

From topic to object: Grammaticalization of differential object marking in Romanian

EDGAR ONEA

University of Graz

edgar.onea-gaspar@uni-graz.at

and

ALEXANDRU MARDALE

Institut National des Langues et Civilisations Orientales

& Laboratoire SeDyL

UMR8202 CNRS

alexandru.mardale@inalco.fr

Abstract

In this paper, we provide new evidence for the hypothesis that in some languages, differential object marking (DOM) may have evolved from marking of the information structural category of topicality. While Iemmolo (2010) suggested that topic left-dislocation might have facilitated the evolution of DOM, and Dalrymple and Nikolaeva (2011) argued for systematic marking of secondary topics as the source of DOM, we suggest a third possible grammaticalization source within the topical domain: the prepositional marking of *embedded topics*. In particular, we show that the Romanian DO (Direct Object) marker *pe* had a function as marker of embedded topics in Old Romanian. Moreover, we show how DOM in Old Romanian may additionally have been facilitated by the re-analysis of embedded topics as THEMES in certain constructions.

Keywords: differential object marking, (embedded) topics, grammaticalization, (Old) Romanian

We gratefully acknowledge that work on this article benefitted from the financial support of the project *Unité et diversité dans le marquage différentiel de l'objet*, carried out at the Laboratoire *Structures et Dynamiques des Langues* (UMR8202) as part of the federal research on *Typologie et Universaux Linguistiques* (FR 2559) at the Centre National de la Recherche Scientifique, France. Moreover, we wish to thank our anonymous reviewers, who greatly contributed to improving the quality of the article. We highly appreciate their patience and invaluable comments. Finally, we wish to thank Igor Yanovich for his help as the editor of our article.

Résumé

Dans cet article, nous offrons de nouvelles preuves pour soutenir l'hypothèse que, dans certaines langues, le marquage différentiel des objets (DOM) a pu évoluer à partir du marquage de la topicalité, une catégorie de structure informationnelle. Tandis que Iemmolo (2010) a suggéré que la dislocation des topiques vers la gauche a pu faciliter l'évolution de DOM, et que Dalrymple et Naikolaeva (2011) ont proposé des arguments voulant que le marquage systématique des topiques secondaires soit à l'origine du DOM, nous suggérons ici une troisième source possible de cette grammaticalisation à l'intérieur du domaine topical: le marquage prépositionnel des topiques des *propositions enchâssées*. En particulier, nous démontrons que le marquer roumain d'objet direct *pe* fonctionnait, dans l'ancien roumain, comme marqueur des topiques des propositions enchâssées. De plus, nous démontrons comment le DOM, dans l'ancien roumain, a également pu être facilité par la réanalyse, dans certaines constructions, des topiques des propositions enchâssées comme thème.

Mots-clés: marquage différentiel de l'objet, topic, propositions enchâssées, grammaticalisation, (ancien) roumain

1. INTRODUCTION

Languages of the world tend not to mark the direct object (DO) at all with overt case markers and if they do, they usually do not mark them across the board (e.g., Sinnemäki 2014). This phenomenon, where in some languages DOs receive overt case-marking only under certain conditions, is called differential object marking (henceforth DOM, a term coined by Bossong 1985, 1991). The conditions under which DOs end up case marked are chiefly determined by the semantic properties of the DO. These are usually aligned as implicational scales in the sense of Comrie (1979), Aissen (2003) and Croft (2003). The two main scales are the animacy scale in (1) and the definiteness scale in (2).¹

(1) Animacy scale: [+human] > [+animate] > [-animate]

(2) Referentiality scale: pronoun > proper name > definite DP > indefinite DP

Languages usually single out segments of these scales for which marking of DOs is either necessary, optional or excluded. These segments are generally connected and do not exhibit gaps. Moreover, both scales are often used simultaneously, as discussed in Aissen (2003). For example, in modern Romanian, with some exceptions, only [+human] DOs can be marked (with the differential object marker *pe*), and while the marking is optional with indefinite DPs, it is obligatory with pronouns and proper names (see Farkas 1978, Dobrovie-Sorin 1994, Mardale 2009 and others).

The diachronic change associated with DOM usually proceeds as a gradual spreading from the left to the right on these scales. For example, von Heusinger

¹Languages may use much more fine-grained versions of these scales such as distinctions between specific and non-specific indefinites, and various types of pronouns, on the referentiality scale, or special treatments for godly entities or totem animals on the animacy scale. Moreover, additional syntactic factors such as incorporation or aspectual alternations appear to play a role as well. See Kalin (2018) for a recent syntactic approach to the latter.

and Kaiser (2005) and von Heusinger (2008) show for Spanish, as von Heusinger and Onea (2008), Avram and Zafiu (2017), Hill and Mardale (2019) show for Romanian, how DOM spreads, over several centuries, from mainly marking pronouns and proper names to marking most definite DPs, and finally to also marking some indefinites.

Two questions regarding the historic development of DOM immediately come to mind. The first one is why and how DOM would spread from left to right on the aforementioned scales. While, of course, it is hard to give definitive answers to why-questions in this domain, there are some entirely plausible assumptions in the literature. For example, von Heusinger (2008) and von Heusinger and Onea (2008) suggest that the spread proceeds by creating (more or less) ad hoc semantic distinctions within cells on the scales in which DOM is not yet systematically marked. At least descriptively, the expansion of DOM along the scales is sufficiently well-documented to be taken for granted in this article, and we will have nothing to add to this matter.

The second question, which is more important to our article, is why and how DOM would arise in a language in the first place. The literature offers a couple of potential answers. One important observation is that the animacy and referentiality scales are good indicators of prototypical subjects (Klein and de Swart 2010). Hence, for example, [+animate] [+definite] DPs tend to be more likely subjects than DOs, as compared to [-animate] [-definite] DPs (see, e.g., Øvrelid 2005 for Norwegian). Then, one may insist that the reason why DOs that are high on these scales tend to be case marked is that they are prototypical for subjects in terms of their animacy and referentiality, and thus easy to be taken for subjects in communication, resulting in misunderstanding (see, e.g., Comrie 1979 and Aissen 2003). Alternatively, one may suggest that DOs that come with such animacy and referential features, that place them high on the relevant scales, are indexed for their properties based on general notions such as *affectedness*, *transitivity* or *animacy*, as suggested by Hopper and Thompson (1980), Naess (2004), de Hoop and Malchukov (2008), and in subsequent literature. While different instantiations of either type of approach make subtly different predictions regarding the distribution of differential case marking in general, for DOM, they tend to make the same predictions (see de Swart 2003, Malchukov 2008): DOM should start at the left of the animacy and/or definiteness hierarchy.

Still, such theories usually neither explicitly attempt to explain, nor lend themselves readily to explaining, the exact mechanism by virtue of which DOM would emerge in a language. After all, disambiguating subjects and DOs is not necessary in all that many cases. Other cues such as word order or syntactic strategies, like passivization, are available to avoid ambiguity. One should not forget that most languages, as mentioned above, do without any DO marking. Moreover, there are also empirical issues for these theories: for example, Iemmolo (2010) shows that in some Romance languages DOM starts with precisely those pronouns that had retained overt accusative case markers from Latin; the need for disambiguation may not have been very great, for these pronouns. Having said this, even if disambiguation may have been necessary, one would still need to explain why some particular grammatical device (e.g., a preposition, as in the case of Romanian *pe*) would start being used as a DO marker. For indexing approaches, a similar issue arises: why would indexing of DOs in particular, for some semantic property, emerge, and why

would it begin in the way it did? Therefore, it seems reasonable to interpret the above-mentioned theories as explaining the existence or development of a functional pressure to develop DOM, but not as actual explanations for how exactly DOM developed in any particular language.

One natural way one could try to answer the latter question is to suggest that DO markers may originally have marked something else that was correlated with the left edge of the animacy and/or referentiality scale. The notion of topicality has most prominently been connected to the origin of DOM. It has been suggested that secondary topics (Dalrymple and Nikolaeva 2011) and left-dislocated topics (Iemmolo 2010) may have initially been marked in some languages with the soon-to-be differential object marker (see also Escandell-Vidal 2009 for effects of overt topic-marking, especially topic left-dislocation). Recently, Seržant (2017) suggested that another information structural construction, clefts, may also serve as a grammaticalization source for differential case marking. These information structural notions clearly correlate with some notion of prominence, and arguably coincide with highly animate/referential DOs sufficiently often to trigger the grammaticalization process. After all, on both the disambiguation and the indexing approaches, marking DOs high on the animacy and referentiality scales would have a clear functional advantage. Once the grammatical role of a DO marker has been established, the expansion process along the universal implicational scales mentioned above can easily kick in.

Iemmolo (2010) may be the best example of a theory reconstructing in detail how such a process of grammaticalization may have emerged in Romance languages. In particular, he argues that DOM emerges in the context of personal pronouns in dislocations. In these contexts, the Latin *ad* could be used as a topic marker, which was eventually re-analyzed as an object marker (after a stage of acting as a goal/dative marker) for precisely those direct objects that had the correct referential features to trigger the emergence of DOM. Arguably, such a story could be generalized to Persian *-râ* (Karimi 1990) and Hindi *-ko* (Montaut 2018).

In this article, we provide an analysis of the development of DOM in Romanian that resembles Iemmolo's analysis of DOM in Italian dialects in one important respect, while it differs from it in another – critical – respect. Specifically, we argue that DO marking in Romanian with the preposition *pe* may have developed from a topic-marker as well (Mardale and Onea 2017). However, we lack any evidence that Romanian *pe* may have been a general-purpose left-dislocation topic marker² or that *pe* may have been, at any point, a dative marker. Instead, we make two important observations about *pe* in Old Romanian.

Our first observation is that *pe* can be used in Old Romanian to mark what we call *embedded topics*, that is, constituents surfacing in a superordinate CP that mark the topic of an embedded clause. We will refer to these *embedded topics* as E-TOPICS and suggest that E-TOPIC is a semantic role expressing the topic of embedded propositional content, roughly equivalent with the semantic contribution of *about* in (3).

²Though fronting seems to have played a role in the early distribution of DOM in Old Romanian, see Hill and Mardale (2019).

(3) *Steve complained about Donald_i that he_i works too little.*

The fact that *pe* could mark E-TOPIC is in itself remarkable. To arrive at this, we offer a three-step argumentation. A) In (3), Donald is in no way to be understood as a topic of the matrix clause, but rather as the topic of the complaint made by Steve, supposedly an independent speech act uttered by Steve with its own, independent, discursive properties. Hence, we may assume that E-TOPICS do not share the discursive properties of topics within the discourse context of the matrix clause. However, they are marked for being topics of some other propositional content, and hence one may expect that an E-TOPIC may share the prototypical referential and argument structural properties of topics (Givón 1983,1984). B) Topics share prototypical referential properties with subjects (e.g., Comrie 1979). Incidentally, in (3) not only is *Donald* co-indexed with the subject of the embedded clause *he*, an AGENT, but it is hard to even imagine how one could complain about a non-agent. C) As already stated in Farkas (1978), DOM (in Romanian) applies to prototypical subjects, as is entirely expected (Aissen 2003). Therefore, if *pe* could be used to mark E-TOPIC, we have a direct link between *pe* and prototypical instances of DOM. Thus, the fact that a preposition can mark E-TOPIC may constitute in itself a reasonable level of analysis at which topicality leads to DOM.³ However, while we fundamentally uphold this analysis, it is also true that such an analysis remains on a fairly general/abstract level without specific links demonstrating how the transition from E-TOPIC marking to DO marking may have taken place. One wonders whether an even more fine-grained reconstruction of the grammaticalization process could be possible. Fortunately, our second observation may suggest a positive answer.

We found a remarkable case of what seems to be (or is at least stunningly similar to) an argument alternation with a limited set of verbs that may have led to the re-analysis of an E-TOPIC-marking construction as a transitive construction, thus enabling the evolution of a topic-marker into a DO marker. Schematically, the type of structure we have in mind is presented in (4a) and (4b) respectively, where V is the same word in both examples⁴:

- (4) a. AGENT:S V THEME⁵:DO/REFL-ACC E-TOPIC: [*pe* DP]
 b. AGENT:S V THEME:DO/[*pe* DP]

(4a) suggests that the verb V can be transitive. Otherwise, reflexivization with an accusative reflexive pronoun would not be possible. This schema also shows that in (4a) the *pe*-marked DP cannot be the DO; it must be something else. Since we know that *pe* is a preposition, the natural conclusion appears to be that (4a) involves a PP. Semantically, we will argue that this PP has the role of E-TOPIC. Representation (4b), however, lacks anything that would count as a DO except for the *pe*-marked DP. Hence, it may appear simplest to assume that the *pe*-marked DP is the DO of the transitive verb. The grammaticalization path that we suggest is, thus, a re-analysis of a PP

³We are indebted to an anonymous reviewer of this article for helping us clarify this point.

⁴We will discuss examples for such alternations in sections 3.2 and 5.2.

⁵In some cases, theme could be replaced by patient without changing anything in the overall architecture of our argument, as both are prototypical DO roles.

representing an E-TOPIC as a DO. While we will not have sufficient evidence to claim that such a re-analysis is the only factor triggering DOM in Romanian, we do suggest that it may have had a decisive role in its onset.

Accordingly, the main contribution of our article is to suggest a new possible path of grammaticalization for DOM, via argument-structural re-analysis, that has not yet been identified or explored in the literature. At the same time, our article also comes with some deeper theoretical implications with regard to the relation between information-structure and argument structure.

Casually speaking of an argument~E-TOPIC alternation, as we have seemingly done above, may strike the reader as a terrible idea. The traditional view is that information structure and argument structure are really different aspects of grammar that are – usually – dealt with using different resources (e.g., adpositions and case for argument structure vs. intonation and word order for information structure) even though categories of information structure do get morphosyntactic marking in many languages. This strict theoretical distinction would not only hold for radical theories keeping information structure outside of core grammar (e.g., Fanselow 2006), but also for cartographic orthodoxy (see Rizzi 1997 and subsequent literature), and for LFG accounts of information structure (see Dalrymple and Nikolaeva 2011). Indeed, we agree with this distinction in general. However, with verbs that have content/propositional arguments, we believe the situation to be different: there, topicality within the CONTENT-argument may be reflected by the semantic role E-TOPIC at the level of the matrix verb. Indeed, we introduce the notion of embedded topic (E-TOPIC) as a semantic argument of some verbs and their nominal counterparts, suggesting that, in cases of syntactic embedding of propositional content, information structural categories can become semantic roles. However, when they do, they are no longer information structural categories within the matrix clause. In other words, the categories *information structure* and *argument structure* are not really blurred.

Argument structure is at the core of what is being said, and information structure is a natural property of the act of saying, for example in terms of relating the said proposition to the context, information status, etc. Information structural categories such as topic are naturally construed as the way in which the speaker presents information, and should be represented as factors in the act of uttering, rather than as factors of the proposition itself. The linguistic form of utterances (syntax) reflects information structure at a higher level which is usually immaterial to semantic role assignment. However, we suggest that this distinction does not hold entirely when it comes to embedded content, as verbs that take propositional content as arguments may be sensitive to information structural characteristics of that content beyond the well-known cases of focus-sensitivity (see Beaver and Clark 2008). E-TOPIC is one specific way in which verbs semantically encode their sensitivity to the information structure of their CONTENT-argument. In particular, if a sentence *p* has a topic *t*, and *p* gets embedded under a matrix verb *v*, it is possible that the matrix verb *v* will show this fact by assigning *t* the semantic role of E-TOPIC.

When it comes to the presentational roadmap of this article, we face a difficulty. On the one hand, we want to focus on the analysis of the alternation schematically represented above, which we think is the best available cue to how and why DOM

may have emerged in Romanian. On the other hand, in order to be able to provide such an analysis, we will need to introduce and motivate a new semantic notion, E-TOPIC, which is neither specific to Romanian nor to the problem at hand. We deal with this problem in the following way: first, in section 2, we provide a brief background on DOM in present-day Romanian and Old Romanian and – alongside that – we present the type of corpus data we worked with. In section 3, we present the data that are relevant for our proposal regarding the grammaticalization path of *pe* to a DO marker in Old Romanian, along with an outline of how grammaticalization could have worked in general. In section 4, we introduce the notion of E-TOPIC as a necessary tool to better understand the data revealed in section 3. In section 5, we show that this notion is well-suited to capture the Romanian data. Moreover, we show in detail what kind of re-analysis may have happened in Old Romanian that may have led to the reanalysis of *pe* as a DOM marker. In section 6, we discuss some strengths and weaknesses of our proposal, also clarifying and further strengthening our argument. We conclude with a section in which we come back to the broader picture of how topicality relates to DOM.

2. BACKGROUND

In this section, we briefly describe the main synchronic and diachronic distributional facts concerning DOM in Romanian. We thereby highlight a few generalizations important to the following discussion, and present the data used in this article. A secondary goal of this section is to confirm, by virtue of a small corpus annotation, that our data exhibit distributional patterns entirely compatible with the general diachronic picture suggested in the existing literature.

2.1 Synchronic DOM in Romanian

In present-day Romanian, DOs may appear either without any overt case marking, as in (5a), or marked with the preposition *pe* (5b–c). In addition, in many cases, DOs are also doubled by a weak clitic pronoun, as in (5c). The preposition *pe* also has additional usages as a locative preposition, as exemplified in (6), and further discussed in section 2.3.

- (5) a. *Am văzut (*pe) un film.*
 have.1SG seen pe a movie
 intended: ‘I have seen a movie.’
- b. *Am văzut pe un tânăr.*
 have.1SG seen pe a young man
 ‘I have seen a young man.’
- c. *L=am văzut pe un tânăr.*
 CL.3SG=have.1SG seen pe a young man
 ‘I have seen a young man.’

- (6) *Am dormit pe un covor.*
 have.1SG slept pe a carpet
 'I have slept on a carpet.'

The exact generalizations covering the distribution of DOM are complicated by a range of exceptions, which appear to be orthogonal to our article. However, the main tendencies seem very clear: only [+human] and at least [+specific] DOs can be marked with *pe*. Pronouns and proper names are always marked with *pe*. Definite DPs that are not weak in the sense of Carlson et al. (2006) are always marked with *pe* unless in possessor raising constructions (see Onea and Hole 2017). Indefinites are marked with *pe* if they are specific (see Farkas 1978 and subsequent research). We will not enter into a more detailed synchronic description here.

There is, however, one significant synchronic observation that we will later need: the preposition *pe* preserves syntactic features typical of prepositions in Romanian even when clearly acting as a DO marker (Mardale 2009, Hill and Mardale 2019). We illustrate this in (7). In Romanian, definite DPs are generally marked by the enclitic definite article as in (7a) irrespective of the syntactic role of the DP (Dobrovie-Sorin and Giurgea 2006). However, if the definite DP contains only a noun without any modifiers and the DP is headed by a preposition (except the prepositions *cu* 'with' and *de-a* 'as'), the definite article cannot be overt, as shown in (7b). This phenomenon is known as *article drop* (Dobrovie-Sorin 2007, Mardale 2008, Onea and Hole 2017). Crucially, this phenomenon, otherwise limited to prepositions, also applies to clear cases of DOs marked with *pe*, hence showing that at least some prepositional features of the differential object marker *pe* are preserved.

- (7) a. *Am dormit lângă copac-ul cel mai verde.*
 have.1SG slept near tree-DEF that most green
 'I slept near the greenest tree.'
- b. *Am dormit lângă copac(*-ul).*
 have.1SG slept near tree(-DEF)
 'I slept near the tree.'
- c. *L=am văzut pe senator-ul cel mai corupt.*
 CL.3SG=have.1SG seen pe senator-DEF the most corrupt
 'I saw the most corrupt senator.'
- d. *(L-)am văzut pe senator(*-ul).*
 CL.3SG=have.1SG seen pe senator(-DEF)
 'I saw the senator'

From this observation one could conclude that *pe* as a DO marker did not undergo a profound diachronic change with regard to its original locative prepositional variant. Instead, one could conjecture that *pe* may have been able to maintain its prepositional status because the process of becoming an object marker was neither abrupt nor did it involve any radical step of (syntactic) reanalysis. This, in turn, suggests that a search for the diachronic source of DOM in Romanian should naturally begin with the very semantics of *pe* as a locative preposition.

2.2 Diachronic Data

Given the fact that we do not have a large amount of written material from Romanian before the 16th century, the history of DOM in Romanian cannot be fully reconstructed on solid empirical grounds. The diachronic spreading of DOM since the 16th century has been, however, fairly well documented (von Heusinger and Onea 2008; Stark 2011; Hill 2013; Mardale 2015; Avram and Zafiu 2017; Hill and Mardale 2017, 2019, and others). In particular, it seems that in the 16th century, DOM is attested mainly with proper names and pronouns, and is beginning to spread to definite DPs. Later, especially in the 19th and 20th century, DOM spreads towards (specific) indefinite DPs. This diachronic process appears to be very similar to the one described by von Heusinger and Kaiser (2005) and von Heusinger (2008) for Spanish.

The data we consider for this article are the oldest original texts in Romanian, the *Documente și însemnări românești* [Romanian Documents and Notes], 1521–1625, (edited by Chivu et al. 1979) To provide a brief overview of the diachronic situation reflected by this particular dataset⁶, we have manually annotated the first 400 uncontroversial DOs expressed as DPs (including quantificational and negative phrases as well as personal and relative pronouns, but not including clitic pronouns) for referential categories, *pe*-marking and animacy.⁷ Our annotation did not include cases in which the argument structure of the verb was not clear, but this in no case led to ignoring an argument marked by *pe*. Hence, the proportion of unmarked DOs may be higher than reflected by the data we report further below. Notice that more detailed reports on the distribution of DOM and clitic pronouns in Old Romanian (partly based on the same corpus) are given in Mardale (2015) and Hill and Mardale (2017, 2019). Our report here is limited to our manually annotated data.

We found only one instance of *pe*-marking for non-human DOs: an instance of a country-name, provided in (8). It is not entirely obvious whether this example can be treated as an error (since in all other cases country names as DOs were not used with *pe*), or whether some variation existed as to whether countries could in certain cases be treated as [+human] (e.g., in a metonymic sense).

⁶The texts are mostly official documents such as contracts and complaints, as well as additional judiciary and diplomatic documents. Most texts are short, usually containing fewer than 300 words; many contain repetitive structures and enumerations of goods, lands or subjects, for example, lists of witnesses or items sold. Some of the longest diplomatic documents contain 900–1200 words.

⁷An anonymous reviewer raised the question of why we limited the analysis to the first 400 DOs. The answer is twofold. We did not annotate more DOs because the data show a clear picture that is widely in line with the findings in the literature; there is no reason to assume that annotating more DOs would change anything substantial in the resulting overview. More importantly, since the texts are ordered chronologically and we were mainly interested in the oldest attestations, annotating the first 400, instead of a random collection of 400 DOs, gave us better insight into the state of the language we were interested in.

	pers. pronouns	proper names	def. DP	indef. DP	others	Total
[+pe] [+cl]	5	1	2	0	0	8
[+pe] [-cl]	4	25	19	0	1	49
[-pe] [+cl]	0	0	1	0	0	1
[-pe] [-cl]	0	1	34	2	29	67
Total	9	28	56	2	30	125

Table 1: *Pe*-marking for human DOs in Old Romanian (absolute numbers)

- (8) *de-au scosu pre⁸ țeara Moldovei den mâna tătarâlor*
 that=have.3PL liberated pe country Moldova.GEN from hand.DEF
 Tatars.GEN
 ‘...that they have liberated Moldova from the Tatars’

In Table 1, we report the results of our annotation for human DOs. The data widely confirm older observations by von Heusinger and Onea (2008) from a corpus of Bible translations: in general, *pe*-marking was fairly obligatory with personal pronouns and obligatory with proper names while it was optional with definite descriptions and very uncommon or even ungrammatical with indefinites and other types of expressions.

We provide an example of *pe*-marked proper names and definite descriptions in (9) and of *pe*-marked personal pronouns in (10). The only example of a non *pe*-marked proper name in our data is the one given in (11). While we have annotated the expression *Dumnezeu* (‘God’) as a proper name, because it was generally both marked with *pe* and written with a capital initial letter, it is by no means clear that it was always considered a proper name in common usage.⁹ Moreover, the example immediately continues with a Slavonic passage, which may also suggest that there might be grammatical contamination due to code switching. Hence, we cannot consider (11) a sufficient reason to dismiss the generalization that *pe*-marking was obligatory with proper names in Old Romanian.

- (9) *să pomenească și pre Radu (...) și pre părinții lu[i] (...) în sfânta liturghie*
 SUBJ mention also pe Radu and pe parents.DEF his in holy.DEF mass
 ‘(they) shall also mention Radu and his parents during the holy mass’
- (10) *și va să ne piarăză și pre noi*
 and will SUBJ CL.1PL destroy also pe us
 ‘and they want to destroy us’

⁸In the 16th century, *pe* appears in a number of variants including *pre* and the regional versions *pi*, *pă*, *pri*, *piră*.

⁹Incidentally, in (13) we have one more example, outside of the annotated part of our corpus, in which a DO realized as a proper name is not marked with *pe*. Again, in this example, the proper name is a godly entity: *Hristos* (‘Christ’).

- (11) *ei încă roagă Dumnezeu*
 they still ask God
 ‘...they still ask God...’

Table 1 also suggests that clitic doubling with *pe*-marked DOs was not very common in Old Romanian, at least before 1600. Clitic doubling was only common when the DO was a stressed personal pronoun, and quite rare with proper names and definite descriptions. In fact, all instances in which a co-referential clitic pronoun occurred with a proper name or with a definite description were cases of clitic-left-dislocation, as in (12) (see also Escandell-Vidal 2009 and Iemmolo 2010 for similar observations about early Catalan and Hill and Mardale 2017, 2019 about early Romanian).

- (12) *iară pre acei boiari i=au datu în chezășie*
 and pe those nobles CL.3PL=AUX given in hostage-ness
 ‘and they gave those nobles as hostages’

3. SETTING THE STAGE: *PE* IN OLD ROMANIAN

If one is to investigate the way in which a preposition, in our case *pe*, ends up as a DO marker, a natural place to start is to consider the various usages that preposition had in the earliest available attestations. In section 3.1, we do just that. Thereby, we point out that *pe* had some abstract usages it widely lacks in present-day Romanian – usages lost because of the evolution of *pe* into a DO marker. In section 3.2 we focus further on these usages and present some pairs of examples in which the very same verb (at least in its overt form) arguably takes DOs marked with *pe* in one case and non-DOs marked with *pe* in the other. We will suggest that this is a natural point of departure for understanding the mechanisms of grammaticalization, even though we still lack a clear semantic understanding of what exactly is happening in these data at this stage of the analysis.

3.1 Other usages of the preposition *pe* in Old Romanian

Alongside its usage as a DO marker, *pe*, as shown above, most commonly appears as a locative preposition with a considerable range of locative functions. These are exemplified in (13).

- (13) a. *iară până au pus Hristos pe cruce 5533* (ON)
 again until have put Christ pe cross 5533
 ‘and again, until they have put Christ on the cross, in 5533’
- b. *pre ceastă vreme și iute și rea* (ON/AT)
 pe this time and bitter and bad
 ‘in these bitter and bad times’
- c. *și purtat pre munți* (ACROSS)
 and carried pe mountains
 ‘and she carried (me) across the mountains’
- d. *Și se=au dus în sus pre Dunăre* (ALONG)
 and REFL.3PL=have gone in up pe Danube
 ‘and they went up along the Danube’

It is not the goal of this article to provide a sensible spatial-semantic analysis of the meaning of *pe* in Old Romanian by providing a spatial model, for example in the sense of Jackendoff (1990) or Kracht (2002). Part of the difficulty would be that there may have been significant speaker variation within the data, and there also seems to be some amount of polysemy. For example, (13a) appears to introduce a goal/place function, whereas the examples in (13c) and (13d) appear to introduce the path of some movement. Such patterns would pose a challenge even in a synchronic setting, where we could rely on speaker intuitions to delimit the range of possible readings.

Moreover, there is a range of usages in different conceptual domains that appear to be closely related to the basic locative reading, some of which are exemplified in (14).

- (14) a. *ȕi=am tremes în tot chip pre Iurgachi.* (THROUGH)
 CL.2PL.DAT=have sent in every way *pe* Iurgachi
 ‘we sent you (the result) by all means through Iurgachi’
- b. *noi am dat pre ei tot galbeni* ((IN) EXCHANGE FOR)
 we have given *pe* them also gold
 ‘we also spent gold for them’
- c. *am judecat pre drept și pre lege* (BY/ACCORDING TO)
 have judged *pe* justice and *pe* low
 ‘(we) judged according to the justice and to the law’
- d. *și pre mai mare credințe ne=am pus* (WITH THE AIM OF)
 and *pe* more big hope CL.1PL.ACC=have put
și pecețile
 also seals.DEF
 ‘and, with the aim of higher hope, (we) apposed our seals’
- e. *Deci am un frate de tată pre nume Du <MI>tru* (BY (THE NAME OF))
 so have a brother of father *pe* name Dumitru
 ‘So I have a brother on my father’s side whose name is Dumitru’

Some of these readings can be reconstructed either within the framework of conceptual metaphors (Lakoff and Johnson 1980) or within some more elaborate localist theories of semantic roles (e.g., Ostler 1979). For example, one could at least imagine that (14b) involves a goal specification within some transactional domain, or that (14c) would involve a conceptualization of the legal basis of judiciary decisions as *the grounds* of judgments, that is, the place on which these judgments stand. That these readings are closely connected to the core locative meaning of *pe* (‘on’) is evidenced also by the fact that we find these readings in distantly related languages for the corresponding preposition. For example, while English *on* cannot be used in (14b) with the verb *pay*, the expression *spend money on something* implies that *on* can be used in English to express what we glossed as ‘in exchange for’. Similarly, the German preposition *auf* (‘on’) can express the conventional meaning captured by (14d) in the construction *auf dass* (‘with the aim of’).

Upon closer inspection, there is a wide variety of usages of *pe* in Old Romanian that appear to be connected to the reading in (14b), but which extends beyond the domain of money and spending, covering a reading ranging between *for*, *concerning*, and *about*, that English *on* also shares in certain contexts. We provide a few examples in (15) and English examples in (16):

- (15) a. *Și cheltuiala măruntă pe treaba lu voievod*
 and expense.DEF small pe business.DEF his ruler
 ‘And small expenses for / concerning the ruler’s activity’
- b. *am scris aceasta carte lu Panaet de Hotărani,*
 have.1SG written this book/missive he.DAT Panaet of Hotarani
pe ocină ce i=am vândut iară în Hotărani.
 pe land that CL.3SG.DAT=have.1SG sold again in Hotarani
 ‘I wrote this letter to Panaet from Hotarani, about/concerning the land I have sold to him in Hotarani’
- c. *că am avut pîră amîndoi pe nește rumîni din Răsnicel*
 that have.1SG had complain both pe some servants from Rasnicel
 ‘because we both had a complaint about/concerning some servants from Răsnicel’
- d. *de ne=au tocmit pe acești rumîni ce sînt mai susscrîși*
 that CL.1PL.acc=have negotiated pe these servants that are more above-written
 ‘so that they have negotiated with us concerning the above-mentioned servants’

- (16) a. *It was the kind of moment I’d dreamed of having ever since I read my first book on bats in high school.*

Dan Riskin: *Mother Nature Is Trying to Kill You*

- b. *More specifically, a detained person can make an oral or written complaint on any matter concerning his or her detention to the Chief Custody Officer at any time and he or she may be represented by their counsel.*

Róisín Mulgrew, Denis Abels eds. *Research Handbook on the International Penal System*

These readings are interesting for two reasons. Firstly, it is not obvious how exactly such readings could be connected to the spatial meaning of *pe*. But at the same time, such a connection probably exists at some cognitive level – be it a conceptual or a metaphorical level – because different languages tend to exhibit the same or similar patterns. Secondly, these readings are no longer readily expressed with *pe* in modern Romanian. The usual defaults include *despre* (‘about’) (17a), used entirely differently in Old Romanian and *asupra* (‘concerning’) (17b). Crucially, using *pe* in such constructions is not always entirely out in modern Romanian, but examples get better and less archaic in cases in which the argument of *pe* is [-animate] and thus could not be a *pe*-marked DO, as, for example, in (17c) and (17d). Hence, it looks like *pe* has at least partially lost some of its usages transitioning from Old Romanian to present-day Romanian.

- (17) a. *Troțki, colegul său de revoluție, a scris o carte despre(??pe) Baudelaire.*
 Trotsky, colleague his of revolution has written a book despre/pe
 Baudelaire
 ‘Trotsky, his revolutionary colleague, has written a book about Baudelaire.’
 George Friedman: Puncte de presiune. Despre viitoarea criză din Europa
- b. *A făcut niște observații asupra sintaxei (/?pe sintaxă).*
 has made some remarks asupra syntax.DAT pe syntax
 ‘S/he made some remarks on syntax.’
- c. *Mircea e consilier pe probleme de securitate internă.*
 Mircea is advisor pe problems of security internal
 ‘Mircea is an advisor on internal security problems.’
- d. *Bunicii i=au făcut acte/ documente pe terenul de la mare.*
 Grandparents.DEF CL.3SG.DAT=have made papers/documents pe land.DEF from at sea
 ‘His grandparents made him documents regarding the seaside land.’

On the one hand, these readings may not be the best places to look for a source of grammaticalization of *pe* as a DO marker because they do not seem to be obviously connected to expressions which have high values for animacy or definiteness. But on the other hand, the fact that some of these readings have been lost in modern Romanian seems to be a good indication that they may have been at least impacted by the evolution of DOM.

3.2 Argument alternation and grammaticalization

With a limited set of verbs, the abstract readings suggested above come in interesting pairs. Such pairs¹⁰ are exemplified in (18) and (19).

- (18) a. *s=au jeluit sluga noastră Ștefan pre Samoil*
 CL.REFL.3SG=has complained servant.DEF our Stephen pe Samuel
 ‘And our servant Stephen complained about Samuel’
- b. *jeluiaste sluga noastră Ștefan Moimăscul pre niște cuconi mici*
 mourns servant.DEF our Stephen Moimăsc.DEF pe some children little.PL
 ‘Our servant Stephen Moimăsc mourns for some (little) children’
- (19) a. *se=au tocmit pre megiiăș[i]*
 CL.REFL.3PL=have haggle pe neighbours
 ‘They haggled over their neighbours’
- b. *amu tocmit pre Mușat postelnic și pre Negoe pîntru*
 I.have hired pe Mușat chamberlain and pe Negoe for
4 rumîni vii și pîntru 6 delniți, firî rumîn
 4 serfs alive and for 6 plots-of-land without serf
 ‘I hired Musat the chamberlain and Negoe in exchange for 4 living serfs and for 6 plots of land, without serfs’

¹⁰Given the scarcity of data, we were unable to find minimal pairs.

Two important facts are remarkable about these data. Firstly, they can be schematically represented as in (20), revealing that the *pe*-marked arguments must have a different argument structure in each of the constructions. The representation in (20a) suggests that the verb can be transitive, because of the presence of the reflexive pronoun, but the *pe*-marked DP cannot be the DO, so it must be a PP. The one in (20b), however, lacks anything that would count as a DO except for the *pe*-marked DP. Hence, under the assumption that the verb is transitive after all, the *pe*-marked DP would have to be analyzed as the DO, and the PP as missing (either no longer part of the argument structure, or not overtly realized.)

- (20) a. S V DO:REFL-ACC [*pe* DP]_{PP}
 b. S V [*pe* DP]_{DO}

The second important fact is that the distribution of this phenomenon is generally limited to verbs that involve propositional or even utterance content, including verbs of bargaining and economic transactions, which involve performative components. Interestingly, each of the two syntactic constructions above lead to different English translations of the respective verbs, while remaining within the same conceptual domain: *complain* vs. *mourn* and *haggle* vs. *hire*.

Hence, one may try to understand this phenomenon as an argument alternation, in the sense of Rappaport and Levin (2008) and others. Levin (2015) provides a number of properties that should hold for argument alternations. For example, there should be a significant meaning overlap between the two variants. Most, if not all, of the same semantic arguments ought to be expressed in both variants. The same pattern of alternation can be more or less exactly reproduced for other, similar, verbs. Prima facies, our examples do not seem to fit directly into this pattern. For example, the fact that the English translations differ should cast doubt on the semantic similarity. Moreover, it is not exactly the same arguments that are expressed in each variant. And finally, we have little evidence that the patterns could in fact be reproduced with other, similar, verbs.

Given the scope of this article, it will not be crucial for us to make the claim that these are bona fide argument alternations; but we suggest that they are at least very close to argument alternations. Firstly, we do not know (given the data situation) whether by fixing all the mismatches between the examples in order to create minimal pairs, we would or would not obtain very similar meanings. We will argue in section 5.2. that this is at least plausible. For now, it should at least be observed that hiring appears to be the result of a haggling process, and mourning and complaining differ mainly in what could be deemed the cause or occasion (some bad event for complaining; death for mourning). Secondly, the fact that the verbs are translated differently into English is unsurprising, given that argument alternations do have semantic impact and languages may differ in containing verbs that express both feature bundles associated with an alternation (see Levin 2008). For example, the contrast between *buy* and *sell* or *borrow* and *lend* has been analyzed as an argument alternation in Puyuma (see Teng 2008, Kuo 2015). Put differently: if in English the same argument alternation does not exist, it would not be surprising that we get different lexical translations. Moreover, the type of verbs in question do

- (24) a. S_i V Arg1: ACC_i Arg2:PP
 b. S V Arg1:zero/ACC Arg2:PP/zero
 c. S V Arg2:ACC

One could object to our hypothetical explanation, however, pointing out that it is not obvious why Arg2:PP would ever be re-analyzed as Arg1:zero from a semantic perspective. Turning to our example, this objection amounts to the following question: How could the role played by the agent in complaining possibly be attributed to the role of the dead person? But then again, if we could convincingly argue that, in fact, in (24b) Arg2 is co-indexed with Arg1, the objection would no longer constitute a problem, as the re-analysis would be – in this case – a very natural step. After all, being the same individual is the best possible reason for reanalysis.

This is precisely what we will eventually suggest, even though the analysis will turn out to be a bit more complex. In order to be able to suggest such an analysis, we need to achieve a deeper understanding of the role played by the PP in these constructions. This is the goal of the following section.

4. A DETOUR ABOUT ABOUT

We have seen in the previous section that *pe* has a variety of readings in Old Romanian. The one of interest for explaining the argument alternation observed in 3.2 seems to involve a meaning closely related to that of the preposition *about*. Thus, in order to be able to say more about what *pe* does in the relevant constructions, it seems reasonable to take a closer look at the meaning of *about*: we will first consider the main existing approaches, and then develop an alternative analysis that improves on these approaches by introducing the semantic role E-TOPIC. Finally, we will add some formal details to our analysis.

4.1 Existing accounts

Consider the example in (3) (structurally similar to examples from Moulton 2009), repeated for convenience as (25), as a starting point of our discussion. The question we are interested in is what the semantic contribution of *about* may be, in this example.

- (25) *Steve complained about Donald_i that he_i works too little.*

According to Moulton (2009) and related research, one may assume that the verb *to complain* in English is a so-called content-verb, that is, it may take a CONTENT-argument. A CONTENT is simply propositional content, such as an embedded question or proposition, the story in a book, the image in a picture, etc.¹¹ This is quite similar to the older, more or less standard, analysis of such verbs as taking propositional THEME-arguments, with the notable advantage that it leads to natural predictions concerning the distribution of *that*-argument-clauses: since not every verb (or nominal)

¹¹Obviously, we do not consider non-propositional content such as fluids in bottles.

has a CONTENT, not every verb may take a propositional argument. For (25), then, the proposition that Donald works too little modifies the CONTENT-argument of the verb *complain*, hence yielding the correct result: that the content of the complaint made by Steve is that Donald works too little. As for the semantic import of the preposition *about* in such sentences¹², Moulton (2009) suggests that it introduces a different argument he calls the *res* argument, reminiscent of Quine (1956) or Cresswell and von Stechow (1982). Moulton is not very explicit about what a *res* is, except that it gives a natural account of *de re* readings associated with *about* arguments.

Rawlins (2013) further elaborates on the meaning of *about* arguments, focusing mainly on the distribution of possible internal arguments of *about*. These are particularly interesting for a compositional semantic analysis because they include embedded questions as in (26b), and DPs as in (26a), which arguably have different semantic types.

- (26) a. *John asked about Joanna.*
 b. *John asked about whether Joanna was going.*

The strategy adopted by Rawlins is to take the question-embedding case as the standard, basic case. He suggests that when the internal argument of *about* is a DP, a question needs to be extracted from that DP. When it comes to the truth conditional import of *about*, the intuition in Rawlins (2013) is roughly this: a question A can be *about* a question B if at least some resolution of A helps in resolving B. For (26a), that would suggest that the question actually asked by John was such that it would help resolving some other question constructed from some salient property of Joanna. For (26b), in the same vein, the content of the speech act performed by John was helpful in determining whether Joanna was going or not. Given that Rawlins uses a version of Inquisitive Semantics (see Groenendijk and Roelofsen 2009, Ciardelli et al. 2013, Ciardelli et al. 2018) underlying his type-logic, he can extend this analysis also to examples like (25). The reason for this is that questions (i.e., sets of sets of worlds) and propositions in the classical sense (i.e., sets of worlds), have the same semantic type in Inquisitive Semantics, namely sets of sets of worlds, and, thus, share some properties. *That*-clauses and questions are both propositions in Inquisitive semantics.

At least at an intuitive level, it seems to us that the analysis of *about*-arguments as introducing a different semantic role than CONTENT appears to be correct. However, when it comes to the technicalities, we are skeptical about the analysis outlined in Rawlins (2013). Rawlins uses a symmetrical implementation of what it means for a question to be about another question using a notion of *orthogonality* from Lewis (1988).¹³ Essentially, for a question to be non-orthogonal to another question, it will suffice if one partition cell, in the sense of Groenendijk and Stokhof (1984),

¹²Older literature generally focused only on usages of *about*-arguments without an additional *that*-complement (see e.g., Pesetsky 1982, Boër 1978), and will not be further discussed here. See Rawlins (2013) for discussion.

¹³This does not mean that CONTENT arguments and *about*-arguments can be interchanged. The reason for this is that the CONTENT argument is explicitly marked as such, say by a *that*-

does not overlap. This predicts truth conditions that are too weak. For example, this would predict that Dan could report the situation in (27a) by uttering (27b). After all, the salient question about Peter is given in the context of (27b), and this question is clearly not orthogonal to the question of whether Alfred and Berta slept well – at least given reasonable assumptions about parties disturbing neighbors. Intuitively, John did not ask anything about Peter in the described situation, regardless of whether any question about Peter’s properties/actions – orthogonal or not to his question – may be relevant in context. One obvious reason for this is that John does not even know that Peter exists; hence, John cannot be involved in speech acts about Peter.

- (27) a. Situation: John asked whether Alfred and Berta slept well. John does not know anything about Peter. Alfred and Berta are the people living underneath Peter’s apartment.
- b. Dan: *I was wondering whether Peter had a party last night. I am not sure, but I recently found out that John asked about Peter.*

Similarly, it is not obvious why one would attempt to unify the polymorphic nature of *about* and its internal arguments to a basic, propositional case, as opposed to a DP-denotation. It is not obvious that examples like the ones provided in (28) require any kind of context and allow any non-trivial coercion of DP-arguments of *about* to questions. Yet these examples are prototypical for the usage of *about* in English.

- (28) a. *I read a book about Paris.*
- b. *So happy to meet you, Mr. Peterson. I heard about you on a radio talk show.*
- c. *A book about bats usually contains pictures of bats.*

4.2 About embedded topics

Given these difficulties, we will slightly modify Rawlins’ approach in this section. While we agree with both Moulton (2009) and Rawlins (2013) that *about* introduces a semantic argument of the verb, we will assume that it is neither a question nor a *res*-argument, even though, arguably, what we say could in fact be compatible with Moulton’s analysis.

Instead, we suggest that the way in which a complaint by Steve is about Donald in (25) is not very different from the way in which utterances like those in (29) are *about* Donald. In other words, for (25), we suggest that (on the most natural interpretation) *about* Donald suggests that Steve made a statement in which Donald was the *aboutness*-topic (Reinhart 1981) and whose content was that he works too little. In other words, if Steve made an utterance like any of those in (29), (25) is true.

- (29) a. Steve: *Donald, he really works too little.*
- b. Steve: *Donald works too little.*
- c. Steve: *As for Donald, he should work more.*

clause, or is entirely missing. As such, some level of asymmetry between the CONTENT argument and the *about*-arguments remains.

Before turning to additional details, there are several intuitive arguments supporting this analysis. For one thing, if one were to imagine a list of complaints, say in court archives, one could well imagine that complaints are ordered as file-cards. A complaint about Donald would then be a complaint whose file-card is entitled “Donald”. This is – of course – quite exactly the intuition captured in the theory of topicality put forth by Reinhart (1981). In other words, our approach essentially just captures the meta-language usage of *aboutness*-topic as a theoretical notion. Moreover, one can also imagine that if X is the topic of a complaint, the complaint, as a discourse segment in the sense of Asher and Lascarides (2003), would have precisely X as a discourse topic in the sense of Asher (2004) (see also Roberts 2011). Moreover, this also – at least intuitively – captures an old puzzle about topicality. Consider the examples in (30).

- (30) a. *In planning the restoration of Poplar Island, for example, participants extensively negotiated **about** what percentage of the island would be restored as wetlands.*
Beneficial Uses of Dredged Materials Case Study: Poplar Island, Chesapeake Bay,
www.epa.gov
- b. *However, Peter had not **negotiated about** the terms with the client for any further enhancements or features being added to the software.*
www.managementstudyguide.com/free-courses/Negotiation-Skills-Basics.pptx

In general, the *about*-argument of *negotiate* is typically an overt embedded question as in (30a), or a concealed question, as in (30b). This seems to fit naturally with another implementation of the notion of topic suggested by van Kuppevelt (1995) and subsequent research, according to which discourse topics can be some questions the discourse is about. This notion has never been seriously connected to *aboutness*-topic in the sense of Reinhart (1981), however (see Roberts 2011 for an overview). Thinking of *about*-arguments as topics of embedded content which may consist of one or more speech acts, however, seems to provide empirical evidence that both notions actually capture the same intuition and may have a common core.

After this intuitive captatio benevolentiae regarding our analysis, let us become a bit clearer on our proposal, before outlining, in the next section, its possible formalization. We suggest that there is a semantic role that we would like to call E-TOPIC (embedded topic), that is introduced in English (among other languages) by the preposition *about*. If A is the E-TOPIC of a verb V with B as a CONTENT, then B must be the content of some discourse D that verifies the respective utterance. A, then, is the discourse topic of D. If D happens to be exactly one sentence S, then A will naturally be the sentence topic of S, which is usually an *aboutness*-topic.

Since a discourse topic can be either a question (van Kuppevelt 1995) or an individual (Asher and Lascarides 2003, Asher 2004), it is not surprising that *about* can take either of these as internal arguments. Whether or not it is a sensible approach to reduce either of these two topical categories to the other is a question we do not wish to discuss in this article (pace Rawlins 2013). It is sufficient to notice that DPs can be topics, regardless of whether they have erotetic implications or not.

Finally, we wish to clarify that while we say that E-TOPIC is a semantic role, we are not in the business of the evaluation and enrichment of the inventory of semantic

roles in general (see, e.g., Dowty 1991 for discussion). Similarly, we do not make any strong claims about argument structure as a grammatical category nor its role in the architecture of grammar. Both of these notions are very strongly debated. While these debates are both important and interesting in their own right, they go far beyond the scope of this article. There are very restricted usages of semantic roles in the literature, as well as very permissive ones. Similarly, there are approaches that very closely connect semantic roles to the distinction between arguments and adjuncts (see, e.g., Koenig et al. 2003 for discussion). Here, we make a very limited claim: we suggest that some verbs are compatible with E-TOPIC and some verbs are not. To explain this, we will assume that E-TOPIC is compatible with the meaning of some verbs and not with the meaning of others. Moreover, in the particular framework we propose, this means that this compatibility is lexically coded.¹⁴

4.3 E-TOPIC as a semantic role

In this section, we provide a somewhat more explicit discussion of what we mean by E-TOPIC. Our setup will be similar to analyses by Kratzer (1998), Hacquard (2006, 2010) and others, but simplified for the purposes of this article.

Let us assume there is a set of role-functions such as SIZE, COLOR, NAME, THEME, AGENT, DEGREE, LOCATION, PATH, etc. These functions establish a relation between two entities within a semantic model. We assume that role-functions are partial, that is, not every entity is mapped by (say) LOCATION to some entity. We will use lexical semantics to handle the arising partiality as suggested by the examples (31); leaving aside the possibility of creating hierarchies of roles to simplify notations.

- (31) a. $\llbracket dog \rrbracket = \lambda x_{[x \in \text{Dom}(\text{size, location, } \dots)]}. x \in \text{dog}$
 b. $\llbracket \text{love} \rrbracket = \lambda e_{[e \in \text{Dom}(\text{agent, theme, time, } \dots)]}. e \in \text{love}$

For explicitness, we assume that all role-modifications are introduced by a role-head which can be either a syntactic operator, the lexical meaning of a preposition, or some other morpheme. We give examples of the first two possibilities in what follows.¹⁵

¹⁴Depending on particular assumptions one may have about the overall architecture of grammar with regards to grammatical roles, argument structure, and semantic roles, more or less radical consequences can be drawn from our analysis. We remain silent on such consequences in general. The only point we find essential and wish to draw attention to is this: in our analysis, E-topic is a verbal semantic category, while topic itself is an information structural category. Since the implication of X being an E-TOPIC of a verb V essentially boils down to X being a topic of some content in the information structural sense, this raises the question of whether a general analysis of topics could be possible, in which topics eventually turn out not to be categories of information structure, but rather verbal categories. We hope to pursue this in future research.

¹⁵To deal with quantification, we will need a slightly more involved theory of composition with partial functions. Consider, for example *Every house is an object*. The problem here is that houses may be modified by location but this is not required for objects in general. In other words, one needs to deal with the fact that some individuals are not in the domain of *house* but may well be in the domain of *object*. Off-the-shelf compositional systems that easily

- (32) a. $[[dance]] = \lambda e_{[e \in Dom(agent, \dots)]}. e \in dance$
 b. $[[John]] = j$
 c. $[[AGENT]] = \lambda E. \lambda x. \lambda e_{[e \in Dom(E)]}. E(e) \wedge agent(e) = x$
 d. $[[AGENT danced]] = \lambda x. \lambda e_{[e \in Dom(agent, \dots)]}. e \in dance \wedge agent(e) = x$
 e. $[[John [AGENT danced]]] = \lambda e_{[e \in Dom(agent, \dots)]}. e \in dance \wedge agent(e) = j$
- (33) a. $[[house]] = \lambda x_{[x \in Dom(loc, \dots)]}. x \in house$
 b. $[[the hill]] = iy. y \in hill$
 c. $[[on]] = \lambda y. \lambda F. \lambda x_{[x \in Dom(F)]}. F(x) \wedge loc(x) = above(y)$
 d. $[[on the hill]] = \lambda F. \lambda x_{[x \in Dom(F)]}. F(x) \wedge loc(x) = above(ix. x \in hill)$
 e. $[[house [on the hill]]] = \lambda x_{[x \in Dom(loc, \dots)]}. x \in house \wedge loc(x) = above(ix. x \in hill)$

With this background, we can define E-TOPIC as a role-function and introduce it into the lexical entry of all expressions that actually have that role. An example is given in (34). We used the symbol ψ to mark the internal argument of *about* as a signal that we do not wish to engage in a discussion about the type of this internal argument. We will limit ourselves here to individuals. Of course, instead of the verb *to complain*, we could have any topic-compatible verb or nominal expression. See Rawlins (2013) for a tentative list of such expressions.

- (34) a. $[[complain]] = \lambda e_{[e \in Dom(agent, content, etopic)]}. e \in complain$
 b. $[[about]] = \lambda \psi. \lambda E. \lambda e_{[e \in Dom(E)]}. E(e) \wedge etopic(e) = \psi$
 c. $[[about John]] = \lambda E. \lambda e_{[e \in Dom(E)]}. E(e) \wedge etopic(e) = j$
 d. $[[complain [about John]]] = \lambda e_{[e \in Dom(agent, content, etopic)]}. e \in complain \wedge etopic(e) = j$

So, what are the exact truth conditional implications of the E-TOPIC role? We will need to assume that verbs and nominals that can have an E-TOPIC come with a hidden relational argument R . Some examples are given in (35).

- (35) a. complain: $R = \lambda x. \lambda y. agent(y)$ expressed x
 b. talk: $R = \lambda x. \lambda y. agent(y)$ expressed verbally x
 c. book/letter: $R = \lambda x. \lambda y. x$ is depicted inside y .

We do not wish to dwell upon the exact status of this relational argument in the semantic composition, as this is a complication orthogonal to our aims. The only thing we need is that the head noun or verb will make this relational predicate salient, regardless of the way this is achieved at a technical level. This will allow us to explicate the E-TOPIC role as in (36).

- (36) The E-TOPIC postulate:
 $etopic(a)=b$ iff there is a discourse D such that $content(a)=content(D)$ and b is the

combine with such a theory include Beaver and Krahmer (2001) and Onea (2015). Solving these issues does not seem to be very difficult even in more standard setups, however.

discourse topic of D and $R(D, a)$, where R is a relational variable lexically determined by the matrix predicate

To see what this postulate yields, consider two examples:

- (37) a. *John complained about Peter.*
 b. *John read a book about Peter.*

In (37a), Peter is the E-TOPIC of the complaint that John made. This means that there is some discourse D , such that the content of John's complaint equals the content of D . Moreover, Peter is the discourse topic of D . And finally, due to the relational variable made salient by the verb *complain*, John expressed D .¹⁶ Similarly, for (37b), Peter is the topic of a book. This means that there is a discourse D , such that the content of D equals the content of the book under discussion, Peter is the discourse topic of D . Moreover, due to the lexical import of *book*, we require that D is depicted inside of the book, for example by being written in Latin characters, pictograms or a code nobody except John can understand. In other cases, such as (37a), it is well possible that D contains exactly one sentence. In this case, being the discourse topic would entail being the *aboutness*-topic (in the sense of Reinhart 1981) of the relevant sentence.

Consider now the interpretation of sentences such as the ones in (38), which are translated in this framework into the (simplified) formal representations in (39). Since the lexical semantics of verbs like *tell* or *complain* request that their possible (event-) arguments be in the domain of the E-TOPIC, it follows that every event of *telling* or *complaining* will have such a topic. However, since these sentences do not mark the E-TOPIC, it is not possible to infer what the embedded topic of these verbs is. While one can consider it probable that in (38a) the most likely events that would satisfy the existential quantifier would have Max as an E-TOPIC, it is perfectly possible to imagine plausible scenarios in which this is not the case. For example, Peter's complaint could be about a care-taker whose job was to make sure Max does not dance, or about a particular song that made Max dance, etc. Thus, just because something is the subject of an embedded clause does not entail that it will also be the E-TOPIC of the respective matrix verb.

- (38) a. *Peter complained that Max danced.*
 b. *Peter told Alfred that Max danced.*
- (39) a. $\exists e.e \in \text{complain} \wedge \text{agent}(e, \text{peter}) \wedge \text{content}(e, \text{Max danced})$
 b. $\exists e.e \in \text{tell} \wedge \text{agent}(e, P) \wedge \text{theme}(e, A) \wedge \text{content}(e, \text{Max danced})$

Before closing this section, we wish to stress that we have neither suggested nor do we wish to suggest that *about* can only express E-TOPIC,¹⁷ nor is it the case that

¹⁶We gloss over aspects related to the time and place of John expressing D etc., as they are orthogonal to our goals in this article and do not seem to introduce serious difficulties.

¹⁷In general, we assume that prepositions may encode a multitude of possible roles, and that whatever a preposition modifies, for example a verb, will lexically specify what roles their event argument is guaranteed to be in the domain of. Now, when combining a verb with a preposition, only those readings are computed for which the composition guarantees no clash.

E-TOPIC can only be expressed by *about*. For example, (40a) is a clearly different usage of *about* (see Zaroukian 2013). The analytical options that come to mind for (40b) seem to be quite numerous, and not all of them commit to the *about*-phrase being an E-TOPIC.

- (40) a. *Migrant women represent **about** half of international migrants, but make up over 70% of all trafficking victims and 96% of victims trafficked for sexual exploitation.*
https://refugeesmigrants.un.org/sites/default/files/pv_day2_ag1_unwomen.pdf
- b. *She is very knowledgeable **about** all aspects of the real estate business.*
<https://www.pamboyle.com/testimonial/she-is-very-knowledgeable-about-all-aspects-of-the-real-estate-business/>

Similarly, (41a–b) illustrates that the preposition *on* in English can in some cases mark the same semantic relation as *about*. This is not incidental, as we find this pattern in other languages. French *sur* ('on') is the default marker of the E-TOPIC role, as in (41c), similarly, the German *über* ('on'/'above') that is used to mark the E-TOPIC relation has a clear special meaning similar to *on*, as shown in (41d).

- (41) a. *Some lenders like ING Direct do not negotiate **on** the interest rates for their home loans.*
<https://www.homeloanexperts.com.au/managing-your-home-loan/negotiating-your-interest-rate/>
- b. *Let's say you're writing a paper **on** global food distribution, and you've chosen to compare apples and oranges.*
<https://writingcenter.fas.harvard.edu/pages/how-write-comparative-> analysis
- c. *François Hollande prépare un livre **sur** son quinquennat.* (French)
 Francois Holland prepares a book sur his five-year term
 'Francois Holland is preparing a book about his five-year term'
 Europe 1, February 13, 2018
- d. *Arnold Stadler hat ein Buch **über** die Liebe geschrieben.* (German)
 Arnold Stadler has a book über the love written
 'Arnold Stadler has written a book about love'
 Die Zeit, June 27, 2007

5. ON THE ROOTS OF DIFFERENTIAL OBJECT MARKING IN ROMANIAN

Having provided an analysis of *about*-PPs as E-TOPICS in the previous section, we are now in the position to provide the missing details of our analysis of how grammaticalization of *pe* could have occurred, sketched at the end of section 3.2. In order to do so, we will first show that in Old Romanian, *pe* was indeed used to mark E-TOPICS. This is not entirely trivial and will require a detailed argumentation.

Many ambiguities will remain, and both syntactic and pragmatic factors may further constrain the eventual empirically attested reading or readings.

Moreover, once this is established, we will consider the broader theoretical implications of this finding in itself. In the next step, in sub-section 5.2, we provide the details of the argument-alternation associated with *pe* outlined in section 3.2 and how the re-analysis of *pe*-PPs suggested there could have happened. Finally, we pull all of the strings of our analysis together in section 5.3, and provide an overall picture of how DOM may have begun to emerge in Old Romanian.

5.1 *Pe* and E-TOPICS in Old Romanian

In this section, we provide evidence that *pe* could be used to mark the E-TOPIC role in certain cases in Old Romanian.

We start with reconsidering the examples given in (15), repeated here as (42). We are now in a position to show that these examples can be naturally analyzed as involving the E-TOPIC role marked by the preposition *pe*. As far as the examples (42b) and (42c) are concerned, it should be self-evident that they could be analyzed as involving the E-TOPIC role, hence we see no reason to comment further on them; this would constitute sufficient proof that *pe* can mark E-TOPIC. However, it seems to us that with some detailed explanation, the other two examples can be understood as involving E-TOPICS as well.

- (42) a. *Și cheltuiala măruntă pe treaba lu voievod*
 and expense.DEF small pe business.DEF his ruler
 ‘And small expenses for / concerning the ruler’s activity’
- b. *am scris aceasta carte lu Panaet de Hotărani,*
 have.1SG written this book dat Panaet of Hotarani
pe ocină ce i=am vândut iară în Hotărani.
 pe land that CL.3SG.DAT=have.1SG sold again in Hotarani
 ‘I wrote this letter to Panait from Hotarani, about/concerning the land I have sold to him in Hotarani’
- c. *că am avut pîră amîndoi pe nește rumîni din Răsnicele*
 that have.1SG had complaint both pe some servants from Rasnicel
 ‘because both of us we had a complaint about/concerning some servants from Răsnicele’
- d. *de ne-au tocmît pe acești rumîni ce sînt mai sus-scriși*
 that us-have negotiated pe these servants that are more above-written
 ‘so that they have negotiated with us concerning the above mentioned servants’

Concerning (42a), one can assume that the noun *cheltuială* (‘expense’) takes a complement that specifies the object that caused the expenses. At first sight, it may not be obvious why the preposition *pe*, with a core locative meaning, can be used to mark this complement of *cheltuială*. Also, since *expenses* do not seem to involve speech acts, the idea that *pe* may mark the E-TOPIC role does not seem very natural either. However, we do find the preposition *on* in English as a more or less idiomatic complement marker of *expense* and more generally of *spend (money)*, as witnessed in the examples in (43).

- (43) a. *Here are a few simple tips for uncovering the future potential expense on your rental property.*
<https://www.biggerpockets.com/renewsblog/2014/12/02/rental-property-expenses/>
- b. *There are some things I never spend money on.*
<https://centsiblyrich.com/38-things-i-never-spend-money-on/>

One possibility may be that these constructions may have evolved from locative constructions as in (44) and may then have undergone reanalysis: locations and land in particular are often not only the place where you spend money, as in (44), but are also prime objects one purchases with money, for example, when buying fields, dominions, etc.

- (44) *These options expanded the amount of money users could spend on the marketplace, which in turn, attracted more users to the site.*
<https://blog.opskins.com/a-year-in-the-life/>

Crucially, however, we wish to suggest that in such constructions the range of actual inferences associated with the *pe*- or *on*-argument of expenditure-verbs is quite large, for *pe* can mark either the object purchased, the object causing the expenditure, or just a general description of the domain in which the expenditure takes place. A simple way to unify these usages is to think of the *on*-argument of expenditure-verbs as the topics of negotiation or transaction (arguably) involved in the expenditure. Of course, transactions involved sufficient speech acts to justify speaking of a topic in the relevant period of time for our Old Romanian data (as compared to current click-and-buy-expenditure).¹⁸ Moreover, the fact that we do find examples in Old Romanian with *tocmi* ('negotiate') having *pe*-arguments that mark the object of the negotiation, as in (42d), adds plausibility to our hypothesis, simultaneously providing a unified analysis for the four examples under discussion.

It is important, however, to see that *pe* was not the only expression that could mark E-TOPIC. For example, Mardale (2018) shows that the preposition *spre* ('towards') could be used for this purpose as well. Moreover, while proving a negative with limited historical data is not possible, it seems to us that *pe* could mark E-TOPICS for only a limited range of verbs/nominals. In fact, the empirical picture in Old Romanian appears to be similar to the ability of English *on* to mark E-TOPIC. For example, while we have seen above some examples where *on* can be used to mark E-TOPICS, as in (41), it is easy to find examples where *about* is needed, as witnessed in (45). (Modeling this issue formally, of course, goes beyond the ambitions of this article.)

- (45) a. *Skylar knows about/#on Mary that she is rich.*
 b. *Skylar said about/#on Mary that she is rich.*

The fact that *pe* in Old Romanian could be used to mark E-TOPICS (though it is probably neither the only nor a general marker of this category), has, in itself, important theoretical consequences. In particular, it provides a new and important way in which the notion of topicality in general is connected to the emergence of DOM,

¹⁸We discuss this issue in more detail in the next section; see example (53).

complementing the theoretical inventory hitherto consisting of left-dislocated topics (Iemmolo 2010) and secondary topics (Dalrymple and Nikolaeva 2011). While the Old Romanian data are not sufficient in quantity to establish this specifically for the *pe*-marked cases, at least analytically, we would expect that E-TOPICS have similar referential properties to other syntactically marked topics, such as left-dislocated topics. If so, in the light of any theory that explains the origin of DOM with reference to the prototypical properties of topics, the simple fact that *pe* could be used to mark E-TOPICS may be in itself a sufficient reason for *pe* to grammaticalize into a DO marker, as discussed in more detail in section 5.3.

5.2 Alternation and re-analysis

Having established that *pe* in Old Romanian can be used to mark E-TOPIC, we now turn to the puzzling argument alternation raised in section 3.2. Our first task will be to establish the semantic arguments of the relevant verbs: *jelui* ('complain/mourn') and *tocmi* ('haggle', 'hire'). This will allow us to solidify the facts pertaining to the relevant alternation and to the re-analysis postulated in section 3.2. We do this for the two verbs independently.

We begin the discussion with example (18), repeated as (46).

- (46) a. *s=au jeluit sluga noastră Ștefan pre Samoil*
 CL.REFL.3SG=has complained servant.DEF our Stephen pe Samuel
 'And our servant Stephen complained about Samuel'
- b. *jeluiște sluga noastră Ștefan Moimăscul pre niște cuconi mici*
 mourns servant.DEF our Stephen Moimăsc.DEF pe some children little.pl
 'Our servant Stephen Moimăsc mourns for some (little) children'

We assume that the verb *a jelui* may assign at least four semantic roles: AGENT, E-TOPIC, THEME and CONTENT. The AGENT of the verb is the one who does the complaining. In other words, the agent expresses some content by performing a speech act or a series of speech acts. The CONTENT is the propositional content of the speech act performed by the agent, which of course need not be a quotation or a direct paraphrase; it can also be an interpretation of that speech act. The CONTENT-argument often appears implicit and can be pragmatically reconstructed. The THEME – we argue – is the person affected by the CONTENT. Finally, the E-TOPIC is the (discourse/aboutness) topic of the speech act that was performed by the agent. While three of these roles seem to be quite natural, the idea that the verb *a jelui* may have a THEME argument of the kind suggested may be surprising at first sight. However, it seems to us that good evidence exists in support of our suggestion.

As a preliminary note, we observe that the notion of complaint usually incorporates that the one who does the speech act of complaining is the one affected by whatever the complaint is about. One cannot complain about the weather on Mars, unless one has either real estate, vacation plans (or comparable business) on Mars. However, at least in English, one can complain *on behalf* of someone, as shown in examples such as (47). Especially in the case of (47b), it is clear that this is not a simple *instead of*-reading. This additional argument may not be the most prototypical THEME argument, as one may expect some benefactive/malefactive component.

However, this is a clearly distinct kind of argument compared to typical benefactive/malefactive arguments often marked by *for* in English as in (48).

- (47) a. *The adoptive parents **complained on behalf of** the child that the refusal to allow him entry into the UK infringed his human right to family life under Article 8 of the European Convention on Human Rights.*

Suhraiya Jivraj: The Religion of Law: Race, Citizenship and Children's Belonging

- b. *Ruskin also developed into a reform-minded social critic who **complained, on behalf of** the working class, about capitalist tendencies.*

Kerry Powell and Peter Raby: Oscar Wilde in Context

- (48) a. *John complained **for** Michael.*
 b. *John complained **on behalf of** Michael.*
 c. *John had breakfast **for** Michael.*
 d. # *John had breakfast **on behalf of** Michael.*

Moreover, in several languages, complaint-like verbs are reflexive. Examples include *a se plânge* in modern Romanian and *sich beschweren* in German. In these languages, it seems that it is not idiomatic to add the *on behalf* type of argument. For example, while in German (49a) is grammatical, it simply has the same benefactive reading as in (49b), and (at least for many speakers) not the reading according to which Alfred was affected by something bad and Peter did the complaining without being personally affected or intending to do a favor for Alfred. (Peter might have complained based on a general feeling of injustice.) Comparatively, in English there is a clear difference between simple benefactive arguments as in (48a) and (48c), which are similar to the German examples in (49), and the variants (48b) and (48d), which have no structurally similar German counterparts. The same argument holds for modern Romanian, where *a se plânge* tends to resist a usage in which the agent is not affected by the content of the complaint. Native speakers of Russian and Serbo-Croatian have reported similar judgments to us, where, again, complaint verbs have a reflexive form.

- (49) a. *Peter hat sich **für** Alfred *beschwert*.*
 Peter has REFL.3SG *für* Alfred complained
 'Peter complained for Alfred/on behalf of Alfred.'
 b. *Peter hat **für** Alfred *gefrühstückt*.*
 Peter has *für* Alfred eat-breakfast
 'Peter ate breakfast for Alfred.'

This may naturally be explained by assuming that the reflexive argument grammaticalized the identity between the affected person (the *THEME* argument) and the *AGENT* of the complaint.¹⁹ With this in mind, it seems fair enough to think of the affected person argument of *jelui* in Romanian as the *THEME* argument. Nevertheless, it is

¹⁹Virginia Hill (p.c.) pointed out to us that there could be an alternative to our assumption that the reflexive is actually an argument of the verb in Old Romanian. The analysis of Armstrong (2013) of Agentive Reflexive Clitic vs. Transitive *se*-Clitic constructions in

proper to mention that a full analysis of the semantic content of *on behalf* (which we cannot provide here) would be necessary to complete this argument.

We can now provide a detailed analysis of both structures in (46).

For (46a), we suggest that the AGENT and the THEME are the same person. The reflexive construction amounts to the semantic binding of the reflexive by the subject: the AGENT surfaces as the subject while the THEME is a reflexive object. The E-TOPIC in this case is the oblique PP headed by *p(re)*. The CONTENT of the complaint is implicit, but could be made explicit with an additional subordinate clause headed by (*pentru*) *că* ('because/that') or some other oblique nominal construction. The structure is schematically illustrated in (50a).²⁰

In contrast, for (46b), and very similarly for the modern Romanian structure in (22b), the most natural analysis is suggested to be (50d), in which the THEME is not actually affected by the CONTENT (which would be the bad thing that happened to the theme) but instead by the mourning ritual itself.²¹ But then, there is the question how these entirely different argument structures may have arisen, and why *pe* was used in both. Hence, we assume that in the process of developing the argument structure in (50d), there were two intermediary steps. In particular, we start out with the full four arguments in (50a). Again, the AGENT is expressed as the subject, but this time the THEME and the AGENT are no longer marked as co-indexed, thus they are probably not. Then, the THEME and the E-TOPIC can be assumed to be identical in some cases: these are the cases in which the complaint affects the person who is the topic of the speech act of complaint. The CONTENT, again, is left implicit, but in most cases it is understood as the property of the E-TOPIC having died or having suffered something comparable (gone missing, gone to jail). Clearly, there is a good enough sense in which one can be affected by death and thereby fulfill what we require for the THEME argument. Crucially, we do not assume that in (46b) the THEME and the E-TOPIC are realized by the same expression. We merely suggest that they are referentially identical (i.e., they are the same individual). Only one of them is actually overtly expressed, the other one is implicit. In other words, either

Spanish could potentially be extended to the data we discussed above. This would, however, not contradict our basic suggestion that *a jelui* has a theme argument.

²⁰Representing schematically missing arguments in this way is not meant as a structural analysis. We remain absolutely non-committal with respect to the syntactic representation of non-overt semantic arguments, and to the stage at which the co-reference is established (e.g., whether it is syntactic or pragmatic). It is our aim to remain as general as possible and not make syntactic assumptions unless they are crucial for our argumentation.

²¹The theme seems to be affected by having its chances for a better afterlife or its honour within the community or family boosted. This does seem quite natural, also from the perspective of English. The fact that English can conceptualize the argument of *mourn* as a *for*-PP, may suggest some sort of benefactive reading.

the E-TOPIC or the THEME could be realized by *pre niște cuconi mici*. This gives us two different ways to analyze (46b), schematically represented in (50b) and (50c).

- (50) a. AGENT:S_i jelui THEME:refl-ACC_i E-TOPIC:pe-DP_j CONTENT:zero
 b. AGENT:S_i jelui THEME:zero_j E-TOPIC:pe-DP_j CONTENT:zero
 c. AGENT:S_i jelui THEME:pe-DP_j E-TOPIC:zero_j CONTENT:zero
 d. AGENT:S_i jelui THEME:pe-DP_j

The schema in (50b/c) provides exactly what we hoped for in section 3.2: it turns out that the THEME and the E-TOPIC arguments of (46b) could, at least at some point in the historical development of Romanian, have been co-referential. And if this is true, the re-analysis of an E-TOPIC as a theme for such examples appears to be entirely plausible. In particular, all we need to assume is that a re-analysis of (50b) as (50c) took place before the 16th century.

Some more evidence for our analysis comes from example (51)²², which is particularly intriguing because it is one of the few cases in which we find the verb *jelui* (this time written as *jăluit*, a dialectal variant) in a context in which two arguments are present. The first argument is the person who has been affected by some unfortunate event (chief Grigorie) and the second is the actual unfortunate event that happened to this person (he lost money). We have argued above that when *a jelui* appears as non-reflexive, the affected person is more likely analysed as a THEME, whereas when the verb is reflexive, the affected person is probably analysed as an E-TOPIC instead. Hence, we argued, DOM arises from the re-analysis of an E-TOPIC as a THEME. Crucially, in this example we also find the CONTENT argument, namely what happened, what has been reported during the complaint (as a speech act). Interestingly, in this – rather late – example, the affected argument clearly surfaces as a THEME but we still have a CONTENT-argument, as predicted by our analysis. In other words, we have the situation in (50c).

- (51) *Și au jăluit și pre Grigorie-vodă de bani ce le=au luat*
 and have complained also pe Grigorie-chief of money that CL.3.PL=have taken
 ‘And they have also complained for Chief Grigorie because of the lost money.’

In contrast, in modern Romanian, we do not have a CONTENT-argument marked with *de* but rather an explanative construction, as shown in (52). This clearly suggests that the verb has completely altered its argument structure from E-TOPIC- and CONTENT-marking to simple THEME-marking: from (50c) to (50d).

- (52) *Bunica l=a jelit convinsă că murise*
 Grandmother 3SG.CL.ACC=has mourned convinced that died
 ‘Grandmother has mourned him because she believed that he died.’

Octavian Paler: *Deșertul pentru totdeauna*, 2016

We can essentially repeat this analysis for (19), repeated here as (53). In fact the re-analysis may strike one as even clearer in this case.

²²The example stems from Ureche’s chronicle from the 17th century; see footnote 26 for details.

- (53) a. *se=au* *tocmit pre megiaș[i]*
 CL.REFL.3PL=have haggle pe neighbours
 ‘They haggled with their neighbours’
- b. *amu tocmit pre Mușat postelnic și pre Negoe pîntru 4*
 I.have hired pe Mușat chamberlain and pe Negoe for 4
rumîni vii și pîntru 6 delniși, firî rumîn
 serfs alive and for 6 plots-of-land without serf
 ‘I hired Musat the chamberlain and Negoe in exchange for 4 living serfs
 and for 6 plots of land, without serfs’

For (53a) we suggest that the verb *tocmi* (‘haggle’) assigns the semantic roles AGENT, THEME, E-TOPIC, and CONTENT. To see how the roles are distributed, consider an instance of haggling as a series of transactional speech acts with the following structure: One person says: “*I want you to give me x object for y money*”. The other person says “*I want you to give me y’ money for x object.*” What seems to be clear that the object that is haggled over is the E-TOPIC and the actual offers of the kind mentioned here are the CONTENT. Given the near-symmetry between the roles of both interlocutors it is unsurprising that we get a reflexive structure. What may be surprising is that the structure appears with an accusative and not with a dative reflexive. One conceivable explanation may be that haggling was metaphorically conceptualized (in the sense of Lakoff and Johnson 1980 and subsequent literature) as some sort of a fight or imposing one’s will upon the other in Romanian: thus, in (53a), the AGENT is the one doing the haggling speech act, whereas the THEME is the one who is combatted with in the speech act iteration scenario.²³ The schematic representation in (54a) naturally captures this situation.

Compare (53b), where the verb *tocmi* (‘hire’) changes its semantics and argument structure only slightly. Again, there is a series of speech acts associated with negotiation, but this time the person who is being hired is the same person with whom one negotiates.²⁴ Crucially, this person now has two roles: they are the partner combatted with in the negotiation as well as the target of negotiation. Thus, this time, it is very natural to assume that the E-TOPIC and the THEME are the same entity, as suggested in (54b) and (54c). Probably, in the first step, the E-TOPIC was overtly expressed and the THEME pragmatically reconstructed, but because of the co-reference the structure may have been re-analyzed as *pe*-marking the THEME.

²³As pointed out to us by an anonymous reviewer, it is not obvious that theme is the best choice as a semantic role for our example. From the metaphorical perspective suggested, probably patient would be a better choice. This applies to *tocmi* ‘hire’ as well. We will stick with theme here as a generic notion, as a distinction between these two roles does not seem to add much to our argument.

²⁴Note that, especially in the Middle Ages, this was not always the case: one could well negotiate with someone else about a worker, serf or slave. In fact, probably there might have been a gradual process in which the same verb was used in the face of changing cultural negotiation scenarios – which additionally suggests that the change in semantics mentioned above was likely to occur.

Eventually, in modern Romanian the verb *a angaja* ('to hire') no longer seems to have an E-TOPIC or a CONTENT argument, as suggested in (54d).

- (54) a. AGENT:S-pl_i tocmi THEME:refl-ACC-pl_i E-TOPIC:pe-DP_j CONTENT:zero
 b. AGENT:S_i tocmi THEME:zero_j E-TOPIC:pe-DP_j CONTENT:zero
 c. AGENT:S_i tocmi THEME:pe-DP_j E-TOPIC:zero_j CONTENT:zero
 d. AGENT:S_i angaja THEME:pe-DP_j

5.3 The grammaticalization scheme of *pe* as a DO marker

Above we have shown how using the semantic role E-TOPIC allowed us to find examples in which *pe*-marked DPs could be re-analyzed: from PPs marking the E-TOPIC role to DPs marking the THEME role. In this section, we wish to reflect upon how this could trigger a move towards DOM in Romanian.

To start with, we are not actually claiming that a re-analysis of any of these particular verbs literally triggered DOM in Old Romanian, in the sense that the first instances of DOM came from these alternations. We could imagine that this is the case, but given that we have no data from the time at which DOM began being used in Romanian, we cannot tell with reasonable certainty. What we wish to put forward, therefore, is the more cautious claim that there were a range of verbs which allowed the re-analysis of *pe*-PPs as THEMES (and potentially other semantic roles prototypically appearing as DOs). This re-analysis may have occurred before DOM, as such, emerged. Hence the *pe*-marking THEMES/DOs in the type of structures we discussed may have been available in the language prior to the emergence of DOM.

Some skepticism regarding whether these initial instances of *pe*-marking were DOM seems appropriate. For example, E-TOPICS marked with *pe* did include [-animate] entities which would not have been the usual suspects to be the first marked by DOM. Hill and Mardale (2019), however, do claim that some limited [-animate] *pe*-marked DOs existed in Old Romanian. Additionally, there does not seem to be any well-behaved semantic role that is systematically associated with the notion of being a DO (see, a.o., Kratzer 2002, Hole 2014), even though the role THEME is most intrinsically connected to DOs. Also, we have no evidence that DOM even began by marking THEMES in general in Romanian. If the process was similar to Spanish (see von Heusinger 2008), verb classes may have played a role in the evolution of DOM as well, and we cannot tell, given the late stage of DOM our data covers, whether the initial stages of DOM were even remotely connected to the particular verb types exhibiting the E-TOPIC THEME alternation. Furthermore, we have no evidence suggesting that E-TOPICS are typically personal pronouns or predominantly proper names, entities that are argued to be the initial targets of DOM in Romance languages (though some correlation between such features and being E-TOPIC probably existed).

Therefore, instead of suggesting that the re-analysis of PPs as DOs literally initiated DOM, we are suggesting that the emergence of DOM may have been driven by the three following factors. First, the re-analysis discussed above created

a strong association between *pe* and DOs, an association that was lacking for any other competing preposition.²⁵ Second, this re-analysis provided the grammatical system with examples of marked DOs. Finally, by definition, the E-TOPIC role is closely connected to the notion of topicality in general; we therefore expect that the prototypical E-TOPIC has the same referential and animacy features as the prototypical information structural topic. Hence, the use of *pe* in its function to mark E-TOPICS may have been correlated with the feature bundle that typically triggers DOM in language after language. This may also have facilitated the use of *pe* for marking DOs, in terms of a kind of associative bridge based on feature similarity.

When the functional pressure to mark DOs was high enough, these three factors jointly may have made it possible for the DOM process to begin, and *pe* would have been the most available candidate preposition to become the DO marker. We summarize these findings in one general scheme of the source of grammaticalization of DOM in Romanian in Figure 1. As we see it, the essential cornerstone of this grammaticalization scheme is that *pe* marks the E-TOPIC semantic role.

Finally, we wish to stress, as mentioned in our introduction, that one could imagine a fairly simple theory of the grammaticalization of DOM in Romanian that would go like this: E-TOPICS are, by definition, topics, and thus they share prototypical features with subjects. If *pe* is used to mark E-TOPICS, it does, thereby, also mark prototypical subjects which are – in fact – not subjects of the respective content embedding verb. DOM, in markedness reversal theories, can be conceived of as a very similar phenomenon: marking prototypical subjects that are in fact not subjects, but DOs. Then, one could say that this is all that needs to be said about the origin of DOM in Old Romanian. For all we know, this could be a valid story and – certainly – it is compatible with all findings we presented in this article. Indeed, our analysis could be correct in all details and still this narrative could be told in precisely this way. Moreover, the fact that *pe* could mark E-TOPICS might explain the grammaticalization of DOM in Romanian – at a more general level – even if our particular analysis of the argument alternation was plainly misguided or incidentally immaterial to the historical processes that lead to DOM.

The goal of this article, however, was not to provide a theory which is compatible with the facts and explain the grammaticalization of DOM at a general level, but to try and go as deep as possible into the details of a step-by-step reconstruction of how *pe*

²⁵Another question to ask at this point is why the DOM marker was not a dative marker in Old Romanian, as in other Romance languages. After all, Old Romanian had the prepositional dative marker *la* ('to', 'at') as in the construction *a spus la boiari* ('told to the nobles'), which is still available in modern Romanian. It is not entirely clear whether there was any particular reason for this marker not to become the DOM marker in Romanian. However, Mardale (2015) points out that there is a series of verbs with beneficiary arguments that typically appear in dative such as (*a dăru* 'to give as a present', *a mulțumi* 'to please') which oscillate between morphological dative (*a dăru* *cuiva*) and prepositional accusative (*a dăru* *pe cineva*) instead of prepositional dative (*a dăru* *la cineva*). The fact that prepositional dative was apparently blocked in these constructions may have additionally contributed to choosing a distinct marker for DOs.

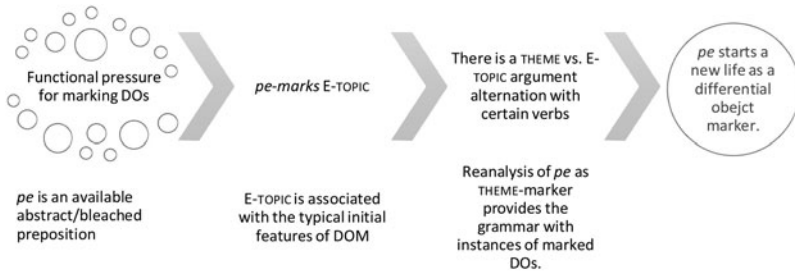


Figure 1: Grammaticalization source of *pe* as a DOM marker in Romanian.

could have ended up as a DOM marker. We believe that the alternations we uncovered may indeed be a valid grammaticalization mechanism, and that analogous constructions should be scrutinized cross-linguistically.

6. DISCUSSION

In this section, we provide some additional empirical motivation for our approach and discuss a number of problems our analysis potentially has. We start with a few examples that show similar alternations either in Old Romanian or diachronically, from Old to modern Romanian, and thereby provide further support for our general line of attack in this article.

We first discuss some examples from the corpus we based our study on. The crucial aspect in example (55) is that we find *pre* as marking the argument of a reflexive verb *milostivească* ('to be merciful'). This cannot be the DO, because the verb is reflexive.

- (55) *să se milostivească pre noi și pre ceastă săracă de țeară*
 that REFL give-mercy pe us and pe this poor of country
 'the he should be merciful with us and our poor country'

The verb *milostivească* and variants such as *a milui* ('to give/have mercy') exhibit, even in modern Romanian, exactly the same argument alternation we discussed above (and in this case there is little doubt that these are argument alternations). This is shown in (56a–b). One should note that the version with a DO in (56a) is mostly confined to religious contexts. Moreover, the reflexive construction with a *de*-marked argument in (56b) is not viewed as idiomatic by all speakers (though is found quite frequently in a web-based survey and also occasionally in standard newspapers). Finally, there is a nominal construction variant with a light verb with a very similar meaning, which is highly frequent in modern Romanian and given in (56c). Here, again, we find a *de*-marked argument of the predicate.

- (56) a. *M=a miluit pe mine.*
 CL.1SG.ACC=has felt-pity pe me.acc
 'He helped (had mercy on) me' (lit.: he gave-mercy to me)

- b. *femeia s=a miluit de el*
 woman.DEF REFL.3SG=has felt-pity de he
 ‘The woman helped him’

Adevărul, November 25, 2009

- c. *I=a fost milă de mine.*
 CL.3SG.DAT=has been mercy de me
 ‘He was merciful to me’

The crucial observation is that there is a clear contrast between the Old Romanian reflexive construction with *pe* given in (55) and the modern Romanian variants in (56b) and (56c). Arguably, in modern Romanian, in the latter two cases the E-TOPIC role of the verb *milui* (‘to give/have mercy’) and the construction *a fi milă* (‘to feel pity for somebody’) remains overtly expressed, and for this the preposition *de* is used. This is perfectly in line with our observation above that in modern Romanian *pe* no longer systematically marks the E-TOPIC role (with some exceptions, mentioned above in connection with the examples in (17)). Instead this role is marked either with the preposition *de* or the (lexically more explicit) E-TOPIC marker *despre* or *asupra* (see Mardale (2018). However, the variant in (56a) additionally shows that a transitive version of this construction also developed, and it seems reasonable to assume that in modern Romanian we are really witnessing the split in the marking of argument structure: the E-TOPIC gets marked with *de* and the THEME/DO is marked with *pe*. Therefore, our example shows that *pe* has been used in a conceptual domain which facilitates the respective argument alternation, as witnessed by the modern Romanian examples, thus adding further support to the idea that argument-alternating verbs may have been a major source of the grammaticalization of *pe* in Old Romanian.²⁶

Another very similar example from Old Romanian is (57). The construction *sa aibă grijă* (‘take care’) appears with a *pe* argument, which is – again – clearly prepositional. As shown in the examples in (58), from modern Romanian, we find the same alternation between *pe* and *de*, where *pe* marks a DO and *de* a more or less prototypical E-TOPIC.

- (57) *Măria lui cu țara împreună acea grijă să aibă, pre ceaste 2 țări.*
 majesty his with country together that care to have pe these 2 countries
 ‘The king and his country should take care of these two (other) countries.’

- (58) a. *Ion are grijă de mama lui.*
 John has care de mother.DEF his
 ‘John takes care of his mother’

²⁶Of course, one may ask whether the argument of *a se milui/a milui* is really an E-topic. If we take our strict definition, this is probably not the case because it is not obvious that any speech act is involved (though there might be some performative element of showing mercy in older religious contexts). However, as with semantic roles in general, we have more or less prototypical cases, and this example naturally falls in a grey transition zone.

- b. *Ion se îngrijește de mama lui.*
 John refl.3SG. caretakes de mother.DEF his
 ‘John takes care of his mother’
- c. *Ion o îngrijește pe mama lui.*
 John CL.3SG.ACC. caretakes pe mother.DEF his
 ‘John takes care of his mother’

Finally, we consider the example in (59). Again, we find a reflexive structure, suggesting that the verb *a peți* (‘ask for somebody’s hand’/‘propose to someone’) takes a non-DO argument which is marked with *pe*. Crucially, in modern Romanian, we again find a clear *pe*-marked version with DO, as seen in (60). We can again argue that in Old Romanian, the argument of *a peți* is an E-TOPIC. This time this is a fairly clear case: in the context of Old Romanian, the marriage arrangement was a matter of complex negotiation, hence the meaning of *a peți* was (probably) just a lexically more specialized version of verbs of negotiation for human referents. In comparison, as cultural changes occurred over the history, the act of asking for someone’s hand became a TOPIC-less speech act and the bride became a standard DO, similar to our observation in Footnote 25. We see here a case of historical transition from E-TOPIC to THEME (as a clear reanalysis alongside cultural change).

- (59) *Și, peîndu=se pre ea, lăuda=se că iaste neğuător mare ...*
 And proposing=REFL.3SG pe she, praised=himself that is trader big ...
 ‘And proposing to her, he praised himself as a successful businessman.’
- (60) *Băsescu a pețit=o pe Ela*
 Basescu has proposed=CL.3SG.ACC pe Ela
 ‘Basescu has asked for Ela’s hand (on behalf of someone else).’

Actualități Arad, June 10, 2016

We now move to a few further examples from other texts from Old Romanian from the 17th-18th centuries. While these examples are slightly later than the data discussed above, we deem them useful to complement the diachronic picture we are trying to draw and thereby provide further evidence for our theory. The examples stem from Ureche and Neculce’s chronicles (around 1740) telling the history of Moldavia²⁷. In particular, we provide a few further examples of the use of *pe* in Old Romanian as a marker of E-TOPIC in nominal constructions in (61a), (61b) and (61c).

- (61) a. *cum au început a scrie cărți pre la cele obuzuri...*
 how have begun to write books pe at those troops
 ‘and they started to write books on those troops’
- b. *înțelegînd că au fost scos acolo hîrtii pre țară*
 understanding that have been extracted there papers pe country
 ‘understanding that they have taken out some papers about the country’

²⁷See Iordan (1955).

- c. *să=ș* *aducă scrisori ce or avè pre moșiile lor*
 to=CL.refl.3SG bring letters that will have pe lands their
 ‘to bring whatever property letters they have about their lands’

Another example with verbal embedding, in which the *pe*-phrases appear as arguments that can be either analysed as E-TOPIC or as some related, potentially even more abstract, concept of GOAL, is given in (62). Notice, however that we have no evidence that *pe* could be used for GOAL-marking in general in Old Romanian, hence the analysis of these *pe*-phrases as E-TOPICS is more plausible.

- (62) *căci eu n=am venit pe bătaie, ce pre pace*
 because I not=have.1SG come pe fight but pe peace
 ‘... because I did not come to fight but to make peace...’

We come now to a problem worth mentioning. At least some of the verbs we discussed above have Slavonic origins (Densusianu 1961; Pană Dindelegan 1968, 2014; Todi 2001). It could, therefore, be argued that these verbs are special, as they probably needed a certain amount of time to be completely integrated into the Romanian grammar. DOM, on the other hand, grammaticalized without Slavonic verbs playing any role in languages such as Spanish, Catalan, Italian dialects, etc. Hence, what we observe may be an interesting phenomenon, entirely unrelated to the source of the grammaticalization of Romance DOM.

Indeed, as we readily acknowledge, the data we discuss are fairly late in the development of DOM, even though very early in the recorded history of Romanian. Of course, one possibility is that Romanian DOM was indeed facilitated by argument alternating verbs of the kind discussed above. In this case, the argument alternation we saw above is reminiscent of the initial stages of DOM and, to a certain degree, it seems natural that it later disappeared, or was transformed into an argument alternation with *de(spre)* and *pe*.

Alternatively, of course, it is possible that DOM started in some entirely different way in Romanian. There is no way to tell – given the historical data available – which of the scenarios happened. However, there are at least two aspects to consider. First, we do not see any clear empirical evidence for a solid alternative hypothesis. If DOM began another way, how did it happen? What facilitated the choice of *pe*? If indeed, the process was like that in Spanish, why was it not the preposition *la* (which could mark directions and goals) that become the DO marker? What explains that Romanian DOM is so different from Spanish DOM? Second, the fact that *pe* could mark E-TOPIC in Old Romanian seems clearly established based on our findings. While the argument alternation we are discussing appears to happen more often with verbs with Slavonic origin, E-TOPIC marking itself was in no way limited to Slavonic lexical material. Moreover, *pe* is the DOM marker in both Old and in modern Romanian and E-TOPIC marking is the only obvious usage of *pe* that was lost in modern Romanian. Hence, we have good reason to hypothesize that the reason that it lost its function as an E-TOPIC marker (with minor exceptions) was precisely that *pe* took over the job of a differential object marker. Hence, there must have been some interaction between E-TOPIC marking and DOM. Moreover, the alternations we captured in our theory would still need to be analyzed in the same or a

similar way. Finally, even if – historically – the argument alternations did not initiate the emergence of DOM in Romanian, they certainly may shed light on the conceptual closeness between E-TOPICALITY and DOM in general.

7. OUTLOOK AND CONCLUSIONS

In this final section, we wish to discuss the way in which our analysis fits into the larger picture of the emergence of DOM. Obviously, our analysis in this article is closely connected to theories such as those found in Dalrymple and Nikolaeva (2011) and Iemmolo (2010). Such theories see the marking of topicality as the starting point of DOM in certain languages. These two approaches differ from each other, sharing only a common core idea, while our analysis significantly differs from both in details, while simultaneously being connected to what we believe to be the common core of the two approaches.

Dalrymple and Nikolaeva (2011) generally assume an indexing approach to DOM: According to this approach, DOM arises not because of the need to distinguish subjects and DOs, but rather because in certain cases objects high in animacy and referentiality need to be indexed due to the transitivity features of the entire construction. For them, the notion of *secondary topic* is a crucial motivation for the indexing of DOs. Nikolaeva (2001) defines the notion of a secondary topic as an entity such that the utterance is about the relation between it and the primary (*aboutness-*) topic. In other words, a secondary topic is the second most prominent argument in a proposition, provided that the proposition is about two entities. While diagnosing exactly what is and what is not a secondary topic in a sentence is a tedious exercise, crucially, the notion of secondary topic nicely connects to the hierarchies of animacy and salience, just like primary topics, thus being a natural candidate for a reason to mark differential objects.

Iemmolo (2010) argues that in early Romance data, little evidence can be found that secondary topics are the triggers of DOM. Instead, he sees left-dislocation (see also Escandell-Vidal 2009) as the prime type of example that leads to early DOM. Indeed, in left dislocation constructions one can argue that the grammatical status of the left-dislocated individual needs – at least in some cases – disambiguation. This is especially true in cases in which the left dislocation of the object occurs without the respective clitic pronoun, for example in (63), in which *a* is the differential object marker.

- (63) *A me, non (mi) convince questo.*
 A me neg CL.1SG convince:prs.1SG this
 ‘This does not convince me’ (overheard) (Gallo-Italian)

Since left-dislocation is – in Romance – typically triggered by topicality (mainly in the sense of *aboutness*-topic) Iemmolo also naturally connects it to the scales of referentiality and animacy. Again, topics are generally high on these scales.

In this article, we have suggested that it is a different notion of topicality, the E-TOPIC role, that may have played a crucial role in the source of grammaticalization of DOM in Romanian. Crucially, for Romanian, left-dislocation alone does not

provide a sufficient cue to the source of the grammaticalization of DOM. The reason is that left-dislocation alone stands in no obvious relation to the semantics of *pe*, and clitic left-dislocation constructions appear to have evolved later than DOM, see Hill and Mardale (2019). In comparison, we have provided ample evidence that *pe* could be used to mark E-TOPICS. In fact, not only could *pe* be used to mark E-TOPICS, but we have suggested that argument alternations associated with E-TOPICS may have provided the grammatical system, by reanalysis, with good examples of *pe*-marked THEMES/DOs, that directly or indirectly may have initiated DOM in Romanian. These ideas are connected to the analysis of the double object alternation discussed in Mardale (2015), which also connects topicality and argument alternations as a possible source of the grammaticalization of DOM in Romanian. They are similarly connected to the more recent claim in Hill and Mardale (2019) that even the discourse feature of topicality may have been indexed by *pe* in early Romanian.

E-TOPIC, of course, is essentially a discourse topic of embedded content, hence it connects to the scales discussed above in the very same way as secondary topics and dislocated topics. Moreover, our theory is in no way meant to contradict or replace the suggestions of Dalrymple and Nikolaeva (2011) and Iemmolo (2010). On the contrary, this article provides a particularly clear piece of evidence for the general connection between topicality and DOM in Romance languages, and as suggested in section 5.3, our theory would even be compatible with a topic-based grammaticalization of DOM.

REFERENCES

- Aissen, Judith. 2003. Differential object marking: iconicity vs. economy. *Natural Language and Linguistic Theory* 21(3): 435–483.
- Armstrong, Grant. 2013. Agentive reflexive clitics and transitive SE constructions in Spanish. *Borealis: an International Journal of Hispanic Linguistics* 2(2): 81–128.
- Asher, Nicholas. 2004. Discourse Topic. *Theoretical Linguistics* 30(2–3): 163–201.
- Asher, Nicholas, and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.
- Avram, Larisa, and Rodica Zafiu. 2017. Semantic hierarchies in the diachronic evolution of DOM in Romanian. In *Sintaxa ca mod de a fi. Omagiu Profesoarei Gabriela Pană Dindelegan* [Syntax as a way of being. Tribute to Professor Gabriela Pană Dindelegan], ed. Adina Dragomirescu *et al.*, 29–42. București: Editura Universității din București.
- Beaver, David, and Brady Clark. 2008. *Sense and sensitivity: How focus determines meaning*. Oxford: Blackwell.
- Beaver, David, and Emiel Kraemer. 2001. A Partial Account of Presupposition Projection. *Journal of logic, language and information* 10(2): 147–182.
- Boër, Steven E. 1978. ‘Who’ and ‘whether’: Towards a theory of indirect question clauses. *Linguistics and Philosophy* 2(3): 307–345.
- Bossong, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuen iranischen Sprachen* [Empirical universal research. Differential object marking in the new Iranian languages]. Tübingen: Narr.
- Bossong, Georg. 1991. Differential object marking in Romance and beyond. In *New analyses in Romance linguistics: Selected papers from the XVIII Linguistic Symposium on*

- Romance Languages*, ed. Dieter Wanner and Douglas Kibbee, 143–170. Amsterdam: John Benjamins.
- Carlson, Greg, Rachel Sussman, Natalie Klein, and Michael Tanenhaus. 2006. Weak definite noun phrases. *Proceedings of NELS 36*, ed. Christopher Davis, Amy Rose Deal and Youri Zabbal, 179–196. Amherst: University of Massachusetts.
- Chivu, Gheorghe, Magdalena Georgescu, Magdalena Ioniță, Alexandru Mareș, and Alexandra Roman-Moraru, eds. 1979. *Documente și însemnări românești din secolul al XVI-lea* [Romanian documents and notes from the 16th century]. Bucharest: Editura Academiei.
- Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen. 2013. Inquisitive semantics: A new notion of meaning. *Language and Linguistics Compass* 7(9): 459–476.
- Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen. 2018. *Inquisitive Semantics*. Oxford: Oxford University Press.
- Comrie, Bernard. 1979. Definite and animate direct objects: a natural class. *Linguistica Silesiana* 3: 13–21.
- Cresswell, Max, and Arnim von Stechow. 1982. *De re* belief generalized. *Linguistics and Philosophy* 5(4): 503–535.
- Croft, William. 2003. *Typology and universals*. Second edition. Cambridge: Cambridge University Press.
- Dalrymple, Mary, and Irina Nikolaeva. 2011. *Objects and Information Structure*. Cambridge: Cambridge University Press.
- Densusianu, Ovid. 1961. *Istoria limbii române* [History of the Romanian language]. București: Editura Științifică.
- Dobrovie-Sorin, Carmen. 2007. Article drop and extended heads. In *Pitar Moș: a building with a view. Papers in honour of Alexandra Cornilescu*, (ed.) Gabriela Alboiu, Andrei A. Avram, Larisa Avram and Dana Isac, 99–107. București: Editura Universității din București.
- Dobrovie-Sorin, Carmen, and Ion Giurgea. 2006. The suffixation of definite articles in Balkan Languages. *Revue Roumaine de Linguistique* 51(1): 73–104.
- Dobrovie-Sorin, Carmen. 1994. *The Syntax of Romanian*. Berlin: Mouton de Gruyter.
- Dowty, David Roach. 1991. Thematic proto-roles and argument selection. *Language* 67(3): 547–619.
- Escandell-Vidal, Victoria. 2009. Differential object marking and topicality: the case of Balearic Catalan. *Studies in Language* 33(4), 832–844.
- Fanselow, Gisbert. 2006. On pure syntax (uncontaminated by information structure). In *Form, structure and grammar*, ed. Patrick Brandt and Eric Fuss, 137–157. Berlin: Akademie Verlag.
- Farkas, Donka. 1978. Direct and indirect object reduplication in Romanian. *Papers of the Chicago Linguistic Society* 14: 88–97.
- Givón, Talmy. 1983. *Topic continuity in discourse: a quantitative crosslinguistic study*. Amsterdam: Benjamins.
- Givón, Talmy. 1984. Direct object and dative shifting: a semantic and pragmatic case. In *Objects. Towards a theory of grammatical relations*, ed. F. Pank, 151–182. London: Academic Press.
- Groenendijk, Jeroen, and Martin Stokhof. 1984. Studies on the semantics of questions and the pragmatics of answers. PhD thesis, University of Amsterdam.
- Groenendijk, Jeroen, and Floris Roelofsen. 2009. Inquisitive semantics: Two possibilities for disjunction. In *Meaning, content, and argument: Proceedings of the ILCLI international workshop of semantics, pragmatics, and rhetoric*, ed. Jesus M. Larrazabal and Larraitz Zubeldia. 41–72. Vitoria: University of Basque Country Publication Service.

- Hacquard, Valentine. 2006. *Aspects of modality*. Doctoral dissertation. Massachusetts Institute of Technology.
- Hacquard, Valentine. 2010. On the event relativity of modal auxiliaries. *Natural Language Semantics* 18(1): 79–114.
- von Heusinger, Klaus. 2008. Verbal semantics and the diachronic development of differential object marking in Spanish. *Probus* 20(1): 1–31.
- von Heusinger, Klaus, and Edgar Onea. 2008. Triggering and blocking effects in the diachronic development of DOM in Romanian. *Probus* 20(1): 67–110.
- von Heusinger, Klaus, and Georg A. Kaiser. 2005. The evolution of differential object marking in Spanish. In *Proceeding of the workshop Specificity and the evolution / emergence of nominal determination systems in Romance*, ed. Klaus von Heusinger, Georg A. Kaiser, and Elisabeth Stark, 33–69. Konstanz: Fachbereich Sprachwissenschaft der Universität Konstanz.
- Hill, Virginia. 2013. The direct object marker in Romanian: a historical perspective. *Australian Journal of Linguistics* 33(2): 140–151.
- Hill, Virginia, and Alexandru Mardale. 2017. On the interaction between differential object marking and clitic doubling in Romanian. *Revue Roumaine de Linguistique* 62(4): 393–411.
- Hill, Virginia, and Alexandru Mardale. 2019. Patterns for differential object marking in the history of Romanian. In *Journal of Historical Syntax* 3(5): 1–47.
- Hole, Daniel. 2014. *Dativ, Bindung und Diathese* [Dative, education and diathesis]. Berlin: Mouton de Gruyter.
- de Hoop, Helen and Andrej Malchukov. 2008. Case-marking strategies. *Linguistic Inquiry* 39(4): 565–587.
- Hopper, Paul J., and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2): 251–299.
- Iemmolo, Giorgio. 2010. Topicality and DOM: Evidence from Romance and beyond. *Studies in language* 34(2): 239–272.
- Jordan, Iorgu. 1955. *Ion Neculce, Letopisețul Țării Moldovei* [Chronicle of the country of Moldavia]. București: Editura de Stat.
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.
- Kalin, Laura. 2018. Licensing and differential object marking: The view from Neo-Aramaic. *Syntax* 21(2): 112–159.
- Karimi, Simin. 1990. Obliqueness, specificity, and discourse functions: *Râ* in Persian. *Linguistic Analysis* 20(3–4): 139–191.
- Klein, Udo, and Peter de Swart. 2010. Case and Referential Properties. *Lingua* 122(1): 3–19.
- Koenig, Jean-Pierre, Gail Mauner, and Breton Bienvenue. 2003. Arguments for adjuncts. *Cognition* 89(2): 67–103.
- Kracht, Marcus. 2002. On the semantics of locatives. *Linguistics and Philosophy* 25(2): 157–232.
- Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In *Proceedings of the 8th Semantics and Linguistic Theory Conference*, ed. Devon Strolovitch and Aaron Lawson, 92–110. Cambridge (MA): Massachusetts Institute of Technology.
- Kratzer, Angelika. 2002. The event argument and the semantics of verbs. Ms. University of Massachusetts at Amherst.
- Kuo, Jonathan Cheng-Chuen. 2015. *Argument alternation and argument structure in symmetrical voice languages: A case study of transfer verbs in Amis, Puyuma, and Seediq*. Doctoral dissertation. University of Hawaii.

- van Kuppevelt, Jan. 1995. Discourse structure, topicality and questioning. *Journal of Linguistics* 31(1): 109–147.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- Levin, Beth. 2008. Dative verbs: A crosslinguistic perspective. *Linguisticae Investigationes*, 31 (2): 285–312.
- Levin, Beth. 2015. Semantics and pragmatics of argument alternations. *Annual Review of Linguistics* (1): 63–83.
- Lewis, David. 1988. Relevant implication. *Theoria* 54(3): 161–237.
- Malchukov, Andrej. 2008. Animacy and asymmetries in differential case marking. *Lingua* 118 (2): 203–221.
- Mardale, Alexandru. 2008. Prépositions et article défini en roumain. In *Actes des 21èmes Journées de Linguistique*, ed. Adèle de Saint-Pierre et Mélanie Thibeault, 78–93. Québec : Centre interdisciplinaire de recherche sur les activités langagières, Université de Laval.
- Mardale, Alexandru. 2009. *Les prépositions fonctionnelles du roumain: Études comparatives sur le marquage casuel*. Paris: L'Harmattan.
- Mardale, Alexandru. 2015. Differential object marking in the first original Romanian texts. In *Formal approaches to Old Romanian DP*, ed. Virginia Hill, 200–245. Leiden: Brill.
- Mardale, Alexandru. 2018. Sur l'emploi des prépositions *p(r)e* et *spre* en ancien roumain du 16ème siècle. In *Studii lingvistice. Omagiu Valeriei Guțu Romalo* [Linguistic studies : Tribute to Valeria Guțu Romalo], ed. Gabriela Pană Dindelegan, Rodica Zafiu and Isabela Nedelcu, 217–231. București : Editura Universității din București.
- Mardale, Alexandru, and Edgar Onea. 2017. Prepositional semantics in the grammaticalization of DOM in Old Romanian. Paper presented at the 19th Diachronic Generative Syntax Conference (DiGS 19). Stellenbosch University.
- Montaut, Annie. 2018. The rise of differential object marking in Hindi and related languages. In *Diachrony of differential argument marking*, eds. Ilja Seržant and Alena Witzlack-Makarevich, 281–313. Berlin: Language Science Press.
- Moulton, Keir. 2009. *Natural selection and the syntax of clausal complementation*. Doctoral dissertation. University of Massachusetts at Amherst.
- Næss, Åshild. 2004. *Transitivity: From semantics to structure*. Doctoral dissertation. Radboud University.
- Nikolaeva, Irina. 2001. Secondary topics as a relation in information structure. *Linguistics* 39 (1): 1–49.
- Onea, Edgar, and Daniel Hole. 2017. Differential object marking of human definite direct objects in Romanian. *Revue Roumaine de Linguistique* 62(4): 359–377.
- Onea, Edgar. 2015. Why indefinites can escape scope islands. *Linguistics and Philosophy* 38 (3): 237–267.
- Ostler, Nicholas. 1979. *Case-linking: A theory of case and verb diathesis, applied to Classical Sanskrit*. Doctoral dissertation, Massachusetts Institute of Technology.
- Øvrelid, Lilja. 2005. Animacy classification based on morphosyntactic corpus frequencies: Some experiments with Norwegian nouns. In *Proceedings of the Workshop on Exploring Syntactically Annotated Corpora*, ed. Kiril Simov, Dimitar Kazakov, and Petya Osenova, 24–34. University of Birmingham.
- Pană Dindelegan, Gabriela. 2014. Variație în construcția verbului în româna veche [Variation in the construction of the verb in Old Romanian]. In *Limba română: Diacronie și sincronie în studiul limbii române* [Romanian: Diachrony and synchrony in the study of the

- Romanian language], ed. Rodica Zafiu, Adina Dragomirescu and Alexandru Nicolae, 155–176. Bucharest: Editura Universității din București.
- Pană Dindelegan, Gabriela. 1968. Regimul sintactic al verbelor în limba română veche [The syntactic regime of verbs in Old Romanian]. *Studii și cercetări lingvistice* 19(3): 270–291.
- Pesetsky, David. 1982. *Paths and categories*. Doctoral dissertation, Massachusetts Institute of Technology
- Quine, Willard Van Orman. 1956. Quantifiers and Propositional Attitudes. *Journal of Philosophy* 53(5): 177–187.
- Rappaport-Hovav, Malka, and Beth Levin. 2008. The English dative alternation: The case for verb sensitivity. *Journal of Linguistics* 44(1): 129–167.
- Rawlins, Kyle. 2013. About about. In *Proceedings of Semantics and Linguistic Theory XXIII*, ed. Todd Snider, 336–357. Fort Washington PA: CLC Publications.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27(1): 53–94.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. L. Haegeman, 281–337. Berkeley: Kluwer.
- Roberts, Craige. 2011. Topics. In *Semantics: An international handbook of natural language meaning*, ed. Claudia Maienborn, Klaus von Stechow and Paul Portner, 1908–1934. Berlin: Mouton de Gruyter.
- Seržant, Ilja. 2017. Differential object and subject marking from clefts. Paper presented at the international workshop *Diachrony of DOM*. Paris: INALCO.
- Sinnemäki, Kaius. 2014. A typological perspective on differential object marking. *Linguistics* 52(2): 281–313.
- Stark, Elisabeth. 2011. Fonction et développement du marquage différentiel de l'objet direct en roumain, en comparaison avec l'espagnol péninsulaire. *Mémoires de la Société de Linguistique de Paris 19 (L'évolution grammaticale à travers les langues romanes)*, 35–61. Leuven: Peeters.
- de Swart, Peter. 2003. *The case mirror*. Master's thesis, University of Nijmegen.
- Teng, Stacy Fang-ching. 2008. *A grammar of Puyuma, an Austronesian language of Taiwan*. Canberra: Australian National University Press.
- Todi, Aida. 2001. *Elemente de sintaxă românească veche* [Elements of old Romanian syntax]. Pitești: Editura Paralela 45.
- Zaroukian, Erin. 2013. *Quantification and (un)certainly*. Doctoral dissertation. John Hopkins University.