasile, 2017, "Syntactic Effects of Verum Focus in Romanian", *Revue roumaine de linguistique*, 62, 3, 323–338.

Giurgea I., I. Nedelcu, 2009, "Elipsa nominală şi construcţia partitivă", in R. Zafiu, G. Stoica, M.V. Constantinescu (eds), *Limba română: teme actuale. Actele celui de al 8-lea Colocviu al Catedrei de limba română*, Bucureşti, Editura Universităţii din Bucureşti, 109–124.

Grice, H. P., 1975, "Logic and conversation", in: P. Cole, J. Morgan (eds), *Syntax and Semantics*, vol. III, New York, Academic Press, 22–40.

Grosz, P., 2019, "Pronominal Typology and Reference to the External World", in: J. Schlöder, D. McHugh, F. Roelofsen (eds), *Proceedings of the 22nd Amsterdam Colloquium¸* 563–573.

Gundel, J. K., Hedberg, N., Zacharski, R., 1993, "Cognitive status and the form of referring expressions in discourse", *Language*, 274–307.

Haik, I., 1985, "Pronouns of Laziness", *North East Linguistics Society* 16, issue 1, article 15. Available at: https://scholarworks.umass.edu/nels/vol16/iss1/15

Hanink, E., 2017, "The German definite article and the 'sameness' of indices", *Penn Working Papers in Linguistics*, 23, 1, 63-72.

Hanink, E., 2021, " 'Same': Structural sources of anaphora and relativization", *Glossa* 6, 1, 1–50.

Hankamer, J., I. Sag, 1976, "Deep and Surface Anaphora", *Linguistic Inquiry*, 7, 3, 391–428.

Heim, I., 1990, "E-type pronouns and donkey anaphora", *Linguistics and Philosophy*, 13, 137–177.

Heim, I., 2011, "Definiteness and determinacy", in C. Maienborn, K. von Heusinger, P. Portner (eds.), *Semantics. An International Handbook of Natural Language Meaning*, Berlin, Boston, Walter de Gruyter, 996–1025.

Holmberg, A., 1986, *Word Order and Syntactic Features in the Scandinavian Languages and English*, doctoral dissertation, University of Stockholm.

Holmes, P., Hinchliffe, I., 2013, *Swedish: A Comprehensive Grammar*, 3rd ed., Routledge Comprehensive Grammars, Routledge, New York.

Huddlestone, R., G. Pullum, 2002, *The Cambridge Grammar of the English Language*, Cambridge, Cambridge University Press.

Irimia, M., 2020, "Variation in differential object marking: on some differences between Spanish and Romanian", *Open Linguistics*, 6, 424–462. https://doi.org/10.1515/opli-2020-0110.

Ivan, R-R., 2020, *Talking about her(self): Ambiguity Avoidance and Principle B.*, doctoral dissertation, University of Massachusetts, Amherst. https://doi.org/10.7275/19173480 https://scholarworks.umass.edu/dissertations_2/2031

Jaeggli, O., 1982, *Topics in Romance Syntax*, Dordrecht, Foris.

Jenks, P., R. Konate, 2022, "Indexed definiteness", *Glossa*, 7, 1, 1–44.

Johnson, K., 2013, "Pronouns vs. Definite Descriptions", in: J. Grinstead M. Becker, J. Rothman (eds), *Generative Linguistics and Acquisition: Studies in honor of Nina M. Hyams*, Amsterdam, Philadelphia, John Benjamins, 157–184.

Karttunen, L., 1969, "Pronouns and variables", in: R. Binnick, A. Davison, G. Green, J. Morgan (eds), *Papers from the fifth regional meeting of the Chicago Linguistic Society*, Chicago, Department of Linguistics, University of Chicago, 108–116.

Kratzer, A., 2009, "Making a Pronoun: Fake Indexicals as Windows into the Properties of Pronouns", *Linguistic Inquiry*, 40, 187–237.

Kurafuji, T., 1998, "Dynamic binding and the E-type strategy: Evidence from Japanese", in: D. Strolovitch, A. Lawson (eds.), *Semantics and Linguistic Theory (SALT) 8*, Cornell University, 129–144. DOI: https://doi.org/10.3765/salt.v8i0.2802

Levinson, S., 1987, "Pragmatics and the grammar of anaphora", *Journal of Linguistics*, 23:379–434.

Levinson, S., 2000, *Presumptive meanings: The theory of generalized conversational implicature*, Cambridge, Mass., MIT Press.

Llombart-Huesca, A., 2002, "Anaphoric *one* and NP-ellipsis", *Studia linguistica*, 56, 1, 59–89.

Longobardi, G., 1994, "Reference and Proper Names: a Theory of N-Movement in Syntax and Logical Form", *Linguistic Inquiry*, 25, 4, 609–665.

Mandelkern, M., D. Rothschild, 2020, "Definiteness projection", *Natural Language Semantics*, 28, 77–109.

Merchant, J., 2001, *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*, Oxford, Oxford University Press.

Neeleman, A., K. Szendrői, 2007, "Radical pro drop and the morphology of pronouns", *Linguistic Inquiry*, 38, 4, 671–714.

Nunberg, G., 1993, "Indexicality and deixis", *Linguistics and Philosophy*, 16, 1–43.

Panagiotidis, P., 2002, *Pronouns, clitics and empty nouns*, Amsterdam, Philadelphia, John Benjamins.

Partee, B.H., 1978, *Bound Variables and other Anaphors,* in: D. Waltz (ed.), *Proceedings of TINLAP*, 2, 79–85.

Patel-Grosz, P., P. Grosz, 2017, "Revisiting pronominal typology", *Linguistic Inquiry*, 48, 2, 259–297.

Perlmutter, D. M., J. Oresnik, 1973, "Language particular rules and explanation in syntax", in: S. Anderson, P. Kiparsky (eds), *Festschrift for Morris Halle*, New York, Holt, Reinhart, and Winston, 419–459.

Pesetsky, D., 2013, *Russian* Case *Morphology and the Syntactic Categories*, Cambridge, MA, MIT Press.

Pesetsky, D., E. Torrego, 2007, "The syntax of valuation and the interpretability of features", in: S. Karimi, V. Samiian, W. Wilkins (eds), *Phrasal and clausal architecture: Syntactic derivation and interpretation, In honor of Joseph E. Emonds*, Amsterdam, John Benjamins, 262–294.

Postal, P., 1969, "On so-called 'pronouns' in English", in: F. Dinneen (ed.), *Report on the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies*, Washington, DC, Georgetown University Press, 177–206.

Reinhart, T., 1983, *Anaphora and Semantic Interpretation*, London, Croom Helm.

Rizzi, L., 1986, "Null objects in Italian and the theory of *pro*", *Linguistic Inquiry,* 17, 501–557.

Roelofsen, F., 2010, "Condition B effects in two simple steps". *Natural Language Semantics*, 18, 115–140.

Runić, J., 2014, *A new look at clitics, clitic doubling, and argument ellipsis: Evidence from Slavic*, doctoral dissertation, University of Connecticut.

Saab, A., 2019, "Nominal ellipsis", in: J. van Craenenbroeck, T. Temmerman (eds), *The Oxford Handbook of Ellipsis*, Oxford, Oxford University Press, 526–561.

Sauerland, U., 2000, "The content of pronouns: evidence from Focus", in: B. Jackson, T. Matthews (eds), *Proceedings of SALT 10,* Ithaca, Cornell University, CLC Publications, 167–184.

Sauerland, U., 2008, "Pseudo-Sloppy Readings in Flat Binding", in: O. Bonami, P. Cabredo Hofherr (eds), *Empirical Issues in Syntax and Semantics* 7, 331–349. http://www.ssp.nrs.fr/eiss7.

Schlenker, P., 2005, "Minimize restrictors! (Notes on definite descriptions, condition C and epithets)", in: E. Maier, C. Bary, J. Huitink (ed.), *Proceedings of Sinn und Bedeutung 9*, 385–416, www.ru.nl/ncs/sub9.

Schroten, J., 1992, "On Spanish definite determiners: Personal pronouns and definite articles", *Recherches de Linguistique Romane et Française d'Utrecht*, 9–24.

Schütze, C., 2001, "Semantically empty lexical heads as last resorts", in: N. Corver, H. van Riemsdijk (eds), *Semi-lexical Categories. The Function of Content Words and the Content of Function Words,* Berlin, Boston, Mouton de Gruyter, 127–188.

Schwarz, F., 2009, *Two types of definites in natural language*, doctoral dissertation, University of Massachusetts at Amherst.

Schwarz, F., 2019, "Weak vs. strong definite articles: Meaning and form across languages", in: A. Aguilar-Guevara, J. Pozas Loyo, V. Vázquez-Rojas Maldonado (eds), *Definiteness across languages*, Berlin, Language Science Press, 1–37.

Simonenko, A., 2014, *Grammatical ingredients of definiteness*, doctoral dissertation, McGill University, Montreal.

Sommerstein, A., 1972, "On the So-Called 'Definite Article' in English", *Linguistic Inquiry*, 3, 197–209.

Sportiche, D., 1999, "Subject clitics in French and Romance Complex Inversion and Clitic Doubling", in: K. Johnson, I. Roberts (eds.), *Beyond Principles and Parameters*, Dordrecht, Springer, 189–221.

Wechsler, S., 2004, "Number as person", in: O. Bonami, P. Cabredo Hofherr (eds), *Empirical issues in syntax and semantics 5*, 255–274.

Wechsler, S., 2011, "Mixed agreement, the person feature, and the index/concord distinction", *Natural Language & Linguistic Theory*, 29, 999–1031.

Wolter, L. K., 2006, *That's that: The Semantics and Pragmatics of Demonstrative Noun Phrases*, doctoral dissertation, University of California – Santa Cruz.