THE DISCOURSE STRUCTURING POTENTIAL
OF DIFFERENTIAL OBJECT MARKING, THE CASE
OF INDEFINITE AND DEFINITE DIRECT OBJECTS
IN ROMANIAN

KLAUS VON HEUSINGER and SOFIANA CHIRIACESCU

Abstract. This paper investigates pe-marking of postverbal human direct objects in Romanian as an instance of differential object marking (DOM). Specifically, we compare the conditions under which unmodified definite noun phrases are pe-marked with the conditions under which indefinite noun phrases are pe-marked. We argue that pe-marking does not only signal the referential property of specificity for indefinites and referentiality for definites, but that it also signals the discourse-pragmatic property of Discourse Structuring Potential (DSP). DSP is reflected by (i) a higher likelihood of subsequent mention (Givón 1983, Ariel 2001, Arnold 2010) and (ii) a higher topic shift potential (Givón 1983). Discussing the findings from different sentence-continuation experiments we conducted, we argue that pe-marked unmodified definite noun phrases and pe-marked indefinite noun phrases show a higher DSP than their unmarked counterparts.

Keywords: Differential object marking, specificity, Discourse Structuring Potential, discourse prominence, sentence-continuation task.

1. DIFFERENTIAL OBJECT MARKING IN ROMANIAN-OBSERVATIONS AND RESEARCH ISSUES

Pe-marking in Romanian is an instance of differential object marking (DOM), i.e. the marking of the direct object under certain conditions. It is commonly assumed that pe-marking (generally accompanied by clitic doubling) is obligatory with (modified) definite human postverbal noun phrases, as in (1) and with all other forms that are higher on the Referentiality Scale, while it is optional with indefinite noun phrases (2) (Niculescu 1965, Pană-Dindelegan 1997, von Heusinger, Onea 2008, Stark, Sora 2008, Ciovârnache, Avram 2013).

1 Universität zu Köln, Klaus.vonHeusinger@uni-koeln.de; Universität zu Köln and Transilvania University, Brașov, Sofiana.Chiriacescu@uni-koeln.de. We would like to thank Alexandra Cornilesescu, Daniela Dinu, Alina Tigau and their team from the English Department at the University of Bucharest for organizing a wonderful workshop, for editing the present volume and for helpful suggestions. We are grateful to the audience of the workshop The Syntax and Interpretation of Specificity held in December 2012 in Bukarest for constructive comments and considerable assistance. Our research was supported by the German Science Foundation by a grant to the project C2: Case and referential context, as part of the Collaborative Research Center 732 Incremental Specification in Context at the University of Stuttgart.

(1) a. Doctorul îl examinează pe băiatul bolnav.  
   Doctor.DEF CL examines PE boy.DEF sick  
   ‘The doctor examines the sick boy.’

b. # Doctorul examinează băiatul bolnav.  
   Doctor.DEF examines boy.DEF sick  
   ‘The doctor examines the sick boy.’

(2) a. Toţi bărbaţii -o iubesc pe o femeie.  
   All men CL love PE a woman  
   ‘All men love a woman.’ (specific / wide scope)

b. Toţi bărbaţii iubesc o femeie.  
   All men love a woman  
   ‘All men love a woman.’ (specific / wide scope vs. non-specific / narrow scope)

The contrast in (2) is generally associated with specificity. While the pe-marked indefinite direct object in (2a) has only a ( scopally) specific interpretation in the sense that there is one woman such that all men love her, the unmarked indefinite in (2b) allows for a specific or a non-specific reading (Farkas 1978, Dobrovin-Sorin 1994). (It is more accurate to say that pe-marking with indefinites excludes a non-specific reading while unmarked indefinite direct objects are compatible with a specific reading.) There is an interesting exception in the otherwise quite robust distribution of pe-marking with definite noun phrases. Modified human definite direct objects obligatorily receive pe-marking, as in (1a), but the co-occurrence of pe and the definite article is blocked if the noun phrase is not modified. This blocking effect on pe-marking derives from an independent syntactic rule of Romanian which holds for most prepositions, e.g. the preposition la ‘to’ in (3A), but also for pe in its case-marking function, as in (3a). There are two alternative constructions for the ungrammatical sentence (3a), namely (3b), in which the definite article is present and the pe-marker (and the clitic) are absent, and (3c), in which the pe-marker (and the clitic) precedes the noun phrase in the absence of the definite article, yielding a definite reading.

(3) A. Un băiat merge la doctor.  
   a boy goes to doctor.

a. *Doctorul îl examinează pe băiatul.  
   Doctor.DEF CL examines PE boy.DEF  
   ‘A boy goes to the doctor. The doctor examines the boy.’

b. Doctorul examinează băiatul.  
   Doctor.DEF examines boy.DEF  

c. Doctorul îl examinează pe băiat.  
   Doctor.DEF CL examines PE boy  

We can summarize our observations made so far: The alternation of pe-marking with postverbal human direct objects is not restricted to indefinite NPs, but also applies to definite (unmodified) NPs. This observation raises the following main questions:

(i) The function of pe for indefinites is generally associated with the notion of specificity. Can we assign an analogous function to pe-marked definites as for example signaling the contrast between a referential and an attributive reading in the sense of Donnellan (1966)?
(ii) Even if we can find a parallel feature for definites, the contrast between (3b) and (3c) cannot be accounted for in terms of scope or specificity. Thus, we have to find out whether there are other features that are associated with pe-marking.

(iii) A more general question regarding pe-marking and DOM in other languages as well is whether the marker has a genuine (lexical) function, or whether the expressed function is derived from the particular construction and some additional inferences.

In this paper we argue that: a) pe-marking has the same or a very similar function for definites and indefinites in contexts with operators; b) pe-marking expresses specificity (or more exact: incompatibility with non-specificity) for indefinites, and referentiality for definites; c) specificity cannot account for contrasts in transparent sentences, i.e. simple declaratives with no operators; d) in transparent sentences, pe-marking expresses the discourse pragmatic function of Discourse Structuring Potential (DSP) for both, definites as well as indefinites. DSP is reflected by (i) a higher likelihood of subsequent mention (Givón 1983, Ariel 2001, Arnold 2010) and (ii) a higher topic shift potential (Givón 1983). Following the literature (Prince 1981, Ionin 2006) we assume that specificity and DSP are closely related and that both functions are lexically encoded in pe (similar to the lexical contribution of articles, specific adjectives like a certain, or discourse particles like then, therefore, etc.).

The paper is structured as follows: In section 2, we provide a brief overview on the conditions of pe-marking in Romanian and discuss the alternation between pe-marking and the use of the definite article for unmodified definite direct objects, in particular. In section 3, we introduce the discourse-pragmatic concept of Discourse Structuring Potential, and in section 4, we report the findings of two sentence-continuation experiments that tested the discourse behaviour of indefinite and definite noun phrases. In section 5, we discuss these findings and present a general perspective for pe-marking in Romanian and DOM in general.

2. PE-MARKING AND THE REFERENTIALITY SCALE

There is some consensus in the literature that the most important synchronic conditions triggering DOM in Romanian are animacy, definiteness, specificity and topicality (Farkas 1978, Dobrovie-Sorin 1994, Cornilescu 2001, von Heusinger, Onea 2008, Kamp, Bende-Farkas (submitted), among others). We focus on definiteness and specificity, which are mapped to the Referentiality Scale (Aissen 2003 among others), as illustrated in Table 1. We restrict our investigation to human direct objects in postverbal position.

<table>
<thead>
<tr>
<th>pe-marking for human postverbal direct objects</th>
<th>pers. pron.</th>
<th>&gt; PN</th>
<th>&gt; def. NP</th>
<th>&gt; spec.indef NP</th>
<th>&gt; non-spec. indef NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>modified</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>unmodified</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
</tbody>
</table>
Full personal pronouns referring to animate entities are always marked with *pe* and doubled by a clitic in synchronic Romanian\(^2\), as (4) shows. Proper names referring to humans are always *pe*-marked, as in (5). Modified human definite NPs in direct object position are generally *pe*-marked, like in (6a) while the form without *pe* is rather marginal, as in (6b).

(4) Maria îl asculta pe el.
'Mary listens to him.'

(5) Am vazut-o pe Maria.
'I have seen Mary.'

(6) a. Doctorul îl examinează pe băiatul bolnav.
'The doctor examines the sick boy.'

b. #Doctorul examinează băiatul bolnav.
'A boy goes to the doctor. The doctor examines the sick boy.'

2.1. Indefinite NPs

*Pe*-marking of indefinite human direct objects is optional and the literature (Farkas 1978, Dobrovin-Sorin 1994) assumes that specificity is the main triggering parameter for *pe*-marking. Following Farkas (1994) and von Heusinger (2011) we can distinguish between different kinds of specificity. We restrict the discussion of specificity to *scopal specificity*, as in (7), specificity in opaque contexts (*referential specificity*) as in (8), and *epistemic specificity* in transparent contexts, as in (9). Scopal specificity with extensional operators and referential specificity with intensional operators triggers *pe*-marking. While the sentence (7a) is ambiguous between a specific (or wide scope) reading and a non-specific (or narrow scope) reading, the non-specific reading in (7b) is ruled out due to the presence of *pe* (Dobrovie-Sorin, 1994). The same variation in readings between specific and non-specific readings is maintained for constructions with intensional operators, like in (8). To be more exact: non-specificity blocks the appearance of *pe* (see also Tigau 2012 for Romanian; Lenoetti 2004, Lopez 2012 for Spanish).

(7) a. Toţi bărbaţii iubesc o femeie.
'All men love a woman.' (specific vs. non-specific)

\(^2\) DOM in Romanian is generally accompanied by clitic doubling, i.e. the occurrence of a co-indexed weak pronoun. A doubling clitic is optional, obligatory or blocked, depending on semantic features of the head noun and further syntactic constraints. In this paper we will address the phenomenon of DOM in Romanian as a whole, thus, in the following sections, will not make an explicit distinction between clitic doubling and *pe*-marking (but see Dobrovie-Sorin 1994, *Gramatica limbii române* 2005).
b. Toţi bărbaţii -o iubesc pe o femeie.
   ‘All men love a woman.’ (specific)

(8) specificity in opaque contexts (referential specificity)

a. Ion caută o secretară.
   ‘John looks for a secretary.’ (specific vs. non-specific)

b. Ion o caută pe o secretară.
   ‘John looks for a secretary.’ (specific)

Thus, (8a) could have an interpretation where John is looking for a particular secretary (the specific reading) as well as one where any secretary will do (the non-specific reading), while (8b) only allows for the specific interpretation. It seems that we cannot extend this contrast to epistemic specificity, as illustrated in (9). Epistemic specificity is understood as the knowledge of the speaker about the identity of the referent. If pe showed a contrast with respect to epistemic specificity, we would expect the referent associated to the indefinite pe un prieten (‘pe a friend’) in (9b) to be known by the speaker, while the sentence (9a) would allow both for an epistemic specific and an epistemic non-specific reading.

(9) epistemic specificity

a. Petru a vizitat un prieten.
   ‘Petru visited a friend.’

b. Petru l-a vizitat pe un prieten.
   ‘Petru visited a friend.’

Clear judgments for these contexts are difficult to get. It seems that both sentences are compatible with a continuation like: (i) *I do not know the friend*, or (ii) *I do know the friend*. It thus seems that the contrast in (9) is due to a different feature that cannot be thoroughly captured by epistemic specificity. We assume that the formal alternation between the two forms can be associated with the weaker discourse-pragmatic property Discourse Structuring Potential. We will elaborate upon this property in section 3 and 4.

2.2. Definite unmodified NPs

Definite modified direct objects are *pe*-marked in most cases. However, unmodified definite direct objects also show an optional *pe*-marking due to a grammatical rule, which blocks the co-occurrence of the enclitic definite article with *pe* as in (10a). In (10), the discourse referent is given, definite, referential, specific, accessible etc. and can be picked up by the definite noun phrase with the enclitic article in (10b) or by *pe* and the “bare” noun (and clitic doubling) in (10c). Informants do not have clear intuitions about the contrast between the readings of (10b) and (10c) and there are not many similar instances in corpora, where such an option is available.
(10) Context: Un băiat merge la doctor. (A boy goes to the doctor.)

a. *Doctorul îl examinează pe băiatul.
   Doctor.DEF CL examines PE boy.DEF

b. Doctorul examinează băiatul.
   Doctor.DEF examines boy.DEF

c. Doctorul îl examinează pe băiat.
   Doctor.DEF CL examines PE boy

‘The doctor examines the boy.’

The alternation between *pe*-marking and the lack of the definite article vs. the lack of *pe*-marking and the definite article is further restricted by various blocking factors. Here, we provide the possessive dative as one example for illustration (see for more examples Chiriacescu 2007, von Heusinger, Onea 2008, von Heusinger, Chiriacescu 2011). At sentence level, *pe*-marking is ruled out whenever the definite article is modified by a possessive preverbal (11a) or postverbal dative (11b), even in cases where the NP is further modified by an adjective (see for a similar blocking effect on DOM in Spanish by possessive datives, Bruggé, Brugger 1996: 29–30).

(11) a. Maria îşi înţelege *pe buna prietenă.
   Maria.DAT understands PE good.DEF friend

   ‘Maria understands her good friend.’

b. Întelegându-şi *pe frumoasa soţie a făcut […].
   understanding-DAT PE beautiful.DEF wife has made

   ‘Understanding his beautiful wife, he made […].’

In the following we tested whether the typical semantic-pragmatic parameters of definite noun phrases, i.e. (i) anaphoricity, (ii) uniqueness vs. familiarity, (iii) scope, (iv) referential readings in opaque contexts and (v) referential vs. attributive readings in transparent contexts, influence the use of *pe*-marking.

An anaphoric definite noun phrase is linked to its coreferent antecedent, as in (12). In a small survey we found various examples of the form in (12) where we replaced one form for the other, but we did not find any significant difference in meaning between the alternate forms.

(12) Context: Ion a cunoscut un politician şi un scriitor.
   ‘John met a politician and a writer.’

a. A doua zi a văzut politicianul la televizor.
   the next day HAS seen politician.DEF on TV

b. A doua zi l-a văzut pe politician la televizor.
   the next day CL HAS seen PE politician on TV

   ‘He saw the politician on TV next day.’

The following context allows us to make an interesting distinction with respect to (situational) salience or familiarity on the one side and descriptive uniqueness on the other. Both sentences (13a) and (13b) are felicitous in the given context, however, they give rise to different interpretations on the side of the hearer, i.e. whether he can see the escaping
prisoner or not. In the *pe*-marked version (13a) the hearer is not only informed of the existence of the escaped prisoner, but he is also instructed to locate the referent in the immediate situation of utterance. If the policeman within the prison utters sentence (13a), then the policeman outside the prison must see the runaway. Otherwise, he would ask a *wh*-question to find out which prisoner escaped, where he went, etc. However, if the first policeman utters sentence (13b) with the definite article and without *pe*, the hearer does not have to see the runaway to understand what happened and how he should react. He would have to look where (and who) the (unique) runaway actually is.

(13) Context: A policeman is standing guard outside a prison, which is surrounded by a twenty-foot wall. Suddenly, he hears the voice of a colleague policeman from the other side:

a. Prinde *pe* fugar.
catch CL PE runaway

b. Prinde fugarul.
catch runaway.DEF

‘Catch the runaway.’

Definite noun phrases generally show wide scope, but we can construct cases where definite noun phrases receive narrow scope with respect to a universal quantifier, as in (14). Here we find an interesting contrast: In (14a), the phrase *proprietarul* (‘the owner’) could be interpreted as ‘Each renter salutes his/her owner’, even if the noun phrase is further modified by the adjective *bogat* (‘rich’). However, the *pe*-marked definite in (14b) clearly underlines the fact that the mentioned owner is the same for each renter. In both readings the definite description expresses a uniqueness presupposition, but with respect to different domains. The definite article allows for local domains created by the universal quantifier, while the *pe*-marked version only allows for wide scope, similar to demonstrative expression and to the scopal specificity of indefinite noun phrases discussed above.

(14) a. Toţi chiriaşii salută proprietarul bogat.
all renters salute owner.DEF rich

‘All renters salute the rich owner.’

b. Toţi chiriaşii îl salută pe proprietar / pe proprietarul bogat.
all renters CL salute PE owner / PE owner.DEF rich

‘All renters salute the owner / the rich owner.’

Definites in opaque contexts show different readings - they can have a *de re* or a *de dicto* reading (Quine 1956). We cannot go into the details of the analysis of these constructions at this point, but a simplistic interpretation would suggest that the contrast could be reconstructed with scope. For the *de re* reading, the noun phrase takes scope over the intensional operator *seek*, while in the *de dicto* reading, the noun phrase takes narrow scope. Example (15b) with the *pe*-marked definite expresses a *de re* reading, i.e. there is an administrator such that Peter is looking for him or her. The property reading is excluded with *pe* in such contexts, but not in others (see Cornilescu 2013 for a discussion of objects of reflexives and middles). The form *administratorul* without *pe* and with the definite article has a *de re-* and *de dicto* reading. The latter one refers to a situation in which Peter is seeking an administrator, whoever s/he might be.
The classical contrast of definite noun phrases in transparent sentences, i.e. in a simple declarative sentence without further operators, is one between a referential and an attributive reading (Donnellan 1966). In the referential reading, the speaker has a particular individual in mind, while in the attributive reading, any individual that matches the descriptive content qualifies as a good referent. This contrast is very similar to the (epistemic) specific vs. non-specific readings of indefinites (Partee 1970):

(16) a. La inaugurarea aeroportului Braşov, cetăţenii vor invita primarul.  
At inauguration. DEF airport. DEF Braşov citizen. DEF will invite mayor. DEF

b. La inaugurarea aeroportu lui Braşov, cetăţenii îl vor invita pe primar.  
At inauguration. DEF airport. DEF Braşov citizen. DEF CL will invite PE mayor

‘At the inauguration of the Brasov airport, the citizens will invite the mayor.’

Both (16a) and (16b) are felicitous in this context, but with different readings. (16a) refers to whoever person might occupy the mayor position at the time the airport will be inaugurated (a particular function). In (16b), the pe-marked noun phrase is not tight to the function its associated referent designates, but to the individual that occupies this position, say Mr. Jones. Note that only functional nouns like mayor, president, murderer etc. show a clear contrast between a referential and an attributive reading. This observation does not hold for sortal concepts like boy as in example (3), for which another explanation is needed.

2.3. Comparing conditions for DOM with indefinites and definites

We can now compare the referential functions expressed by pe-marking of indefinite direct objects with those referential functions expressed by pe-marking of definite direct objects. While the alternation is well documented and discussed for indefinites, there is not very much work on the alternation with definites. In Table 2, we summarized the observations of the last two subsections. Both pe-marked and unmarked definites can be anaphoric, but only pe-marked definites refer to referents that are directly perceivable, while unmarked ones are uniquely identified by their descriptive content. These observations hold for definite noun phrases only. Pe-marked indefinites and definites show wide scope, while unmarked noun phrases can have wide or narrow scope (at least in the case of the indefinites). Pe-marked indefinites and definites show referential or wide scope (or de re) readings in opaque contexts, and they tend to get a referential reading in transparent contexts. Again, this analysis seems more appropriate for definites than for indefinites.

The overview presented above includes transparent contexts in which we find definite and indefinite direct objects with and without pe-marking. While informants report that there are differences in readings, we cannot find one clear referential property that determines this contrast. We therefore assume that the alternation between the pe-marked
and the unmarked form is due to the discourse function of the (in)definite noun phrase. Definite and indefinite noun phrases do have a “cataphoric” force or a “forward looking function” and thus express a certain (additional) level of activation of the referent they are associated with (see von Heusinger 2007), in a way explained in the next section.

Table 2

Referential properties expressed by pe for indefinites and definities

<table>
<thead>
<tr>
<th></th>
<th>anaphoric</th>
<th>immediate situation</th>
<th>scope</th>
<th>opaque context</th>
<th>transparent context</th>
</tr>
</thead>
<tbody>
<tr>
<td>pe + indef. N</td>
<td>n.a.</td>
<td>n.a.</td>
<td>wide</td>
<td>wide</td>
<td>Spec vs. non-spec</td>
</tr>
<tr>
<td>Ø + indef. N</td>
<td>n.a.</td>
<td>n.a.</td>
<td>narrow</td>
<td>narrow vs.</td>
<td>spec vs. non-spec</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>vs. wide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pe + N-Ø</td>
<td>anaphoric</td>
<td>familiar vs. visible</td>
<td>wide</td>
<td>wide</td>
<td>referential</td>
</tr>
<tr>
<td>Ø + N+def.</td>
<td>anaphoric</td>
<td>unique</td>
<td>narrow</td>
<td>narrow</td>
<td>attributive</td>
</tr>
</tbody>
</table>

3. FORWARD LOOKING FUNCTION AND DISCOURSE STRUCTURING POTENTIAL

A body of linguistic and psycholinguistic research has investigated various factors that influence the comprehension and production of different types of referring expressions (Givón 1983, Gundel, Hedberg, Zacharski 1993, Kehler et al. 2008, Arnold 2010). The majority of these studies focused on anaphora resolution, as it is commonly assumed that reduced referring expressions correlate with highly accessible or prominent entities. Furthermore, to determine the accessibility of a referent, researchers have generally employed a backward-looking perspective, determining the factors that license the usage of a particular type of referring expression at a particular stage in the discourse. In other words, given a certain type of referring expression (e.g. a pronoun), the factors that license its use were investigated.

In contrast to personal pronouns, which refer to previously mentioned and focussed entities, definite noun phrases display different kinds of forward-looking referential properties: First, they can be used for discourse-new entities or for entities introduced by a bridging or inference relation as well (Hawkins 1978, Vieira, Poesio 2000). Second, both types of definite noun phrases (familiar ones and first mentioned ones) change the accessibility or the activation of the associated discourse referents (von Heusinger 2003, 2007). Indefinite noun phrases are more often related to a “forward looking function” as their main function is to introduce a new discourse item that can be used as an antecedent for subsequent anaphoric terms (Karttunen 1976, Heim 1982, Kamp 1981/2013). In this study we extend this forward-looking perspective and test the effects of production-driven biases licensed by pe-marked and unmarked indefinite and definite direct objects in Romanian. We investigate the Discourse Structuring Potential (DSP), which can be best measured by two textual characteristics that pertain to the following discourse (Chiriacescu, von Heusinger 2010, Chiriacescu 2011, Deichsel, von Heusinger 2011).

The first metric, referential persistence, reflects the likelihood that a particular referent will be picked again in the ensuing discourse (Givón 1983, Kehler et al. 2008). The second metric for DSP, topic-shift potential, is defined in terms of the likelihood that a
referent will be mentioned in grammatical subject position. We focus on the subject position because different linguistic and psycholinguistic studies (e.g. Crawley, Stevenson 1990) have shown that referents mentioned in the syntactic subject position are more salient or accessible in a given discourse than referents mentioned in other syntactic positions (e.g. as direct or indirect direct objects). For the sake of simplicity, the first instance in which a direct object referent becomes the grammatical subject in a matrix clause is treated as an instance of topic shift. Despite being mentioned in a rather non-preferential grammatical position (i.e. as a direct object), we expect pe-marked definite and indefinite noun phrases to show higher values for both metrics than their non-pe-marked counterparts.

4. WEB-BASED EXPERIMENT FOR DSP

To investigate whether the presence of the pe-marker boosts the prominence or salience of the referents associated with the direct objects realized as definite unmodified noun phrases, we used the metrics for discourse prominence developed for the experiment with indefinite noun phrases (cf. Chiriacescu, von Heusinger 2010). More precisely, we analysed whether pe-marked definite noun phrases are (i) referentially persistent in the subsequent discourse (i.e. whether the referent headed by pe is likely to be continued), and (ii) more susceptible to shift the topic (i.e. in the sense of Givón 1983, Ariel 2001, among others) of the current discourse. Let us now consider how we predict participants’ responses to pattern with respect to the two different metrics we tested. First, in light of the findings from the pe-marking experiment with indefinite noun phrases (Chiriacescu, von Heusinger 2010) and other experimental investigations, which showed that accessible or salient referents are more likely to be subsequently mentioned (Givón 1983, Gernsbacher, Shroyer 1989, Arnold 1998, among others), we predict that referents headed by pe will be referentially more persistent in the ensuing discourse, compared to referents marked with the simple definite article.

Second, given the observation that important or salient referents tend to be mentioned in topic position (which in English generally corresponds to the grammatical subject position, e.g. Ariel 2001, Arnold 1998), we predict that in comparison to their unmarked counterparts, pe-marked direct objects will (i) be mentioned more often in the subsequent text, and will (ii) become the new topic in the following discourse.

4.1. Method

Participants: Twenty native speakers of Romanian participated in the experiment on pe-marking with indefinite noun phrases and other twenty native speakers of Romanian participated in the experiment on pe-marking with definite noun phrases. They received no incentive for taking part in the survey. It took about twenty minutes to complete an experiment.

3 Independently of these two textual characteristics, we considered the type of referring expression used to pick up the referent of the direct objects. We will not discuss the findings of this metric in this paper, but note that the likelihood of subsequent mention does not point in the same direction as the likelihood of being realized with a pronoun (see also the discussions in Kehler et al. 2008, Kaiser 2010, Chiriacescu, von Heusinger 2010). For the purposes of this paper, just note that pronominalization does not reflect the salience or discourse prominence to the target referents.
Materials: The methodology used in this experiment was an open-ended sentence-continuation task. Participants were presented target items consisting of mini-discourses, as in table 4 and 5. Their task was to read the given story fragments and add five logical and natural-sounding sentence continuations for each of them. The first two sentences of each test item set the context of the story, and contained individual references to two characters.

The first character was the clearly established topic of the mini-discourse, as it was mentioned in subject position at least once and was the referent the story was about. In the last sentence of each mini-discourse, the critical referent was introduced as an indefinite or definite noun phrase in direct object position. We only manipulated the morphological realisation of the target referents, which resulted in two conditions for each experiment, i.e. one condition in which *pe* heads an indefinite unmodified noun phrase (in Exp1) and a definite unmodified noun phrase (in Exp2) (see the left columns of Tables 3 and 4 below), and another condition in which the same direct objects are not *pe*-marked, i.e. they are headed by the simple indefinite article in Exp1 and followed by the enclitic definite article in Exp2 (see the right columns of Tables 3 and 4).

Table 3

<table>
<thead>
<tr>
<th>Sample experimental items from Exp1 on indefinite NPs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>pe-condition</strong></td>
</tr>
<tr>
<td>Aseară a fost extraordinar de cald. Pentru că nu mai rezista în casă, Graur s-a hotărât să iasă în oraș. Pe drum l-a văzut pe un copil intrând într-un magazin. “It was extraordinarily warm outside yesterday evening. Because it was unbearable for him to stay home anymore, Graur decided to go downtown. On his way there he saw a child entering a store”.</td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th>Sample experimental items from Exp2 on definite NPs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>pe-condition</strong></td>
</tr>
<tr>
<td>La petrecerea de aseară, Andrei a cunoscut un politician şi un cântăreţ de renume. Astăzi l-a întâlnit pe politician în piaţă. “At yesterday evening’s party, Andrew met a politician and a famous singer. Today he met politician at the market.”</td>
</tr>
</tbody>
</table>

Procedure and data analysis: The first five main clauses (including subordinate ones, if there were any) of each continuation story provided by the participants were analysed. Two independent judges coded for two aspects of the direct objects: (i) their referential persistence and (ii) their topic shift potential. We coded 10 continuations for each condition of the two experiments (i.e. 10 responses for *pe*-marking with indefinite NPs and another 10 responses for non-*pe*-marking with the simple indefinite article in Exp1. In Exp2 we coded 10 responses for *pe*-marking with definite NPs and 10 responses for non-
pe-marking). The global topic of the first and second sentence (e.g. Graur and Andrei in the test items in Table 3 and Table 4) received Subscript 1. Subscript 2 was used for the noun phrase whose form was manipulated in the critical sentence (e.g. pe un copil or pe copilul (‘pe a child’ in Exp1 or ‘pe the child’ in Exp2) in the left columns of Tables 4 and 5. Example (17) represents an example response for test item 3 for the pe-condition, and Table 5 illustrates the coding methods used.

(17) Example responses and coding methods from the story continuation experiment

La petrecerea de aseară, Andrei1 a cunoscut un politician2 și un cântăreț de renume3. Astăzi (pro)1 l-a întâlnit pe politician2 în piată.

‘At yesterday evening’s party, Andrew1 met a politician2 and a famous singer3. Today he1 met the politician2 at the market.’

S1: (pro)1 știe ca asta este șansa lui1,

‘He1 knew that that’s his1 chance.’

S2: Politicianul2 era un pic grizonant, slăbuț, cu accent baritonal.

‘The politician2 had some greyish hair, was thin with baritone voice.’

S3: Andrei1 s-a dus spre el2, și (pro)1 i2-a cerut ajutorul să (pro)1 alegă un pepene bun.

‘Andrei1 went towards him2 and he1 asked (him2) for help to choose a tasty water melon.’

S4: Politicianul2 s-a întors și (pro)2 l-a răspuns cu un aer distrat.

‘The politician2 turned around and (pro)2 responded him1 in a distracted voice.’

S5: Il2 chema don Giuseppe și (pro)2 era inginer zootehnist de meserie.

‘His2 name was don Giuseppe and he2 was a zootechnician engineer.’

Table 6

<table>
<thead>
<tr>
<th>Coding methods</th>
<th>First referent (Subject)</th>
<th>Target referent (Object)</th>
<th>Refer per item / S</th>
<th>Sum</th>
<th>Refer per item / S</th>
<th>Sum</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 Anaphoric forms and grammatical function [pro1] (pron1)</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>Topic 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Sub1] (IO1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2 [def NP2]</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>Topic 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Sub2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3 [PN1, pron2] [pro1, CL2] (pro1)</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>Topic 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Sub1, PP2] [Sub1, IO2] (Sub1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4 [def NP3] [pro2, CL1]</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>Topic 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Sub3] [Sub2, IO1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5 [CL2] [pro3]</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>7</td>
<td>Topic 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[DO2] [Sub2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Referential persistence is measured by referents mentioned per sentence (item / S) and the sum of all items up to S5 (i.e. a cumulative measure). Comparing the sums indicates at what stage in the discourse we have more anaphoric expressions referring to one referent compared to another. We verified in what sentence the target referent (i.e. the politician) becomes the subject and topic of a main clause. In example (17), this happens in sentence continuation 2 (S2).

4.2. Results

40 participants provided continuations for the initial story fragments. The results from the two metrics, referential persistence and topic shift potential reflect the discourse status of the stories’ referents. In the following sections we discuss the findings of the two textual characteristics in detail.

4.2.1. Referential persistence

The first textual characteristic investigated was referential persistence. Figure 1 displays the mean values for referential persistence of all referents of the test items in Exp1 with indefinite noun phrases. For the pe-condition, we notice a strong likelihood of the referent to me mentioned in the following discourse. On the contrary, the direct object referents in the non-pe-marked condition are picked up in the subsequent discourse less often.

![Fig. 1. Referential persistence of object referents in Exp1 on indefinite direct objects.](image)

The same holds for the second experiment (Exp2) on definite noun phrases. Figure 2 shows that the referents of the pe-marked direct objects realized as definite noun phrases are more frequently re-mentioned than the unmarked ones. The predictions concerning this
metric are confirmed, as the *pe*-marked referents were picked up more often in the subsequent discourse than the referents of the unmarked direct objects.

![Image]

Fig. 2. Referential persistence of object referents in Exp2 on definite direct objects.

In sum, participants preferred a continuation story that evolved around the referent of the subject, thus taking it up more often, unless the direct object referent was *pe*-marked. In such a case, the referent of the *pe*-marked referent becomes a better competitor for the subject referent in terms of referential persistence.

### 4.2.2. Topic shift potential

The second textual characteristic investigated was the topic shift potential of direct object referents. Recall that each mention of a direct object in grammatical subject position was counted as an instance of topic shift. The counts for the topic shift potential are cumulative. Figure 3 and 4 indicate the percentage of direct object referents mentioned in topic position (the y-axis) in each continuation sentence (the x-axis). The findings condensed in Figure 3 and 4 reveal several patterns. First, the referent of the *pe*-marked direct object displays a stronger preference to become a subject in the continuation sentences (S1-S5) than the referent of the non-*pe*-marked direct object referent. Second, while almost all participants mentioned the referent of the *pe*-marked direct object sooner or later as a subject in the continuation text, the unmarked direct object became a subject in less than 25% of cases. Third, Figure 5 shows that the referent of the unmarked direct object was never picked up in subject position in the first two continuation sentences (S1 and S2) provided by the participants. On the contrary, the referent of the *pe*-marked direct object was picked up in the first two continuation sentences, even though the rate was not high.
The findings concerning the topic shift potential of direct objects confirmed the initial predictions, as the referents of the *pe*-marked direct objects displayed a higher expectancy to be mentioned again as topics in a main clause (i.e. in subject position) in comparison to the unmarked ones. The observations hold for both Exp 1 with indefinite noun phrases, and Exp2 with definite noun phrases.
4.3. Discussion

The findings with respect to the Discourse Structuring Potential of direct objects realized as definite noun phrases parallel those reported in Chiriacescu, von Heusinger (2010) about the discourse behaviour of indefinite noun phrases in direct object position. Up to the last continuation sentence (S5), the pe-marked referents (i) exceeded their unmarked counterparts in referential persistence (76% vs. 24%) and (ii) became the topic of the discourse more often than the non-pe-marked referents (in 80% vs. 15% of the cases).

Referential persistence and topic shift underline the privileged status of the pe-marked referents (whether expressed by an indefinite NP or by a definite unmodified NP) and thus confirmed Predictions 1 and 2 (cf. Chiriacescu 2011 on similar effects of indefinite-this in English and indefinite-som in German; Deichsel, von Heusinger 2011 and Deichsel 2013 for indefinite dies (‘this’) in German). The likelihood of a referent to be mentioned in the subsequent discourse is not a reflex of the high activation level of that referent, but rather represents a mechanism employed by the speaker to link the hearer’s attention to an entity which will be further elaborated upon. Psycholinguistic research (Levy 2008) has convincingly shown that statistical regularities are observed at different levels of linguistic output. It seems that hearers identify frequency patterns in order to predict what is likely to occur in the following context. The referential persistence of the pe-marked indefinite and definite descriptions analysed here shows that language users make use of such regularities at the discourse level as well.

5. CONCLUSION – THE FUNCTION OF PE IN ROMANIAN

In light of the findings of the experiment presented in section 4, we argue that pe signals the Discourse Structuring Potential of the referent it precedes. More concretely, it was shown that pe-marking has the same or a very similar function for definites and indefinites in neutral (transparent) contexts, as illustrated in (3) and (9). Such referents were shown to be more recurrent in the following discourse and to be more prone to shift the topic of the discourse. These observations add an additional dimension to the analysis of DOM: besides referential (specificity), lexical (animacy), information structural and backward-looking discourse properties, DOM indicates forward-looking discourse properties (i.e. Discourse Structuring Potential) as well. These properties are not pragmatic, but “built in” or semantic, similar to specific adjectives like a certain, or discourse particles like then, therefore, etc.

REFERENCES


Ionin, T., 2006, “This is definitely specific: specificity and definiteness in article systems”, *Natural Language Semantics*, 142, 175–234.


