INDEFINITE NPs AND *PE*-MARKING
IN ROMANIAN[*]

SOFIANA CHIRIACESCU

1. INTRODUCTION

It is well known that Romanian is a language which marks its direct objects differentially by means of the particle *pe* (Niculescu 1965, Chiriacescu 2007, von Heusinger & Onea 2008). As in other Romance languages, direct object case marking is obligatory for some objects, optional for others and excluded for a third set, whereas this phenomenon is mostly accompanied by clitic doubling. The factors that are considered to be the main triggers of the marked direct object form are animacy, definiteness and specificity. DOM-marking starts at the more prominent part of these scales covering areas of different length (Farkas 1978, Dobrovie-Sorin 1994, von Heusinger & Onea 2008). So, whether a direct object will be marked or not depends on the amount of features united in the object in cause.

Up to this point, the literature (Chiriacescu 2007, Stark& Sora 2008, etc.), has focused on the development and on the parameters licensing DOM in Romanian, leaving aside certain constructions that could not be accounted for by means of the above mentioned factors. These constructions involve direct objects realized as post-verbal unmodified definite or indefinite nominal phrases. In such cases, both the marked and the unmarked direct object constructions coexist, whereby the difference in meaning between the two forms is subtle and difficult to analyze.

In the present paper I will focus on direct objects realized as indefinite nominal phrases which involve not clearly explained alternations between a *pe*-marked and an unmarked form. The examples (1a) and (1b) illustrate this variation. The context sentence (A) can be continued either as in (1a) where the indefinite direct object is *pe*-marked, or as in (1b) where the object appears unmarked:

(1)   A: *Ce face Maria? (What does Maria do?)*
       a.   *Maria o   vizitează   pe   o   prietenă*
           Maria CL   visits   PE   a   friend
           'Maria visits a friend.'
       b.   *Maria vizitează   o   prietenă*
           Maria visits   a   friend
           'Maria visits a friend.'

Examples as the one presented above underline the insufficiency of the general acknowledged criteria that trigger DOM to account for the optionality of *pe*-marking in Romanian. This paper proposes the introduction of the additional discourse- based parameter "referential persistence", to explain more subtle differences as those involving indefinite unmodified noun phrases. Referential persistence (Givon 1983, Ariel 1988) designates a discourse pragmatic property that is weaker than topicality and that reconstructs the informal description of "importance for the subsequent discourse". This property, for which a quantitative measure will be offered, indicates that the *pe*-marked indefinite unmodified NP will be taken up in the subsequent discourse with a high probability.

In Section 2 I will briefly present the local parameters that trigger DOM in Romanian, not taking into account global parameters such as the lexical semantics of the verb, secondary predication, etc. In Section 3 I will sketch out the concepts of topic continuity and accessibility, which will represent the staring point for the analysis of the discursive nature of *pe*-marked indefinite NPs. A special emphasis will be put on the concept of discourse prominence and its subcomponent "referential persistence". In Section 4 I will show that *pe*-marked direct objects realized as indefinite NPs are taken up in the subsequent discourse more often than their unmarked counterparts, signalizing a higher degree of activation. Section 5 contains the concluding remarks of this paper, as well as some open remained questions.

2. LOCAL FACTORS DETERMINING DOM

As already mentioned in the introductory part, animacy, definiteness and specificity are the three main factors that determine the *pe*-marking of a direct object in Romanian. In the following, I will briefly sketch the distribution of *pe* as a case marker along these scales, paying special attention to entities realized as indefinite direct objects in postverbal position[1]. Only direct objects specified for the semantic feature [+human] will be taken into account.

2.1 Definite expressions

In nowadays Romanian, *pe*-marking typically targets those direct objects which denote human entities. Personal pronouns referring to animate entities are always marked with *pe* and doubled by a clitic. Direct objects realized as reflexive pronouns, the interrogative and relative pronouns *care* and *cine* ("that/ who") referring to animates as well as inanimates, demonstrative pronouns (except *asta* "this".FEM.SG referring to neuter nouns) are also preceded by *pe*. The negative pronoun *nimeni* ("nobody") and the indefinite pronouns are also differentially marked with *pe* when they replace a noun referring to an individual.

Proper names referring to humans or strongly individuated, personified animals are regularly case marked with *pe* when they appear in direct object position. Exceptions from this rule are proper names referring to names of countries or cities, even if these names are used metonymically, denoting the inhabitants of a city.

In the case of definite unmodified direct objects, Romanian shows a general blocking effect of prepositions upon the definite article. Even though we do not analyze the differential marker *pe* in terms of a preposition, the above blocking phenomenon nevertheless holds, as illustrated in (2).

(2)   *Doctorul   o   examinează   pe   fat(*-a)*
        Doctor.DEF   CL   examines   PE   girl
        'The doctor examines the girl.'

Note that the presence or absence of a clitic pronoun does not improve the acceptability of the sentence; however, this blocking effect disappears if the noun is further modified. While in the case of other prepositions this rule strictly blocks the apparition of the definite article, it renders two possibilities in the case of DOM-marked nouns. The ungrammatical sentence (2) can be replaced in two different ways: as in (3a) where *pe* is omitted and the definite article is kept, or as in (3b) where *pe* is retained but the definite article is omitted:

(3)   A: *O fată merge la doctor. (A girl goes to the doctor.)*
        a.   *Doctorul   examinează   fata*
                Doctor.DEF   examines   girl.DEF
                'The doctor examines the girl.'
        b.   *Doctorul   o   examinează   pe   fată*
                Doctor.DEF   CL   examines   PE   girl
                'The doctor examines the girl.'

---

[1] Because of lack of space I have to generalize over many exceptions but see see Farkas (1978), Gramatica Academiei Române (2005), Chiriacescu (2007), von Heusinger & Onea (2008), Stark & Sora (2008), for a detailed picture of this distribution.

To keep the story simple, I will neither explain nor enumerate the blocking effects and exceptions found within the class of definite unmodified NPs. It suffices to emphasize at this point that, in the case of direct objects realizes by means of a definite unmodified nominal phrase, Romanians can generally choose between two constructions, like those in (3a) and (3b) above. Both sentences are grammatical and have the same propositional content and depending on the context and language register, speakers tend to prefer one construction over the other (for an extensive analysis of examples with definite NPs in direct object position, see von Heusinger & Chiriacescu 2008, to appear).

2.2   Indefinite nominal phrases

For indefinite human direct objects, *pe*-marking is optional; however, the parameters that might influence the DOM-marking are not quite clear, this being a typical instance of "fluid" constraints (see Malchukov & de Hoop 2007).

In what follows, we test different types of specificity: scopal specificity with intensional and extensional operators and epistemic specificity in transparent contexts. Scopal specificity, whether with extensional or intensional operators, triggers *pe*-marking. While the sentence (4a) is ambiguous between a specific reading (or wide scope) and a non-specific (or narrow scope) reading, the non-specific reading in (4b) is ruled out due to the presence of *pe* (Dobrovie-Sorin 1994). The variation between wide and narrow scope is maintained for constructions with intentional operators, like in (5):

(4)   Extensional operators (universal quantifiers)
    a.   *Toți   bărbații   iubesc o   femeie*
        All   men   love   a   woman
        'All men love a woman.' (specific/ non-specific)
    b.   *Toți   bărbații   o   iubesc pe   o   femeie*
        All   men   CL   love   PE   a   woman
        'All men love a/ this woman.' (only specific)

(5)   Intensional operators
    a.   *Ion   caută   o   secretară*
        John   looks for   a   secretary
        'John looks for a secretary.' (specific/ non-specific)
    b.   *Ion   o   caută   pe   o   secretară*
        John   CL   looks for   PE   a   secretary
        'John looks for a secretary.' (only specific)

In a "transparent" context, the contrast between (6a) and (6b) can be (partially) explained by means of epistemic specificity. It seems that the characteristics of the referent *un prieten* ('a friend') are not important in (6a), while (6b) gives the impression that the speaker intends to communicate more information about the direct object.

(6)   Transparent context
    a.   *Petru   a   vizitat un   prieten*
        Petru   Aux.   visited a   friend
        'Petru visited a friend.'
    b.   *Petru   l   -a   vizitat pe   un   prieten*
        Petru   CL   Aux.   visited PE   a   friend
        'Petru visited a friend.'

This very interesting variation hints towards a more complex system of contrasts (see von Heusinger & Onea 2008 for a detailed analysis). In the next sections, we will concentrate on the variation between (6a) and (6b), which is not sufficiently described in terms of epistemic specificity.

2.3   Summary

The next table (7) comprises the referential contexts in which direct objects are *pe*-marked in Romanian. Besides the type of phrase through which the objects are realized, the table also makes a clear distinction in the domain of indefinite nominal phrases with respect to specificity. So, indefinite non-specific NPs are not differentially marked.

(7)   Referentiality Scale for *pe*-marking in Romanian for human direct objects

pers. pron.	*> propr. noun*	*> def. NP*	*> indef. spec. NP*	*> indefinite non-spec.NP*
obligatory		*obligatory (with exceptions)*	*optional*	*ø*

As we could see so far animacy, definiteness and specificity cannot thoroughly account for the distribution of *pe* with the free variation found in the domain of unmodified indefinites in contexts like that in (6). Neither topicality, nor other global parameters (like the lexical properties of the governing verb or secondary predications, etc.) are general enough or useful to explain this variation.

In what follows, I will account for the problematic examples involving indefinites by adding a more general parameter on the list of the factors licensing DOM in Romanian. I will use the gradual concept of "topic continuity" introduced by Givon (1981), to show that *pe*- marked indefinites are more prominent in the discourse than their unmarked counterparts.

3. ACCESSIBILITY AND INDEFINITE REFERENCE

Before the seminal work of Givon, the concept of topic was understood in an intuitive way, a sentence was therefore conceived as containing at most one topic. Givon (1981, 1983) was the first to introduce the graded concept of "topic continuity" (the situation in which the same topic extends over several clauses) for the behavior of discourse referents across more than one sentence. He showed, for example, that an entity realized as a zero anaphor is an accessible topic and is most continuous, while an indefinite NP is less accessible and therefore usually discontinuous.

3.1. Accessibility Hierarchies

Accessibility/ giveness theories offer a procedural analysis of the referring expressions, as marking different degrees of mental accessibility. In this framework, where "accessibility" is regarded as a gradient category rather than a categorical one, as in DRT, a discourse referent can be more or less accessible. The basic idea behind this theory is that referring expressions are actually accessibility markers by giving evidence to the addressee on how to retrieve the appropriate mental representation for an entity. So, the referential form of the referent mirrors its accessibility status and its prominence in the discourse. There have been many attempts to capture the correlation between the accessibility of an entity and the referential expression through which this entity is realized, for example Prince's (1981) "Familiarity Scale", Ariel's (1988) "Accessibility Hierarchy" or Gundel, Hedland & Zacharski's (1993) "Giveness Hierarchy" which is exemplified in (8).

(8)   Givenness Hierarchy (Gundel, Hedland & Zacharski 1993)

in focus >	activated >	familiar >	uniquely identifiable	> referential >	type identifiable
it	*that, this*	*that* N *this* N	*the* N	indefinite *this* N	*a* N

more accessible ⟷ less accessible

This approach suggests that the mental accessibility of an entity has a strong impact upon the reference form which will be chosen to refer to it. In English, for example, there are two determiners which can precede a NP in a specific indefinite context: the indefinite article *a* and the determiner *this* (the referential and not the deictic *this* determiner). However, these two forms cannot be used interchangeably. Ionin (2006) notes that besides their

different scopal behavior (*this*-determiners do not take narrow scope with respect to intensional or modal operators and negations), the two forms also differ with respect to the noteworthiness property. The examples (9a) and (9b) underline the latter difference:

(9)   a.   He put √a/ #this 3$ stamp on the envelope, so he wants to send the letter.
      b.   He put √a/ √this 3$ stamp on the envelope and realized only afterwards that it was worth a 100$.

If the speaker uses *this* over *a* in (9a), s/he conveys additional information about the NP headed by the determiner. Accordingly, the hearer expects that the speaker will talk about the stamp again, perhaps explaining what the noteworthy quality of the stamp is. Because this expectation remains unfulfilled in (9a) in contrast to (9b), the usage of *this* is rendered infelicitous. So, in the so called "transparent context" as in (9), a noteworthy referent can be preceded by *this* if it will be implicitly or explicitly referred to again (Prince 1981). We will see in Section 4. that the apparent optionality of the *pe*-marked construction and the unmarked one can be explained (in most contexts) in a similar manner as the variability presented above.

3.2   Discourse prominence

Indefinite expressions do not "look back" or refer to already introduced referents in the same way as definite expressions. However, as already stated, Gundel et al. (1993) assume two kinds of indefinite NPs - one specific and one non-specific. Givon (1981) and Wright & Givon (1987) give more types of indefinites in order to explain the development of the indefinite article at different stages. They distinguish between specific and non-specific uses of indefinite expressions, among others.

Starting from the premise that thematically more important referents tend to be more anaphorically accessible and cataphorically persistent, Givon develops two measurements for topicality: referential distance ("look back") and topic persistence ("look forward"). The first factor, "referential distance" determines how recently an entity has been mentioned, by looking at the sentences on the left of the referent. The second factor "persistence" measures how long the entity will remain in the discourse after it was introduced for the first time. The way in which an entity is referred to reflects the speaker's intentions about the role this entity will play in the subsequent discourse. Because the two factors often overlap, we will only look at the "referential persistence" of the referent introduced in the discourse.

4. REFERENTIAL PERSISTENCE

In this section I will illustrate the referential persistence of a *pe*-marked referent by comparing this type of construction with the one in which the referent in direct object position is not preceded by *pe*. The first article in (10) contains a direct object that was introduced by means of *pe* in the discourse, whereas in the second article (11), the same indefinite direct object occurs without *pe*. The two article extracts relate the same shooting event in the same way, the only difference being the form of realization of the two objects.

(10) *pe*-marked DO[1]	(11) *pe*-unmarked DO[2]
[1] Neculai Florea, de 40 de ani, viceprimarul satului Horodniceni, şi-a pus poliţia pe cap după ce l-a împuşcat cu un pistol cu gloanţe de cauciuc **pe un tânăr** din localitate.	[1] Viceprimarul Neculai Florea, din comuna Horodniceni, este cercetat de poliţie după ce în noaptea de sâmbătă spre duminică a împuşcat în picior **un tânar** de 24 de ani la discotecă.
[2] Incidentul s-a petrecut în noaptea de 10 spre 11 februarie, la discoteca ce aparţine soţiei viceprimarului Florea şi a fost reclamat la poliţie în cursul după amiezii, la ora 15:40.	[2] Viceprimarul, care este membru PNG, a scos pistolul pentru a interveni într-o încăierare între tineri, care avea loc în discoteca familiei sale.
[3] La ora respectivă, **Vasile M**., de 24 de ani, din	[3] El este asociat unic, iar soţia sa administrator.

comuna Horodniceni, **pro** s-a adresat postului de poliție reclamând că **pro** a fost împuşcat în picior de viceprimarul Neculai Florea.

[4] La Horodniceni s-a deplasat în aceeaşi zi o echipă operativă a Serviciului arme, explozivi, substanţe toxice din IPJ Suceava, pentru a elucida cazul.

[5] Din primele verificări efectuate s-a stabilit că în cursul nopţii, la discoteca viceprimarului, pe fondul consumului de alcool, a avut loc o altercaţie, iar Neculai Florea a folosit pistolul cu gloanţe de cauciuc împotriva lui **Vasile M**., pe care **l**-a împuşcat în picior, rănindu-**l**.

[6] Viceprimarul Neculai Florea susţine că a fost nevoit să facă uz de armă, întrucât a fost agresat **de tânărul în cauză**.

[7] A spus că în cursul nopţii de 10 spre 11 februarie, în discoteca administrată de soţia lui a izbucnit un scandal între două grupuri rivale de tineri.

[8] "Soţia mea m-a chemat şi am intervenit ca să liniştesc apele.

[9] Am încercat să stau de vorbă, să-i calmez, dar **băiatul acela** m-a lovit în piept şi era cât pe ce să [...].

[4] Poliţia a stabilit că **tânărul împuscat, Vasile Mihai**, pe fondul consumului de alcool, **pro** a fost implicat într-un scandal, iar viceprimarul a intervenit pentru a-**l** stopa.

(no further co-referential expressions)

Before analyzing the persistence of the direct objects, it is also important to underline the fact that in (10), it is the other man, Neculai Florea, who is the topic of the discourse, and not the *pe*-marked DO. Furthermore, it is also worth noting that the two examples in (10) and (11) do not contrast in their epistemic specificity.

A striking observation with respect to DOM is the fact that the *pe*-marked direct object in (10) displays a higher discourse prominence than the direct object which is not preceded by *pe* in the discourse, because it displays the potential to generate further co-referential expressions. This feature of DOM marked indefinite direct objects is underlined on the one hand by the fact that the referent of this object is taken up in the next nine sentences 8 times, while the referent of the not *pe*-marked direct object in (11) was mentioned again in the next eleven sentences only 3 times.

The structures of the above given examples are:

(12)   Structure of (10):
        $[_1 ... $ **pe** indef NP$]$ $[_2 $ $]$ $[_3$ PN, pro, pro$]$ $[_4$ $]$ $[_5$ PN, pers.pron, PN $]$ $[_6$ def NP$]$ $[_7 ..]$ $[_8 ..]$
        $[_9$ def NP$]$ $[_{10}$ pron$]$ $[_{11}$ pron$]$

(13)   Structure of (11)
        $[_1$ indef NP$]$ $[_2$ $]$ $[_3$ $]$ $[_4 ...$ (def NP + Adj+ PN), pro, CL$]$

On the other hand, the discourse prominence of the *pe*-marked direct object is evidenced by the first anaphoric item. In article (10), the newly introduced referent *un tânăr* ('a young man') is taken up in the following discourse by a proper name. However, a proper name can be opted for only in cases in which the presupposition licensed by the proper name can be accommodated within the context. This does not hold for the second article (11), in which the referent of the unmarked direct object is mentioned again by means of the definite modified NP *tânărul împuşcat* ('the young man that was shot').

The next table in (14) is a modified version of the table presented under (7) above.

Besides the distribution of *pe*-marking along the Referentiality scale, the table also contains the newly introduced factor "referential persistence":

(14)   Referentiality Scale for *pe*-marking in Romanian for human direct objects

Ref Scale Disc Prom	*pers. pron.*	> *PN*	> *def. NP*	> *indef NP*		> *non-arg NP*
				spec.	*non-spec*	
topic	+	+	+	+	+	n.a.
ref persistence	+	+	+	+	n.a.	n.a.
non-prominence	+	+	+ (±)	-	-	–

Indefinite specific objects which are important for the upcoming discourse are characterized through a high persistence and will therefore be marked by *pe*. Accordingly, indefinite specific objects which are not that relevant for the discourse in question will not be taken up too often in the subsequent discourse. The lack of prominence of such objects is formally expressed by the absence of *pe*.

5. CONCLUSION AND OPEN QUESTIONS

As I have showed in this paper, *pe*-marking expresses different functions. One of them is to indicate a higher activation in terms of referential persistence of the direct object marked in this way. To assume that the direct object preceded by *pe* is more activated, is a weaker claim than to assume that DOM-marking expresses topicality.

While "referential persistence" characterizes definite nominal phrases in direct object position as well (see von Heusinger & Chiriacescu 2008, to appear), it is still open whether this feature is a property that applies only to synchronic Romanian data, or if it also holds for diachronic texts. Furthermore, several problems of the empirical base of the hypothesis still remained unresolved. One of these problems might be the fact that there still exist other parameters that could interact with *pe*-marking (as for example different verb classes (see von Heusinger 2008 for Spanish). Another encountered problem is the fact that I could find only a limited number of instances of *pe*-marking with indefinite direct objects under "controlled conditions" as in the examples (10) and (11) above.

References

Ariel, Mira. 1988. Referring and Accessibility. Journal of Linguistics 24. 65-87.

Chiriacescu, Sofiana. 2007. Pe-Markierung und Diskurs-Prominenz im Rumänischen. Magisterarbeit.
        Universität Stuttgart.

Cornilescu, Alexandra. 2001. Direct Objects at the Left Periphery, the Case of Romanian. In: Bucharest Working
        Papers in Linguistics 3, Bucharest. 1-18.

Dobrovie-Sorin, Carmen. 1994. The Syntax of Romanian. Berlin: Mouton deGruyter.

Farkas, Donka. 1978. Direct and indirect object reduplication in Rumanian. In Papers from the Seventeenth
        Regional Meeting of the Chicago Linguistics Society (CLS 14), 88-97. University of Chicago, Chicago.

Givon, Talmy. 1981. On the Development of the Numeral 'one' as an IndefiniteMarker. Folia Linguistica
        Historia 2. 35-53.

Givon, Talmy. 1983. Topic Continuity in Discourse. An Introduction. In: T. Givón (ed.). Topic Continuity in
        Discourse. A Quantitative Cross-Language Study. Amsterdam; Philadelphia: Benjamins. 1-41.

Gramatica Limbii Romane. 2005. Vol I. Cuvantul. Bucuresti: Editura Academiei.

Gundel, Jeanette & Hedberg, Nancy & Zacharski, Ron. 1993. Cognitive Status and the Form of Referring
        Expressions in Discourse. Language 69. 274-307.

von Heusinger, Klaus. 2003. The Double Dynamics of Definite Descriptions. In: J. Peregrin (ed.). Meaning in
        the Dynamic Turn. Amsterdam et al.: Elsevier. 150-168.

von Heusinger, Klaus. 2008.Verbal Semantics and the Diachronic Development of Differential Object Marking

in Spanish. Probus 20.

von Heusinger, Klaus & Onea, Edgar. 2008. Triggering and Blocking Effects in the Diachronic Development of DOM in Romanian Probus 20.

Ionin, Tania. 2006. This is definitely specific: specificity and definiteness in article systems. Natural Language Semantics. Springer. 14. 175-234.

Kamp, Hans and Agnes Bende Farkas. 2006. Specific Indefinites: Anchors and Functional Reading. Ms. Universität Stuttgart.

Malchukov, Andrei & Helen de Hoop. 2007. On fluid differential case marking: A bidirectional OT approach. Lingua, Vol. 117, Issue 9. 1636-1656.

Mardale, Alexandru. 2002. Note despre construirea obiectului direct prepositional in Romana si Spaniola. In: Studii si cercetari lingvistice 1-2. 77-94.

Niculescu, A 1965.Obiectul direct prepozitional in limbile romanice. In: Individualitatea limbii romane intre limbile romanice. Bucuresti: Editura Stiintifica.

Prince, Ellen. 1981. Toward a taxonomy of given-new information. In:P. Cole (ed.). Radical Pragmatics. New York: Academic Press. 223-255.

Stark, E & Sora Sanda. 2008. Why is there differential object marking in Romance? 29th Annual Meeting of the German Society for Linguistics in Bamberg, Germany.

Wright, S. & Givon, Talmy. 1987. The pragmatics of indefinite reference: quantified text-based studies. Studies in Language 11. 1-33

Indefinite NPs and *pe*-marking in Romanian
(Abstract)

The fact that in Romanian a direct object is sometimes morphologically marked by the particle *pe* and sometimes not, is a long attested phenomenon (Niculescu 1965, Chiriacescu 2007, von Heusinger & Onea 2008, among others). Most studies on Differential Object Marking (DOM) explained most occurrences of *pe* as a case marker in terms of animacy, definiteness and specificity. However, there still remained instances which could not be accounted for by means of the above mentioned parameters. These not elucidated cases are those in which a direct object realized as an unmodified definite or indefinite nominal phrase are optionally marked, while the difference in meaning between the two constructions is subtle.

Focusing on examples in which the direct object is realized as an indefinite unmodified nominal phrase, I underline the necessity of a more detailed analysis of the principles involved in *pe*-marking by taking into account the discourse context in which such indefinites occur. I will show that differentially marked direct objects receive a preferential treatment in the perception and production of a discourse in comparison to their unmarked counterparts. *Pe*-marking displays the property of "referential persistence" (the number of occurrences of co-referential expressions) of a referent introduced by an indefinite direct object in the discourse. Consequently, I propose the introduction of an additional discourse- based parameter to account for *pe*-marked indefinite direct objects.

Sofiana Chiriacescu
Institut für Linguistik/ Germanistik
Universität Stuttgart
Heilbronner Str. 7
70174 Stuttgart
Sofiana.chiriacescu@ling.uni-stuttgart.de