PREDICATE DOUBLING IN ROMANIAN

Anca Sevcenco[*]

Abstract: The current paper analyzes Romanian predicate doubling, a construction that features topicalization of a non-finite form, a supine, that surfaces either as a bare verb or as a vP complete with arguments and adjuncts and is immediately followed by a clausal structure whose fully inflected tensed verb is the lexical copy of the supine. Predicate doubling occurs in a large variety of languages and has been used in syntactic research to support various theoretical accounts such the multiple copy theory of movement developed in Nunes (2004) or late adjunction of the arguments of the fronted predicate (Landau 2007), to name just a few. I argue for a base generation account of Romanian predicate doubling, drawing upon the framework implemented in Muñoz Pérez & Verdecchia (2022). This framework takes into consideration information structure and the way in which discourse develops by answering relevant questions under discussion.

Keywords: predicate doubling, Romanian, topicalization, movement, base generation

1. Introduction

The current study[1] focuses on discussing the syntactic and discourse-related (information structure) properties of Romanian verbal predicate doubling, illustrated in (1a-b) below, and on providing a theoretical account for this type of construction that rejects a derivation resulting from syntactic movement (based on multiple copy spell-out à la Nunes 2004) and argues instead for a discourse-framed explanation as outlined in Muñoz Pérez & Verdecchia (2022).

Predicate doubling, also known in the literature as predicate clefting, is found in a variety of languages: Spanish (Vicente 2009), Brazilian Portuguese (Bastos Gee 2009), Italian (Maiden & Robustelli 2007), Russian (Abels 2001, Aboh & Dyakonova 2009), Bulgarian (Karagjosova & Jasinskaja 2015), Polish (Bondaruk 2009), Yiddish (Cable 2004) and Hungarian (Ürögdi 2006). It involves topicalization of a non-finite verbal predicate, i.e. a bare verb (1a) or an entire verb phrase (1b), or of an adjectival or nominal predicate (see (1c) and (1d) from Gorăscu 2005: 875). The topicalized verb phrase may be also be complete with arguments also adjuncts.

(1)   a.   De căutat,   am   căutat.
                DE search-SUP have search-PTCP
                'As for searching, I did (search).'
        b.   De căutat   un restaurant bun, am   căutat.
                DE search-SUP a restaurant good have search-PTCP
                'As for searching for a good restaurant, I did (search).'

[*] University of Bucharest, Department of English, anca.sevcenco@g.unibuc.ro.

c. De frumoasă, e frumoasă.
    of beautiful is beautiful
    'As for being beautiful, she is.'
d. De șef, a fost șef toată viața.
    of boss has be.PTCP boss all life-the
    'As for being a boss, he's been a boss all his life.'

The paper focuses on verbal predicates and makes no claim about the pattern involving nominal and adjectival predicates. The topicalized predicate projects at least a vP, but never extends so as to include a tense phrase. In most of the languages that allow this type of verbal predicate topicalization, the topic component is lexically realized as an infinitive verb (see Spanish, Brazilian Portuguese, Russian, Polish, Hungarian and Yiddish). In Romanian, another type of untensed form, the supine, which is morphologically marked by the functional preposition *de*, replaces the infinitive. The fully inflected verb in the clause following the topicalized predicate (henceforth the IP doublet) must be the tensed copy of the topicalized supine. This indicates that predicate doubling imposes rather strict lexical identity conditions on the relevant verbs. In terms of information structure partitioning, the fronted predicate behaves as a topic and the IP doublet as comment. Quite importantly, this topic – comment order cannot be reverted. I illustrate this for Romanian (see also Pană-Dindelegan 2003, Gorăscu 2005):

(2)   *Am căutat, de căutat (un restaurant bun).
       have search-PTCP DE search-SUP (a   restaurant good)
      '*I did search for a good restaurant, as for searching.'

At the discourse level, Muñoz Pérez & Verdecchia (2022) suggest that the predicate doubling structure provides a (possible) answer to an immediate question under discussion (QUD, Roberts 1996) that has previously arisen in the preceding context, which is continously updated. Hence, example (1a) could very likely be integrated in the following conversational exchange between A and B:

(3)   A: Ai căutat?
           have search-PTCP
           'Did you search (for it) ?'
      B: De căutat am căutat (prin tot orașul).
           DE search-SUP have search-PTCP (through all town-the)
           'As for searching, I did search all over the town.'

Alternatively, an informative response the question in (3A) could simply be a confirmation or negation of the searching event: *Da, (am căutat)* 'Yes, I did' or *Nu, n-am căutat* 'No I didn't'. If, however, the speaker chooses to formulate an answer with predicate doubling, the listener expects to hear a continuation that sets the event of searching in contrast with an alternative event (for instance, searching vs. finding). The use of an adversative conjunction formally marks the contrast, as indicated in (4):

(4)   De căutat,   am căutat,   dar nu am   găsit.
        DE search-SUP have search-PTCP but not have find-PTCP
        'As for searching, I did search, but I didn't find anything.'

Predicate doubling has raised a lot of interest among syntacticians. Probably the most intriguing question has to do with why lexical identity has to hold between the verbal forms in the topic and the IP doublet. Another point currently under debate is whether it is possible to come up with a single explanation for both bare verb and phrasal vP doubling (see Antonenko 2018 for the proposal that Russian predicate doubling with bare verbs results from movement whereas phrasal vP doubling is best analyzed as base-generated).

It seems that some languages show enough empirical evidence to support a unifying movement derivation for head and phrasal predicate doubling (see, for instance, Vicente 2009 for Spanish or Abels 2001 for Russian, a.o.). Both A-bar movement and remnant movement have been invoked, but the details are not relevant here. Yet other languages do not align with the movement account, one example being Yiddish. In fact, Cable (2004) notes that Yiddish introduces a complex, even paradoxical, situation because the data seem to endorse both a movement and a base-generation analysis. More specifically, Yiddish predicate doubling shows pervasive island sensitivity, and this fact speaks in favor of movement. But, on the other hand, lexical identity effects are not strict in the language. Cable (2004) refers to the cases of loose identity as genus-species effects. He notes that genus-species effects hold on condition that a constituent in the IP doublet gives more specific information than its related constituent in the topicalized predicate. An illustration is given in (5). In (5a), the direct object of the fully inflected verb, *pike*, is a hyponym to the direct object of the fronted predicate, *fish*. Similarly, in (5b), flying to New York is a specific way in which travelling to America generally speaking can be done.

(5)   a.   ? Essen fish est Maks hekht.   Yiddish
                 eat-INF fish eats Max   pike
               'As for eating fish, Max eats pike.'
        b.   ? Forn   keyn Amerike bun ikh gefloygn keyn nyu-york
                 travel-INF to   Amerike am I   fly-PTCP to   New York
               'As for traveling to America, I have flown to New York.'
                                                              (Cable 2004: 8)

Cable (2004) admits that the sentences in (5) sound quite awkward, but he claims that Yiddish grammar licenses them nevertheless.

Considering this roughly outlined background that I have sketched so far, my goal is to integrate the Romanian data in the larger cross-linguistic picture of predicate doubling and find a suitable account for them. To this end, I begin by describing the properties of Romanian predicate doubling in section 2. First, in subsection 2.1, I lay out the descriptive data by making an inventory of the verb classes that occur in this structure and discussing the few restrictions on verb types that apply. Second, in subsection 2.2, I briefly review a previous analysis of the structure under scrutiny here offered in Pană

Dindelegan (2013). I focus on the status of the topicalized supine phrase (hanging topic versus contrastive topic) and I go over the contexts in which this topic construction occurs, laying stress on the way in which information is structured and integrated in the discourse. When possible, I draw comparisons to other languages in order to better highlight what is language specific and what is universal about the Romanian construction. Section 3 goes on to evaluate whether Romanian predicate doubling results from movement of the topicalized supine phrase to a position in the left periphery. I ultimately argue that an explanation along this line fails to capture the empirical data. Section 4 continues the discussion by introducing Muñoz Pérez & Verdecchia (2022)'s framework and subsequently laying down an account for Romanian coached in this framework. Section 5 concludes.

2. Romanian predicate doubling

2.1 The descriptive data

Pană Dindelegan (2003, 2013) and Gorăscu (2005) give an extensive descriptive presentation of Romanian predicate doubling. Pană Dindelegan (2013) mentions that topicalization of the supine form in predicate doubling characterizes standard Romanian. In contrast, in Aromanian, the topicalized predicate is an infinitive, just as in the other Romance languages that allow this structure, i.e. Spanish, Brazilian Portuguese and Italian. The construction belongs to the informal, colloquial speech register (Zafiu 2013). The supine is uninflected for number, person and gender and preceded by the preposition *de*, which Pană Dindelegan (2003) considers to be a topic marker.

Pană Dindelegan (2013) notes that no restriction holds on the type of verbs that occur in predicate doubling. In (6), I offer some examples involving a wide range of verb classes: unergative, unaccusative, modal, aspectual, state and idiomatic verbs (the examples below are Pană Dindelgan's 2013: 152):

(6)   a.   De lucrat,   am   lucrat   destul.   unergative
               DE work-SUP have work-PTCP enough
               'As for working, I worked enough.'
       b.   De căzut,   a   căzut   de nenumărate ori.   unaccusative
               DE fall-SUP have fall-PTCP of countless   times
               'As for falling, (s)he fell countless times.'
       c.   De putut,   sigur că   am   putut.   modal
               DE can-SUP surely that have can-PTCP
               'As for being able to do it, I sure was.'
       d.   De început,   am   început   de mult.   aspectual
               DE begin-SUP have begin-PTCP of long time
               'As for beginning, I did begin a long while ago.'

e.   De părut   bine, sigur că -mi   pare bine.   state
      DE feel-SUP good surely that CL.DAT.1SG feels good
      'As for feeling happy I surely feel happy.'

f.   De venit   în fire mi-   am venit   idiomatic
      DE come-SUP in sense CL.DAT.1SG have come-PTCP
      din   prima clipă.
      from first   time
      'As for coming to my senses, I did come from the first moment.'

Bondaruk (2009) and Vicente (2009) note that no constraints on the verb types that occur in predicate doubling hold for either Polish or Spanish with one exception, though – the verb 'to be'. Vicente (2009: 166) rates (7) as ungrammatical because the topicalized predicate lacks referential power:

(7)   *Ser,   la puerta fue reparada.   Spanish
        be.INF the door   was fix-PTCP
       'As for being (done something), the door was fixed.'

Bondaruk (2009) presents a more nuanced case for Polish. She concurs that bare verb copula 'be' topicalization causes ungrammaticality (see 8), but topicalization of *be* accompanied by its predicative is perfectly acceptable (9).

(8)   *Być   (to) był sławny, ale już   nie jest.   Polish
        be-INF PRT was famous but no longer not is
       'As for being, he was famous but no longer is.'

(9)   Być   sławny (to) był ale już   nie jest.
       be-INF famous PRT was but no longer not is
       'As for being famous, he was famous, but no longer is.'

As already hinted above, the resistance of copula and predicative *be* to topicalization has been put down to the fact that *be* is informationally light and topics must be referential. In (9), copula *be* becomes part of a referential vP, so the ban on topicalization no longer holds. Romanian behaves a little differently in this respect. Pană Dindelegan (2003: 152) gives examples with topicalized existential and bare copulative *be*:

(10)   De fost,   am fost   destul de des.
       DE be.SUP have be.PTCP quite of often
       'As for having been there, I have been quite often.'

(11)   De fost,   am fost   și eu profesor.
       DE be.SUP have be.PTCP too I   professor
       'As for having been a professor, I have been one, too.'

However, she does point to a restriction on the tense of the inflected verb from the IP doublet, which can be only the *perfect compus* tense, as in (10) and (11). The use of any

other tense than *perfect compus* triggers ungrammaticality, as shown in (12)/(13) in which the verb carries present and imperfect inflection, respectively:

(12)   *De fost,   sunt   răbdător.
          DE be.SUP be.PRS patient
        'As for being patient, I am.'
(13)   *De fost,   eram   răbdător.
          DE be.SUP be-IMPF patient
        'As for being patient, I was.'

(Pană Dindelegan 2003: 152)

Pană Dindelegan (2013) accounts for this restriction by calling upon a suggestion advanced in Manoliu (1993). Manoliu (1993: 110) suggests that the past participle of *be* (i.e. *fost* 'been') that is part of the make-up of the supine phrase is perfective and can function as the topicalized part only in contexts that are temporally marked as [+Past].

At this point, it would be useful to add that Hebrew also allows doubling of existential 'be' ('be' referring to location), on a par with what we are seeing in the Romanian example (10):

(14)   lihyot, Gil haya   be- nyu York (aval rak xaci yom).   Hebrew
        be-INF Gil be-PST in New York (but   only half day)
        'As for being, Gil *was* in New York but only half a day.'

(Landau 2006: 41)

Thus, a look at the distribution of Romanian predicate doubling shows that the construction is mostly without specific contraints, excepting the cases in which the topicalized supine is existential or bare copulative *be*. The following subsection presents Pană Dindelegan's (2013) analysis of predicate doubling.

2.2 A previous analysis on Romanian and the current proposal

Regarding the discourse status of the topicalized supine phrase, Pană Dindelegan (2013: 243) proposes that the supine behaves like a hanging topic, "a hanging theme supine". The main supportive argument comes from the observation that the supine is prosodically and syntactically isolated. Prosodically, the topicalized predicate is followed by an intonational break with falling intonation. Pană Dindelegan (2013) takes the possibility to resume in the IP doublet the non-finite form and whatever arguments and adjuncts go with it (see 15) as an indication that the topicalized supine is (also) syntactically unintegrated.

(15)   De mers   la mare, merg   la mare.
        DE gone-SUP to seaside go-PRS.1SG to seaside
        'As for going to the seaside, I will.'

However, I believe that the the topicalized supine does not behave like a hanging topic and is, in fact, a contrastive topic. First, it is generally acknowledged that hanging topics stay at the root, they cannot be embedded, as shown in (16b) with a nominal topic, *Maldive* 'the Maldives':

(16)   a.   Maldive, acolo aș   petrece vacanța de vară.
                  Maldives there COND spend   vacation of summer
                  'The Maldives, I would spend my summer vacation there.'
        b.   *Am mărturisit   că   Maldive, acolo aș   petrece
                   have confess-PTCP that Maldives there COND spend
                  vacanța de vară.
                  vacation of summer
                  '*I confessed that the Maldives, I would spend my summer vacation there.'

Yet, in stark contrast, predicate doubling may occur in embedded contexts:

(17)   Am jurat   că   de căutat,   am căutat   peste tot.
        have swear-PTCP that DE searched-SUP have search-PTCP through everything
        'I swore that as for searching, I did search everywhere.'

In fact, in this respect they behave just like topics introduced by a preposition and are therefore realized as prepositional phrases. For instance, *în Maldive/*'in the Maldives' in (18), is a topicalized PP that is not a hanging topic and occurs in an embedded context (19):

(18)   În Maldive, acolo aș   petrece vacanța de vară.
        in Maldives there COND spend   vacation of summer
        'In the Maldives, I woud spend my summer vacation.'
(19)   Am mărturisit   că   în Maldive, acolo aș   petrece
        have confess-PTCP that in Maldives there COND spend
        vacanța de vară.
        vacation of summer
        'I confessed that in the Maldives, I would spend my summer vacation.'

Also, the received view holds that hanging topics are not integrated in the sentence in which they occur because they lack any kind of syntactic and morphological marking that connects them to a constituent in the respective sentence (see Fábregas 2016, a.o.). Topicalized supines, on the other hand, come with a distinct marker – the preposition *de*.

One last point that I want to bring to attention is that hanging topics never introduce new information. Topicalized supines, on the other hand, can be used to steer the conversation to a different topic – see example (20) from Pană Dindelegan (2003: 157):

(20)   Am vorbit   de câte   și mai câte.
        have speak-PTCP of how many and more how many
        Dar de mâncat,   ai   mâncat?
        but DE eaten-SUP have eat-PTCP
        'We spoke of many things. But as for eating, have you eaten (anything)?'

In (20), the speaker resumes the bit of conversation (s)he previously had with the interlocutor by saying that they spoke of many things and (s)he then switches to something else, the question about eating. In doing so, (s)he sets up a contrast between talking about a lot of things and not getting to chance to eat anything so far; in this context, eating sounds like a good idea.

Considering these three arguments, I propose that the topicalized supine functions as a contrastive topic in the sense that it introduces a contrast to other entities previously accessible in discourse. Moreover, some sort of continuation of the predicate doubling construction is always expected to be made later on in discourse simply because the comparison intrinsic to the contrast needs to be further fleshed out. This is a general characteristic of predicate doubling structures, Muñoz Pérez & Verdecchia (2022) refer to it as a continuation effect, Bastos-Gee (2009) prefers the term "the but-effect".

At this point, it would also be useful to revisit the status of the preposition *de*. As mentioned in the beginning of subsection 2.1, Pană Dindelegan (2013) believes *de* to be a topic marker that surfaces in any type of predicate topicalization, not just supine topicalization, i.e. with adjectival and nominal predicates as well. This means that she lumps together *de* in (21) and in (22):

(21) De căutat,   am   căutat.
      DE searched-SUP have searched-PTCP
      'As for searching, I did search.'

(22) De priceput,   e priceput.
      of skilful-SUP is skilful
      'As for being skilful, he IS (indeed).'

However, *de* in (21) cannot be just a topic marker. The Romanian supine may take on nominal or verbal morphology. Nominal supines co-occur with the definite article and select genitive-marked arguments. Verbal supines (the type found in predicate doubling) assume a morphological form resembling that of an invariant past participle (i.e. a participle in the default singular, masculine form) and combine with arguments that bear accusative case. The difference between nominal and verbal supines is illustrated in (23) and (24) respectively, from Hill (2002: 496). (23) shows the nominal supine *culesul* 'the gathering', derived by means of the suffix *-s* (*cules*), accompanied by the enclitic definite article *-(u)l* 'the' and selecting the genitive marked argument *porumbului* 'of the maize'; (24) features the verbal supine *de cules* 'DE gathered', preceded by *de* and selecting an accusative argument, *porumbul* 'the maize'.

(23) Culesul   porumbului e din ce în ce mai greu.
      gathering-the maize-GEN is from that in that more difficult
      'The gathering of maize is more and more difficult.'

(24) E din ce în ce mai greu   de cules   porumbul.
      is from that in that more difficult de gather-SUP   maize-the
      'It is more and more difficult to gather the maize.'

As already shown in (24), verbal supines need to be preceded by a prepositional complementizer *de* (Hill 2002, 2013). In some way, *de* is similar to the preposition *a* 'to'

that encodes non-finiteness in infinitival clauses (see also Pană Dindelegan 2005 on the similarity between these two prepositions). Consequently, *de* in the topicalized supine phrases is syntactically a complementizer.

The last aspect that I touch upon in this section concerns an inventory of the contexts in which predicate doubling occurs and a close look at the way information structure is organized, i.e. further details on the realization of topic and the type of focus found in the IP doublet. I will start with discussing the received view in this respect, as reflected in work by Pană Dindelegan (2003, 2013) and in Gorăscu (2005). Afterwards, I will offer what I consider to be a more economic and unifying perspective on the contexts under scrutiny here.

Pană Dindelegan (2003) identifies three discourse situations that feature predicate doubling: in the first two, the topicalized supine resumes previously introduced predicates and thus behaves as a discourse device that maintains topic continuity. What keeps these situations distinct is that in one case the topicalized supine resumes a declarative sentence (25), *să plece* 'to leave', and in the second, an interrogation (26), *te-a durut*? 'did it hurt?' – the resumed sequences are highlighted in bold.

(25) Și- a   umblat **să**   **plece**, el știe   pe un' se   umblă, da'
      and has walk-PTCP SBJV leave he knows where IMPERS walks but
      pân' la ormă de plecat   n- a   plecat.
      until at end DE leave-SUP not has leave-PTCP
      'And he went around, only he knows where (one usually goes around), but, as for leaving, he didn't leave in the end.'
                                        (G. Adameșteanu, in Pană Dindelegan 2003: 155)

(26) A: **Te**   **-a**   **durut?**
           CL.ACC.2SG has hurt-PTCP
           'Did it hurt?'
      B: De durut,   sigur că m-   a   durut,   dar
           DE hurt-SUP sure that CL.ACC.1SG has hurt-PTCP but
           nu ca să -mi   dau   sufletul.
           not that SBJV CL.DAT.1SG give up soul-the
           'As for hurting, it surely did, but it's not that I couldn't live with it.'
                                        (Pană Dindelegan 2003: 156)

In the third context, the topicalized supine behaves differently: it breaks topic continuity by introducing a new topic – see (20) repeated for convenience as (27). But in doing so, it still contributes to setting up an opposition, a contrast between having talked about a whole lot of things, but not getting the chance to eat anything.

(27) Am   vorbit   de câte   și   mai câte,
      have speak-PTCP of how many and more how many
      Dar de mâncat, ai   mâncat?
      but DE eat-SUP have eat-PTCP
      'We spoke of many things. But as for eating, have you eaten (anything)?'
                                        (Pană Dindelegan 2003: 157)

I believe that it is of no notable theoretical consequence whether the predicate doubling construction resumes a declarative or interrogative sentence. What matters, on the other hand, is that in both the topic continuity and topic shift contexts the topicalized supine instantiates a constrative topic that resumes an immediate QUD and the IP doublet offers an answer to that question. This is, in fact, what Romanian has in common with Spanish. Muñoz Pérez & Verdecchia (2022) propose that, in Spanish the topicalized verb/vP in predicate doubling functions as a contrastive topic as defined by Büring (2003). As such, first of all it brings to attention an immediate QUD that has been implicitely (or explicitly) formulated in the preceding chunk of discourse. In addition, the topicalized verb/vP also brings up a set of alternative questions to that QUD. I will offer a more detailed discussion of Muñoz Pérez & Verdecchia (2022)'s framework in section 4.1. For the time being, to conclude this subsection, I would suggest that Romanian predicate doubling behaves similarly to its Spanish counterpart. Before elaborating on this proposal based on a discourse account, I will first determine in section 3 whether Romanian predicate doubling finds a comprehensive explanation in terms of syntactic movement of the topicalized verb/vP to a left-peripheral position.

3. Is it movement?

Pană Dindelegan (2003: 153) observes that Romanian licenses long distance predicate doubling out of complement clauses. The examples below are hers:

(28) a.   De văzut,   pot să   vadă și   ei.
             DE seen-SUP can SBJV see   also they
             'As for seeing, they can see too.'
      b.   De fugit,   nu se   gândea să   fugă.
             DE run-SUP not REFL.CL.3SG think   SBJV run
             'As for running, (s)he didn't think of it.'

Interestingly, she also brings to attention ungrammatical examples of predicate doubling, which are, in fact island violations; (29) illustrates this point with a *wh*-island:

(29)   *De ascuns,   l-   au   găsit   acolo unde s-   a
          DE hide-SUP CL.ACC.3SG.M have find-PTCP there where REFL.CL.3SG has
        ascuns.
        hide-PTCP
        'As for hiding, they found him where he hid.'

                                                              (Pană-Dindelegan 2003: 153)

A deeper investigation into island effects confirms that Romanian predicate doubling is island-sensitive. Examples (30) and (31) show violations of a complex NP and a coordinate structure, respectively, and both are ungrammatical:

(30)   a.   *De angajat, umblă   zvonul   că   angajează.
              DE hire-SUP circulates rumor-the that hire-PRS
              'As for hiring, the rumor that they are hiring has been circulated.'
       b.   *De speriat   turiști, circulă știrea   că
              DE scare-SUP tourists circulate news-the that
              îi   sperie incendiile.
              CL.ACC.3PL.M scare wildfires-the
              'As for scaring tourists, the news that wildfires scare them has been
              circulated.'
(31)   a.   *De rezolvat, s-   a   informat   și   a   rezolvat
              DE solve-SUP REFL.3SG has inform-PTCP and has solve-PTCP
              problema.
              problem-the
              'As for solving (the problem), he looked up information and solved it.'
       b.   *De rezolvat problema, s-   a   informat   din
              DE solve-SUP problem-the REFL.CL.3SG has inform-PTCP from
              cărți și a   rezolvat. -o.
              books and has solve-PTCP CL.ACC.3SG.F
              'As for solving the problem, he looked up information from books and
              solved it.'

The data in (28)-(31) seem to reliably point to the conclusion that Romanian predicate doubling must be the result of some type of movement because it shows sensitivity to islands, i.e the complex NP and wh-islands, and ungrammaticality of extraction from a coordinate structure. But the story just does not end here, as I will try to demonstrate in the ensuing discussion.

The first point that raises concern about a potential movement derivation comes from morphological mismatches effects. Vicente (2009: 171) notes that in Spanish, the topicalized verb in predicate doubling most of the times occurs as a bare infinitive. However, if the fully inflected verb in the IP doublet is in the passive voice, the topicalized predicate must surface as a past participle, not an infinitive, and must agree in number and gender with the subject of the passive sentence – an agreement mismatch explains the ungrammaticality of (32b):

(32)   a.   Reparada,   la puerta ha sido reparada.   Spanish
              fix-PTCP.SG.F the door has been fix-PTCP.SG.F
              'As for being fixed, the door has been fixed.'
       b.   *Reparado,   la puerta ha sido reparada.
              fix-PTCP.SG.F the door   has been fix-PTCP.SG.F

In Romanian, the topicalized predicate will always be a supine irrespective whether the verb in the IP doublet is in the passive or active voice. And the supine surfaces as an invariant, uninflected form, as already mentioned. Consequently, if the fully inflected verb is passive, no agreement holds between the supine and the subject of that passive, unlike in Spanish:

(33)   a.   De reparat,   uşa   a   fost   reparată.
              DE fix-SUP door-the has be.PERF repair-PTCP.SG.F
              'As for being fixed, the door has been fixed.'
       b.   *De reparată,   uşa   a   fost   reparată.
              DE fix-SUP.SG.F door-the has be.PERF repair-PTCP.SG.F
(34)   a.   De găsit,   a   fost   găsită   în cele din urmă.
              DE find-SUP has be.PERF find-PTCP.SG.F in those from last
              'As for being found, it was found eventually.'

                                                              (Gorăscu 2005: 874)

       b.   *De găsită,   a   fost   găsită   în cele din   urmă.
              DE find-SUP.SG.F has be.PERF find-PTCP.SG.F in those from last

This is a first indication that the topicalized supine predicate did not initially originate in a position within the IP doublet. Additional support for this view comes from the presence of genus-species effects, similar to those reported for Yiddish (Cable 2004), Brazilian Portuguese (Cable 2004, Bastos-Gee 2009) and also for Spanish (Muñoz Pérez &Verdecchia 2022). Let us consider (35) and then, for the sake of comparison, (36) from Muñoz Pérez & Verdecchia (2022: 32):

(35)   De mâncat peşte, mănânc doar păstrăv.
        DE eat-SUP fish   eat-PRS only trout
        'As for eating fish, I only eat trout.'
(36)   Leer   libros, leo   solo novelas.   Spanish
        read-INF books read-PRS only novels
        'As for reading books, I (only read novels).'

The narrow focus in (35) goes to the direct object *păstrăv* 'trout' which is in a hyponymy relation with the direct object in the topicalized predicate phrase, *peşte* 'fish'. The same semantic relation holds between *libros* 'books' and *novellas* 'novels' in (36). Muñoz Pérez & Verdecchia (2022) suggest that for genus-species effects to be licensed, the hyponymy constraint must apply. The prerequisite for a hyponymy relation in this case dovetails nicely with Cable (2004)'s earlier observation that the constituent in the IP doublet must provide more specific information than its counterpart in the topicalized predicate.

   Pană Dindelegan (2003) notes that, on occasion, in Romanian, the IP doublet need not necessarily contain a lexical copy of the supine. More specifically, she says that a verbal anaphor could resume the topicalized supine phrase. Invariably, this anaphor is the verb *a face* 'to do' preceded by the invariable unstressed clitic *o* 'it'.

(37)   De redus salarii, o   vor face cu   siguranţă.
        DE cut.SUP wages, CL.ACC.3SG will do with certainty
        'As for cutting down wages, they will surely do it.'

Interestingly, Pană Dindelegan goes on to say that reverting the topic-focus order produces ungrammaticality:

(38) *O vor face cu siguranţă, de redus salarii.
 CL.ACC.3SG will do with certainty DE cut.SUP wages

In principle, the topic-focus order cannot be reversed anyway, but, in (38), the fact that the verbal anaphor precedes its antecedent adds up to the ill-formedness of the utterance. It is worth mentioning that Brazilian Portuguese also has resumption of the topicalized infinitive by the verbal anaphor *fazer isso* 'do it' (Bastos-Gee 2009: 170):

(39) Vacinar cachorro, eu conheço um veterinário que faz isso. Portuguese
 vaccinate-INF dog I know a veterinarian that does it
 'As for vaccinating dogs, I know a veterinarian that does it.'

Resumption by means of a verbal anaphor gives further supporting evidence to the idea that the topicalized supine phrase did not belong in the IP doublet at the onset of the syntactic derivation and hence cannot have moved from there. Additionally, there is a fourth argument provided by the pronominalization patterns observed in predicate doubling constructions. In some cases, a DP constituent occurring in the topicalized supine phrase is resumed by a pronominal clitic form in the IP doublet. Consider, for instance, (40):

(40) De văzut pe fiica preşedintelui,
 DE see-SUP PE daughter-the president-GEN
 am văzut -o.
 have see-PTCP CL.ACC.SG.F
 'As for seeing the president's daughter, I did see her.'

At first blush, it seems that *fiica preşedintelui* 'the president's daughter', the definite, [+human] direct object DP occurring with the supine verb is clitic doubled by the singular feminine accusative clitic *o* 'her' in the IP doublet. In truth, specific [+human] direct objects are differentially object-marked in Romanian by the functional preposition *pe* and also clitic doubled (41b):

(41) a. *Am văzut fiica preşedintelui.
 have see-PTCP daughter-the president-the
 'I saw the president's daughter.'
 b. Am văzut -o pe fiica preşedintelui.
 have see-PTCP CL.ACC.3SG.F pe daughter-the president-the

The only way to derive by movement the structure in (40) would be to assume that the supine phrase initially started out in a position selected by the fully inflected verb (see (42)), possibly a big DP, as Vicente (2009) assumes for Spanish and subsequently moved to the left periphery of the IP doublet:

(42) De văzut pe fiica preşedintelui am văzut -o
 DE see-SUP PE daughter president have see-PTCP CL.ACC.3SG.F

[am   văzut-   o   ~~de văzut   pe fiica   președintelui~~]
have see-PTCP CL.ACC.3SG.F DE see-SUP PE daughter president

But if this is indeed so, we would predict that further extraction of any constituent belonging to the topicalized supine should be prohibited. This is actually the Freezing effect, a ban on extraction out of a phrase that has undergone syntactic movement. Consequently, we would expect (43) to be ungrammatical because the definite DP has undergone further topicalization, i.e. A-bar movement, but it is not:

(43)   Pe fiica   președintelui, de văzut,
        PE daughter-the president-the DE see-SUP
        am   văzut-   o   ieri   la mall.
        have see-PTCP CL.ACC.3SG.F yesterday at mall
        'As for seeing the president's daughter, I saw her yesterday at the mall.'

        Saab (2017) gives an extensive discussion of pronominalization patterns in Rioplatense Spanish vP topicalization (predicate doubling, in fact) and argues that they match extrasentential rather than intrasentential anaphoric relations. This entails that whatever anaphoric elements happen to occur in the IP doublet will find their antecedents in the preceding discourse, more precisely, in the nominal phrases inside the topicalized verb/vP. His observation applies to clitic pronouns as well as to full pronouns and other anaphoric expressions such as epithets. Let us consider the Romanian data:

(44)   Am   citit-   *(o) cartea.
        have read-PTCP it   book-the
        'I read the book.'
(45)   De citit   cartea,   am   citit   -o   ieri.
        DE read-SUP book-the have read-PTCP CL.ACC yesterday
        'As for reading the book, I read it yesterday.'

The definite DP *cartea* 'the book' in (44) cannot be differentially object-marked and clitic doubled by the accusative clitic *o* 'it', it lacks the [+human] specification. This further suggests that the accusative clitic *o* 'it' in (45) is simply a pronominal anaphor that takes the DP *cartea* 'the book' as its extrasentential antecedent. Put differently, the relation between *the book* and *it* in (45) is identical to the relation between these two constituents in (46):

(46)   Am   văzut   cartea.   Am   cumpărat -o   imediat
        have see-PTCP book-the have buy-PTCP CL.ACC immediately
        'I saw the book. I bought it at once.'

The other options of anaphoric nominals with extrasentential antecedents mentioned by Saab for Spanish are also available in Romanian: see (47) for a tonic pronoun and (48) for an epithet example:

(47) De vorbit   cu vecinul,   am vorbit   ieri   cu   el.
      DE speak-SUP with neighbour-the have speak-PTCP yesterday with him
      'As for speaking to the neighbour, I spoke with him yesterday.'
(48) De vorbit   cu vecinul,   am vorbit   ieri   cu
      DE speak-SUP with neighbour-the have speak-PTCP yesterday with
      idiotul   ăla.
      idiot-the that
      'As for speaking to the neighbour, I spoke yesterday with that idiot.'

In (47), the pronoun *el* 'him' inside the prepositional phrase *cu el* 'with him' takes the DP *vecinul* 'the neighbour' from the topicalized supine verb as its extrasentential antecedent. The same relation holds between the epithet in (48), *idiotul ăla* 'that idiot' and the nominal *vecinul* 'the neighbour'.

The present section started out with a question regarding the feasability of a movement derivation for Romanian predicate doubling. Even if long distance doubling is allowed and sensitivity to islands seems to be in place (but see the discussion in section 4 too, in regard to islands), I have shown in this section that there are other pieces of evidence that undermine this view. They have to do with: morphological mismatch effects, genus-species effects, resumption by the verbal anaphor *a o face* 'do it' and the pronominalization pattern in predicate doubling, which has the properties of extrasentential anaphoric relations. In a nutshell, the evidence presented here points to the conclusion that movement is not tenable and that the topicalized supine phrase must be base-generated. The next section outlines an account for Romanian predicate doubling in discourse terms, following Muñoz Pérez & Verdecchia's (2022) explanation for Spanish, and sheds light on the issues of island sensitivity and genus-species effects.

4. It is base generation

4.1 The framework: Muñoz Pérez &Verdecchia (2022)

Muñoz Pérez & Verdecchia (2022) explore predicate doubling (bare verb and vP topicalization) in Spanish, decide against a movement derivation in terms of multiple copy spell-out and settle for a base-generation explanation. But they also took things a step further from assessing only the syntactic facts and set the predicate doubling construction against the discourse backdrop that contains it. This led to a discourse explanation of the structure under scrutiny.

The account makes use of the notion of contrastive topic, as defined in Büring (2003), and of the relation between constrastive topics and focus envisaged in the same work. It also draws upon the Question Under Discussion model of discourse put forth in Roberts (1996), according to which sentences represent answers to explicit or implicit questions that come up as the discourse gradually unfolds. By and large, a piece of discourse addresses what Roberts (1996) calls a Big Question (for instance, *What is the way things are?*), i.e. a question that covers multiple aspects and whose final answer(s) will settle a certain matter that has been of interest. Roberts (1996) distinguishes between

super-questions and sub-questions and proposes that the former entail the latter. Büring (2003) adopts the super-question/sub-question distinction[2] in order to advance a discourse model based on a hierarchical structure represented as a tree diagram, i.e. a discourse/D-tree. To illustrate these theoretical claims in simple, intuitive terms, let us consider an example of a Big Question given by Muñoz Pérez & Verdecchia (2022): *What is Cosmo like?* In order to explore what Cosmo is like, a sub-question, *What does Cosmo like?*, is going to be useful, and so is its answer: *Cosmo likes chicken.* More sub-questions and their answers will come up until the Big Questions gets settled, i.e. we find out what kind of guy Cosmo is. Put in general terms, in this model, the resolution of a piece of discourse arises when all the relevant sub-questions have been exhaused and have received an answer. In Büring's (2003) D-tree account, contrastive topics relate a sentence to a set of alternative questions whereas focus connects sentences to a set of alternative propositions.

Starting from the assumption that discourse chunks are built around relevant set(s) of immediate questions under discussion, Muñoz Pérez & Verdecchia (2022) propose that the topicalized predicate in predicate doubling instantiates a contrastive topic (CT) (defined as in Büring 2003) that calls attention upon an explicit or implicit immediate QUD and feeds the formulation of other sets of questions alternative to the initial QUD. More specifically, as summarized in (49) from Muñoz Pérez & Verdecchia (2022: 1176), their proposal rests on the empirical observations: (i) that the topicalized verb/vP, Predicate 1, functions as a contrastive topic and (ii) Predicate 2 marks narrow focus on a constituent inside the vP or *verum* focus on the main verb and provides answers to the relevant QUD:

(49)   Predicate 1,   [$_{CLAUSE}$ … Predicate 2 … X^0/XP]
         contrastive topic   focus

In technical terms, both the contrastive topic and the focus behave as variables. Discourse resolution arises after the application of the Constrastive Topic Formation algorithm, CT-Formation, taken over from Büring (2003: 519):

(50)   a.   Replace the focus with a wh-word and front the latter; if the focus marks
              the finite verb   or negation, front the finite verb instead.
       b.   Form a set of questions as a result of (49a) by replacing the CT with
              some alternative to it.

For this algorithm to work with contrative topics realized as verbs or vPs as is the case in predicate doubling constructions, a stipulation is necessary (Muñoz Pérez & Verdecchia: 1174):

(51)   If the CT-marked constituent is dislocated outside the clause, replace it with its
         correlate within the clause in order to form the QUD.

---

[2] In fact, Büring (2003) uses the terms question vs. sub-question.

The (strict or sometimes loose) lexical identity between Predicate 1 and Predicate 2 results from an additional stipulation, the Congruence Condition for Predicate Doubling, which requires that the answer in the IP doublet (lexically realized as an assertion) include Predicate 1, the predicate in the topicalized constituent (Muñoz Pérez & Verdecchia: 1180):

(52)   Given a sentence with the structure in (49), there must be a question Q with PREDICATE 1 as its main predicate such that $[\![Q]\!] \subseteq [\![CLAUSE]\!]^{f}$.

Muñoz Pérez & Verdecchia (2022) claim that their Congruence Condition is semantic in nature, not pragmatic.

The next section offers a detailed presentation of how this discourse account of predicate doubling applies to the Romanian data.

4.2 The current proposal at work

By way of exemple, let us see how this proposal works by having a close look at (53), in which the IP doublet contains a DP marked with narrow focus/F, *rezumatul* 'the abstract':

(53)   A: Ce a citit George? Cartea sau rezumatul?
                'What did George read? The book or the abstract?'
        B: De citit,   a   citit   [rezumatul]_F
                DE   read-SUP has   read-PTCP   abstract-the
                (iar cartea   a   lăsat   -o   deoparte).
                 and book-the has leave-PTCP CL.ACC.3SG.F aside
                'As for reading, the did read the abstract (and the book, he put aside).'

By applying the CT-Formation algorithm in (50), the narrow-focused constituent in (53B) gets replaced by a wh-word, *ce* 'what' (54a), and then a set of alternative questions to the QUD in (53A) is formulated (54b) (R marks the verb variable in Muñoz Pérez & Verdecchia 2022's notation).

(54)   a.   [De citit], ce a citit George → Ce [a citit]_CT George?
                 As for reading, what did George read? → What did George read?
        b.   Ce R George? → {Ce a citit George?, Ce a lăsat deoparte George? … }
                What R George → {What did George read?, What did George put aside ? … }

These questions have to be at-issue in the sense of Simons et al. (2010), i.e. a proposition *p* is at-issue if *?p* deals with the QUD. While addressing the question (53A), B leaves room for the alternative relevant QUD *Ce a lăsat George deoparte*? What did George put aside?' with the possible answer *A lăsat deoparte cartea* 'He put aside the book'. This alternative gives rise to the continuation effect that characterizes predicate doubling.

And now, let us see what happens when the focus part gets the *verum* focus (capitalized) interpretation as in (55):

(55)   A: A citit George rezumatul?
                'Did George read the abstract?'
        B: De citit   rezumatul,   l-   a   CITIT.
                DE read-PTCP abstract-the CL.ACC.3SG.N has read-PTCP
                'As for reading the abstract, he DID read it.'

Following CT-Formation, the finite verb gets fronted (56a) and then the predicate in the CT gets replaced by alternatives (56b) – P marks the predicate variable:

(56)   a.   [De citit rezumatul]CT, l-a CITIT ⇸ [A citit rezumatul]CT George?
                'As for reading, he did read it ⇸ Did George read the abstract?'
        b.   P George? ⇸ {A citit rezumatul George?, A răsfoit cartea George? … }
                'Did George P? ⇸ {Did G. read the abstract?, Did G. browse the book? … }

In this case, the predicate doubling structure provides an answer to the QUD: *A citit George rezumatul* 'Did George read the abstract?' Further on, other questions with other verbs that alternate with that signalled by the topicalized predicate will be formulated: *A frunzărit George rezumatul* 'Did George browse the abstract? or *A înțeles George rezumatul* 'Did George understand the abstract?', etc.

The presence of island effects gets an explanation based on the Congruence Condition in (52). In Muñoz Pérez & Verdecchia's (2022) framework the presence of island effects is not to be traced back to syntactic reasons. Instead, they are triggerd by a disruption in discourse. Island structures disobey the Congruence Condition. The assertions embedded in them provide answers to irrelevant, contextually inappropriate QUDs and this leads to a pragmatic crash in discourse. To see how this works out for Romanian, consider the *wh*-island violation in section 2, i.e. example (29), repeated for convenience as (57):

(57)   *De ascuns,   l-   au   găsit   acolo
          DE hide-SUP CL.ACC.3SG.M have find-PTCP there
         unde   s-   a   ascuns.
         where CL.REFL has hide-PTCP
         'As for hiding, they found him where he hid.'

The Congruence Condition requires that (57) answer a QUD that is about hiding and has *hide* as predicate. Hence, the QUD should be something like: *S-a ascuns?* 'Did he hide?' The relevant, at-issue answer is the proposition *p S-a ascuns* 'Yes, he did hide', or *Nu s-a ascuns* 'No, he didn't hide'. Instead, a second, irrelevant QUD cuts in: *'Did they find him?'*, followed by an irrelevant answer *'They found him where he hid'*:

(58)   ⟦S-a ascuns? ⟧ ⋢ ⟦L-au găsit unde s-a ascuns⟧^{f}
        'Did he hide?   They found him where he hid.'

This account makes an interesting prediction. Namely, if the main clause that introduces the island gets a paranthetical interpretation, the ungrammaticality associated with the island disappears. To this extent, Muñoz Pérez & Verdecchia (2022: 1186) discuss the case of main clauses whose verbs function as evidential makers, like *hear*, and which introduce a complex NP island:

(59)   A: ¿Qué compró el vecino?   Spanish
            'What did the neighbor buy?'
       B: Escuché el rumor (de) que compró una Ferrari.
            heard   the rumour of that bought a Ferrari
            'I heard the rumor that he bought a Ferrari.'

The verb *hear* behaves as an evidential marker because it identifies the speaker as the the source that offered the information expressed in the complement clause *he bought a Ferrari*. Example (59) features a case in which well-formedness is not compromised because the main clause *I heard the rumour* functions as a paranthetical element. In this framework, a paranthetical is a structure which does not contain the main point of discourse; in fact, the main point is made in the embedded clause. Put differently, the real at-issue answer in (59) is given in the island-embedded IP doublet.

This type of repair strategy smoothes out the island effects I discussed earlier in regard to Romanian. Consider (60):

(60)   A: Se mai angajează la stat acum? Guvernul reduce drastic cheltuielile?
            'Is the state still hiring now? Is the government cutting down dramatically on expenses?'
       B: De angajat, am auzit zvonul că se angajează.
            DE hire-SUP have hear-PTCP rumour-the that IMPERS hire
            'As for hiring, I heard the rumour that they are hiring.'

In (60B) there is a complex NP island which embeds the at-issue answer to the first QUD, *Se angajează la stat?* 'Is the state still hiring now?'. To me, (60B) does not sound ungrammatical. To futher flesh out the preceding discourse, I also introduced a second QUD – *Guvernul reduce drastic cheltuielile?* 'Is the government cutting down dramatically on expenses?' – which is supposed to get an answer later on in discourse, i.e. *Da, se vor reduce drastic cheltuielile* 'Yes, they will cut down on expenses dramatically', for instance.

The pattern with resumption by the anaphoric verbal expression *a o face* 'do it' finds a trivial account in this framework as well. See (37) repeated for convenience as (61):

(61)   De redus salarii, o   vor face cu siguranță.
       DE cut-SUP wages CL.ACC will do with certainty
       'As for cutting down wages, they will surely do it.'

The QUD that the predicate doubling structure in (61) answers is *Are they going to cut down wages?* The IP doublet *o vor face cu siguranță* 'they will surely do it', makes a

reasonable, at-issue answer to this QUD; the absence of lexical identity between *a reduce* 'to cut down' and *a o face* 'to do it' is not a problem because the invariant anaphoric clitic *o/it* can resume topicalized supine predicates.

It also looks like anaphoric resumption by *a o face* 'to do it' works fine in islands as well. The example (39) provided by Bastos-Gee (2009: 170) for Portuguese, repeated here as (62), clearly indicates this. The immediate QUD should be: *Do they vaccinate dogs?* The main clause, *I know*, functions as an evidential, it asserts the speaker's claim about the reliability of the information source. This allows the main point of discussion to be settled in the embedded relative clause, i.e. the answer to the QUD, *yes, a veterinarian vaccinates dogs* follows in the relative clause.

(62)   Vacinar   cachorro, eu conheço um veterinário que faz isso. Portuguese
        vaccinate-INF dog   I   know   a   veterinarian that does it
        'As for vaccinating dogs, I know a veterinarian that does it.'

The ungrammatical Romanian examples with 'be' topicalization could also be explained along similar lines. Have a look at (12) and (13) repeated here as (63) and (64):

(63)   *De fost,   sunt   răbdător.
         DE be.SUP be-PRS patient
        'As for being patient, I am.'
(64)   *De fost,   eram   răbdător.
         DE be.SUP was-IMPF patient

                                                        (Pană Dindelegan 2003: 152)

The QUD that the topicalized supine in (63) calls attention on is *Were you patient?* and it is anchored in the past. Therefore, an answer with the verb in the present *Sunt răbdător* 'I am patient' is not particulary adequate[3]. Nor is an answer with an imperfective verb *Eram răbdător* 'I used to be patient'. This explanation fits with Manoliu's (1993) insight that the past participle that goes into the morphological make-up of the supine is compatible only with [+Past] contexts. Something similar goes on in (64).

The last point that I want to bring up concerns genus-species effects like (35), repeated as (65):

(65)   De mâncat pește, mănânc doar păstrăv.
        DE eat-SUP fish   eat-PRS only trout
        'As for eating fish, I only eat trout.'

---

[3] One reviewer asked about the other possible QUDs, such as *Eşti răbdător* 'Are you patient?' or *Erai răbdător* 'Did you use to be patient?'. To my mind, in the case of the first QUD, with the verb in the present, an answer like *De fost răbdător, am fost* 'As for having been patient, I was' sounds odd because its reference is set to past. I would opt in this case for an answer involving a topicalized adjectival predicate, i.e. *De răbdător, sunt răbdător* 'As for being patient, I am'. Regarding the second QUD, I would go for the same choice with a topicalized AP: *De răbdător, eram răbdător* 'As for being patient, I used to be'. Further study of this issue will definitely be necessary and useful.

Muñoz Pérez & Verdecchia (2022) suggest that, in this case, the IP doublet offers a coherent answer to the QUD instantiated in the topicalized predicate simply because a hyponymy relationship holds between the internal arguments of the topicalized and the fully inflected predicate, respectively, i.e. *păstrăv/'trout'* has the properties of *fish*, and it brings an additional specification about the type of fish.

To conclude so far, I have shown in this section an alternative account of Romanian predicate doubling, which is not based on syntactic movement, is tenable. Base generation of the topicalized supine verb phrase coupled with sensitivity to discourse contraints can very well account for Romanian predicate doubling. I also raised the issue that island effects may be illusory in Romanian as well, following Muñoz Pérez & Verdecchia's (2022) suggestion (restricted to Spanish) that they are, in fact, caused by disruptions in the coeherence of discourse. Last, I have pointed out that the genus-species effects observed for Romanian also find an explanation in terms of semantic restrictions, i.e. the hyponymy relation.

5. Conclusions

This paper has offered a presentation and discussion of Romanian predicate doubling, a construction that occurs most frequently in the colloquial register of the language and consists of a topicalized supine verb (bare verb, or a verb accompanied by arguments and adjuncts) followed by a fully inflected clause whose verb is lexically identical to the supine form. Even if it might be tempting to use this construction as evidence for the multiple copy theory of movement, as it has been done for predicate doubling in other languages, like Spanish (Vicente 2009), Polish (Bondaruk 2009), and Russian (Abels 2001), to name just a few, the empirical evidence from Romanian speaks against it. First, I reviewed this evidence and established that the presence of morphological mismatches effects, genus-species effects, the possiblity to resume the topicalized non-finite verb by means of the verbal anaphoric expression *a o face/'to do it'* and the pronominalization pattern present in predicate doubling, which characterizes extrasentential rather than intrasentenional anaphoric relationships, clearly point to the absence of syntactic movement, at least in this language. Then, I argued that the Romanian data finds a comprehensive explanation in the approach advocated for in Muñoz Pérez &Verdecchia (2022) for Spanish, according to which the topicalized non-finite form (a contrastive topic) calls attention to an implicit or explicit QUD in discourse and gives rise, in its turn, to another set of other questions that are alternative to the initially posed QUD. The answer to the QUD highlighted by the topicalized predicate is to be found in the fully inflected clause that follows. This approach comes with a nice account for the apparent island effects, explains away the genus-species effects and more adequately fits in with a base-generation analysis for predicate doubling structures. It has the additional advantage that it integrates predicate doubling in the larger context of discourse.

References

Abels, K. 2001. Predicate cleft construction in Russian. In S. Franks & M. Yadroff (eds.), *Proceedings of Formal Approaches to Slavic linguistics 9*, 1-19. Blooomington, IN: Michigan Slavic Publications

Aboh, E. & Dyakonova, M. 2009. Predicate doubling and parallel chains. *Lingua* 119: 1035-1065.

Adameşteanu, G. *Dimineaţă pierdută*. Polirom.

Antonenko, A. 2018. Predicate doubling in Russian: One process or two?. Handout from Formal Approaches to Slavic Linguistics 27 at Stanford University.

Bastos-Gee, A. 2009. Topicalizations of verbal projections in Brazilian Portuguese. In J. Nunes (ed.), *Minimalist Essays on Brazilian Portuguese Syntax*, 161-191. Amsterdam/Philadelphia: John Benjamins.

Bondaruk, A. 2009. Constraints on predicate clefting in Polish. In G. Zybatow, U. Junghanns, D. Lenertová & P. Biskup (eds.), *Proceedings of Formal Descriptions of Slavic Languages 7*, 65-78. Bern: Peter Lang.

Büring, D. 2003. On D-trees, beans and B-accents. *Linguistics and Philosphy* 26: 511-545.

Cable, S. 2004. Predicate clefts and base generation: Evidence from Yiddish and Brazilian Portuguese. Ms, MIT.

Fábregas, A. 2016. Information structure and syntactic manifestation in Spanish. *Borealis: An International Journal of Hispanic Linguistics* 5: 1-109.

Gorăscu, A. 2005. Organizarea informaţională a discursului. In V. Guţu Romalo (coord.), *Gramatica limbii române*, vol. II, *Enunţul*, 864-886. Bucharest: Editura Academiei Române.

Hill, V. 2002. The gray area of supine clauses. *Linguistics* 40 (3): 495-517.

Hill, V. 2013. The emergence of the Romanian supine. *Journal of Historical Linguistics* 3 (2): 230-271.

Karagjosova, E. & Jasinskaja, K. 2015. Predicate clefts in Bulgarian. In G. Zybatow, P. Biskup, M. Guhl, C. Hurtig, O. Mueller-Reichau & M. Yatrebova (eds.), *Slavic Grammar from a Formal Perspective*, 231-245. Bern: Peter Lang.

Maiden, M. & Robustelli, C. 2007. *A Reference Grammar of Modern Italian*. London and New York: Routledge.

Landau, I. 2006. Chain resolution in Hebrew V(P) fronting. *Syntax* 9 (1): 32-66.

Landau, I. 2007. Constraints of partial VP-fronting. *Syntax* 10 (2): 17-164.

Manoliu, M. 1993. *Gramatică, pragmasemantică şi discurs*. Bucharest: Editura Litera.

Muñoz Pérez, C. & Verdecchia, M. 2022. Predicate doubling in Spanish: On how discourse may mimic syntactic movement. *Natural Language and Linguistic Theory* 40: 1159-1200.

Nunes, J. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.

Pană Dindelegan, G. 2003. *Elemente de gramatică. Dificultăţi, controverse, noi interpretări*. Bucharest: Humanitas.

Pană Dindelegan, G. 2005. Supinul. In V. Guţu Romalo (coord.), *Gramatica limbii române I*, 509-524. Bucharest: Editura Academiei Române.

Pană Dindelegan, G. 2013. The supine. In G. Pană Dindelegan (ed.), *The Grammar of Romanian,* 233-244. Oxford: Oxford University Press.

Roberts, C. 1996. Information structure: Towards an integrated formal theory of pragmatics. In J. Yoon & A. Kathol (eds.), *OSUWPL Volume 49: Papers in Semantics*, 35-57. Columbus, OH: Ohio State University Department of Linguistics.

Saab, A. 2017) Varieties of verbal doubling in Romance. *Isogloss. A Journal on Variation in Romance and Iberian Languages* 3 (1): 1-42.

Simons, M., Tonhauser, J., Beaver, D. & Roberts, C. 2010. What projects and why. In N. Li & D. Lutz (eds.), *Proceedings of SALT 20,* 309-327. Ithaca, NY: CLC Publications.

Ürögdi, B. 2006. Predicate fronting and dative case in Hungarian. *Acta Linguistica Hungarica* 53 (3): 291-332.

Vicente, L. 2009. An alternative to remnant movement for partial predicate fronting. *Syntax* 12 (2): 158-191.

Zafiu, R. 2013. Information structure. In G. Pană Dindelegan (ed.), *The Grammar of Romanian,* 568-574. Oxford: Oxford University Press.

A COMMON VIEW ON BROAD AND CONTRASTIVE FOCUS EVENTS

Doina Jitcă[*] and Samuel Maruşcă[**]

Abstract: This paper proposes a cognitive view on sentence stress patterns to discuss focus elements in both broad and narrow focus contexts. The cognitive perspective is based on the hypothesis that prosodic phrases correspond at the cortical level to cognitive binary relations between speech objects of utterances. On this view, cognitive relations are produced by a generic information packaging (IPk) mechanism that pairs constituents with different cognitive functions. At the utterance level, cognitive relations are implemented by prosodic phrases (relations) where different pitch features mark their two functional constituents. Our proposal is to assign sentence stress patterns with corresponding cognitive structural patterns of utterances. One of the two constituents of cognitive and prosodic relations is nuclear and projects its cognitive function to the whole cognitive unit which it belongs to. The paper proposes a nuclear accent analysis by connecting the cognitive functions of constituents with their phonetic/phonological features. The contours analyzed in the paper as hierarchies of cognitive/prosodic relations are selected from those used by Ladd (2008) to exemplify sentence stress patterns in broad focus statements with ascending and descending contours, and in contrastive focus statements. We conclude that, in the new perspective, different cognitive structural patterns can be assigned to contrastive/broad focus statements in different semantic contexts.

Keywords: cognitive relation, prosodic phrase structure, nuclear element, prominence, focus

1. Introduction

The nuclear accent position is an important feature that characterizes the prominence pattern of prosodic phrases. Ladd (2008: 215-216) presents two competing approaches of prominence patterns of contours. The first one is the Normal Stress (NS) view Chomsky (1968), Cinque (1993) and Zubizarretta (1998) and the second one is the Highlighting (H) view, presented in the Bolinger's works; e.g. Bolinger (1965). Referring to the NS view, Ladd concludes that "there is one pattern of prominence that can be specified by rule for every sentence. This pattern assigns a single most prominent stress – primary stress – to one word of the sentence. Normal stress has no meaning or function: it is simply the result of phonological rules on surface syntactic structures". The phonological rules aim to identify the nuclear accent position in order to associate it with the primary stress of utterance. The NS view is applied only to utterances without contrastive focus on their constituents. The following rule is formulated in Bocci et al. (2020) based on Katz & Selkirk (2011): "If the sentence does not contain any occurrence of the [focus] feature, the nuclear pitch accent is assigned to the rightmost element" which has a certain acoustic prominence.

Ladd (2008) analyzes different intonational contour types by using the NS view in order to identify the sentence stress of the related utterance. In the case of the rise-fall-rise contour of *yes-no* question (1b) he concludes that the primary accent is on the word *driving* and the word *instructor* bears the L phrase accent Ladd (2008: 46, 143). This leads to the sentence-initial position of the normal stress corresponding to rising-falling

---

[*] Institute of Computer Science of the Romanian Academy, Iaşi Branch, doina.jitca@iit.academiaromana-is.ro.
[**] Department of Linguistics, SOAS, University of London, maruscasamuel@yahoo.com.

pitch movements. The word *driving* is the first and the last word which bears a pitch accent within the contour, the Normal Stress Rule is satisfied and the nuclear accent is accepted on this word.

(1)   a.   I hear Sue's taking a course to become a driving instructor.
       b.   **A driving** instructor!?

The rise-fall-rise contour is also analyzed in Ladd (2008: 144) within the broad focus statement (2b) with four constituents. An L* pitch accent is applied to the word *dancing* and the H phrase accent is applied to the last word *tonight*. In the NS view the first constituent *I thought* bears the sentence stress because its pitch accent is more prominent then the L* pitch accent on *dancing* having a wide pitch range and the former one could be interpreted the constituent with the primary accent of utterance. The verbal phrase of the subordinate clause *she was dancing* is treated as a post-focal constituent with small variation in pitch range.

(2)   a.   I hear Sue's taking a course to become a driving instructor.
       b.   **I thought** she was dancing tonight.

The present paper proposes a cognitive view that takes into account all accented words as possible candidates for the nuclear position including the word *instructor* in the first case, and *dancing* and *tonight*, in the second. At the cognitive level, all words of utterances are speech items that compete for the nuclear position. We claim that the results of the cortical competition is reflected by utterances at the prosodic level and the nuclear positions can to be deduced by interpreting the prosodic features of constituents.
     In the Highlighting (H) view, a nuclear accent and a sentence stress deduced by a phonological rule is not accepted. Those that support the H view claim that the nuclear position can be assigned on any word which the speaker intends to highlight. In Ladd (2008: 216) this claim is illustrated using Bolinger's work:

> focused words are marked by pitch accent, all pitch accents are individually meaningful, and no one of the pitch accents in an utterance is primary. … In all utterances speakers highlight words according to what they want to say in a specific context.

Ladd illustrates the limitations of this view by using the broad focus statement (3b). The phonological rule identifies sentence stress on the last constituent *francs* bearing an accent with a wide pitch range. Ladd argues that in the case of the word *five* bearing the "information of interest" he has no other rule for identifying the normal stress on the word *five* because the pitch accent with the wide pitch range is also on the last word *francs*. Therefore, he considers the H view cannot deduce in this case the sentence stress on *five* when the pitch accent with significant pitch range is on *francs*.

(3)   a.   What did they give you for participating in the experiment?
       b.   Five **francs**.

Ladd considers that the Focus-to-Accent (FTA) theory "narrows the gap between NS theory and H theory" because the new theory distinguishes between semantic focus and pitch accent and treats the parts of utterances with contrastive focus (CF) involving narrow focus (semantic focus) and the parts of utterance without CF as phrases with normal stress that projects its "focus" function to the whole phrase (broad focus). In Ladd's opinion the "focus" constituent that projects its semantic function to the whole phrase is determined by "language-specific rules or structural principles". In the FTA view the sentence (3b) is analyzed as a narrow focus statement in the particular context of the new information on *five*, even the answer **Five francs** has no contrastive focus meaning. In a narrow focus statement, the sentence stress is on the focus word *five* without engaging the NS rule.

The present paper proposes a cognitive basis for the normal accent rules that can be applied in both broad and narrow focus contexts in order to identify the nuclear constituents of utterances as it results after the speech object representation. The main hypothesis about intonational contours is that they convey the cognitive structure of cortical representations of the respective utterances. Thus, the pitch movements during prosodic words are the result of the F0 frequency modulation by the neuronal output variations of the respective speech constituents during the speech generation process. The cognitive structure of speech object representations is conveyed by the prosodic structure that organizes the corresponding words at the utterance level. In this view, the normal stress position within contours is the consequence of speech object packaging (information packaging) and can be identified by applying a cognitive perspective on F0 contours of utterances.

The cognitive model presented in section 2 involves a cortical mechanism of Information Packaging (IPk) which is responsible for binding speech objects into hierarchically organized cognitive relations marked at the utterance level by prosodic phrases (relations). The cognitive model consists of the functional category definitions including the category of nucleus, and rules for nucleus identification at prosodic phrase level. The cognitive model was also used to explain to explain phrasing and nucleus position in the main Romanian contour types Jitcă (2019) or in contours of English sentences with different information structures Jitcă (2020, 2022).

In section 3 several contours with different focus positions in broad and narrow focus contexts are discussed, in order to identify their cognitive structure differences. Important conclusions summarize the results of this research which lead to a better understanding of information structure of utterances in a direct relationship with their intonational contours.

2. The Information Packaging model

Section 2 summarizes the main aspects of the cognitive model and defines the functional categories of the two structural levels of cognitive relations and their related prosodic phrases (relations). Cognitive relations are binary information units with two contrasted constituents and we named them Cognitive Units (CU). At the cognitive level, utterances are structured by CU hierarchies reflected by prosodic phrase hierarchies, at F0

contour level. The contrast between CU constituents is a functional one and is conveyed by their contrasted tonal features. In the cognitive view, prosodic words are marks applied to words at the utterance level in order to encode the structural information of the corresponding speech objects from the cognitive level. The aim of utterance partitioning consists in identifying the CU hierarchy of utterances.

2.1 Structural levels of cognitive relations

The cognitive model defines the two structural levels of cognitive relations within perceptual object representations by using a set of four functional categories. As the block diagram in Figure 1 presents, the predicate and argument categories defines the first structural level and "theme" and "rheme" categories, the second level.

Regarding predicate-argument structure, Quilty-Dunn (2020) states that "Perceptual Object Representation (PORs) comprise of separate constituents for individuals and properties". He exemplifies this claim by using the sentence *This is a fish* where the constituent *This* corresponds to the individual, *fish* corresponds to the property 'fish', and the syntactic relation between them functions to express the instantiation of fish by the individual. This sort of structure is a canonical example of predicate-argument structure (where *fish* functions as predicate and *This* as argument).

Hurford (2003) considers the predicate-argument structure as "the core of phylo-genetically and ontogenetically primitive (pre-linguistic) mental representations" and claims that "structures of modern natural languages can be mapped onto these primitive representations".
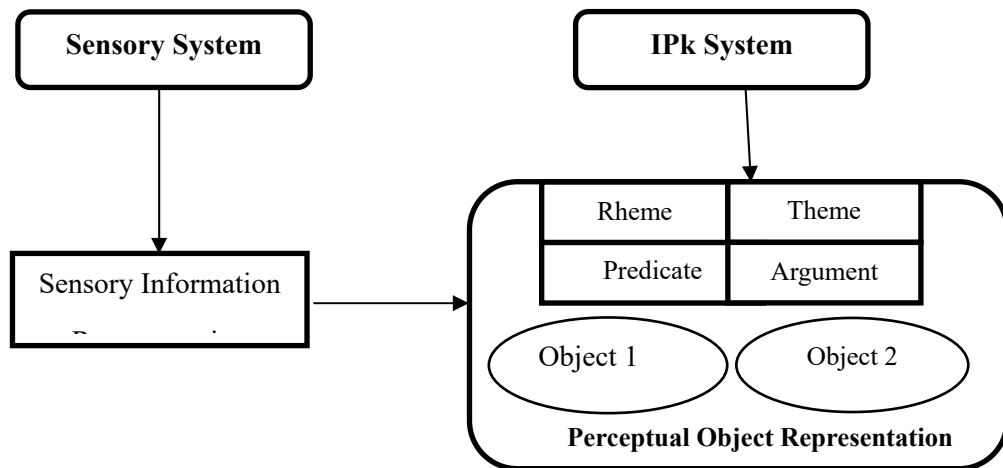
Figure 1. Perceptual Object Representation

At the perceptual object representation level, speech objects become events with a temporal unfolding that are packed into cognitive relation hierarchies. Discussing the common structures of event representations at both perception and memory levels, Zacks (2020) presents the part-subpart structure of two related events as the first structural level

of perceptual representations of events. The part-subpart structure has exactly the same meaning as the argument-predicate structure which describes in Quilty-Dunn (2020) relations within perceptual object representations.

The cognitive model defines CU_argument and CU_predicate functional categories for describing the structure argument-predicate o cognitive relations. CU prefix indicates that the functions have cognitive meaning. At the utterance level, the CU_predicate constituent is marked by the lower target tone, while the CU_argument element is marked by the higher target tone of prosodic phrases (relations).

Gabelentz' model summarized in von Heusinger (2002) presents the argument-predicate structural level in terms of a psychological subject (PS) related to "that about which the hearer should think", and a psychological predicate (PP) related to "what he should think about it". This means that CU_argument is equivalent to the PS category and CU_predicate, to the PP.

In line with Zacks's view presented in Zacks (2020), the part-subpart structure is determined by the spatial delimitation of events in the representational space of their phonetic features and he proposes a second structural level determined by the existence of the temporal delimitation between events on the microbiological time axis. One constituent of cognitive relations is the "cause" event and the other event of the respective relation, is the "effect" event. In this cause-effect view on the order of events, the former event precedes the latter event on the time axis at the neural level. With respect to language, we can speak about the "theme" and " rheme" events, even the semantic theme event does not always precede the semantic rheme event on the time axis of the speech output.

In other words, we can say that the cognitive events related to the speech objects of one utterance are represented in the space of their spatial and temporal features during the construction of their cognitive relation hierarchy or we can say that the temporal evolution of the IPk process of one utterance is encoded in the resulting cognitive representation.

The cognitive model introduces the CU_theme and CU_rheme categories to describe prosodic phrases (relations) at the second structural level. They are marked by different temporal features/shapes of pitch movement during the corresponding prosodic words; e.g. CU_theme is usually marked by slow pitch variation and the CU_rational element is marked by abrupt pitch movements.

The overlapping of the two structural levels is possible at the intonational contour level because the two contrasts are encoded by different acoustic cues of F0 contour involving the two dimensions of pitch variation: tonal target levels and shape or slope of pitch excursion (temporal features). The two structural levels, CU_predicate-CU_argument and CU_theme-CU_rheme, proposed by the cognitive model is a basis for the utterance partitioning description allowing to discuss functional elements in direct relationship to their prosodic features without invoking linguistic aspects.

2.2 Nucleus identification rules

Another aspect of the model refers to the nuclear element of cognitive relations that project its cognitive functions to the whole unit to which it belongs. Based on this

property of nuclear elements, lower-level CUs become functional elements with cognitive functions at the next higher-level relations. In this paper, we suggest the existence of one competition between constituents before their merging into a new unit and, after the nuclear element wins the competition, an exclusion of the non-nuclear elements from the high-gamma activity results but the nuclear one remains in the competition   for higher-level nuclear position.

Two nucleus identification rules are formulated and they correspond to the two types of nuclei: emphasized (prominent) and non-emphasized (non-prominent). Emphasis or prominence of phrases is present on CU_argument constituents when they are marked by high pitch accent with the target tone followed by a falling pitch variation during the last part of the accented syllable and/or on the next non-accented syllable(s). We formulate in (4) and (5) two Nucleus Identification Rules related to the two cognitive relation types: NIR_E, for relations with prominent CU_argument constituent and NIR_NE, for relations without prominent CU_argument constituent, respectively.

(4)   NIR_E: If the CU_argument of cognitive relations is an elementary constituent marked for prominence, then it bears the nuclear function in the current phrase. If an utterance has two constituents marked by prominence, then it is structured by two nested relations where the prominent constituents are the local and global nuclear elements.

(5)   NIR_NE: In cognitive relations without prominence on none of constituents, the CU_predicate bears the nuclear function.

In the present paper, NIR_E and NIR_NE rules are used to deduce local and global nuclei of prosodic phrases (relations) produced by the intonational contours analyzed in section 3.

2.3 The description system of IPk partitions

In view of the IPk model presented in this paper any simple or complex utterance may include a hierarchy of CUs each of them with its own partition. P and A labels were introduced to annotate CU_Predicate and CU_Argument constituents, and T and R labels to annotate CU_Theme and CU_Rheme within IPk partition descriptions. In the proposed description system, two labels are used for annotating one element of partition because it has functions at the two structural levels. Labels are linked by "+" and enclosed in round parentheses.

The description of IPk partitions is a sequence of two pairs of round parentheses separated by slash corresponding to the two CU constituents. In (6) all four possible IPk partition variants for one CU are presented:

(6)   a.   (A+T)/(P+R)
         b.   (A+R)/(P+T)
         c.   (P+T)/(A+R)
         d.   (P+R)/(A+T)

The description of one CU with lower level CU(s) as constituents encloses the description of lower level IPk partitions between brackets and places a functional label in the indices position after the right bracket.

The description of nested CUs needs to identify nuclear contituents at each level of utterance treee in order to deduce the cognitive functions of local CUs at the next higher level. In the cognitive descriptions of the contours analysed in the paper, the nuclear constituent is annotated by n and N labels in the local and global phrases, respectively.

3. Cognitive interpretation of focus in broad and narrow focus contexts

The section presents seven contours extracted extracted from the utterances of the seven sentences also presented in Ladd (2008) in the context of the discussions about their sentence stress patterns. We selected two sentences with rise-fall-rise contours, and five sentences including the syntactic group *five francs* in different semantic contexts. The cognitive descriptions of the contours consist of prosodic phrase (relations) hierarchies and the related nuclear accent hierarchies.

3.1 Nucleus identification in sentences with rise-fall-rise contours

The first sentence with rise-fall-rise contour is the echo *yes-no* question (1b) and it was selected for cognitive analysis in order to demonstrate that NIRs can be applied to both assertive and interrogative sentences. The contour of sentence (1b) is illustrated in Figure 2. The sequence of pitch accents H* L* applies to the contour the CU_argument-CU_predicate structure at the cognitive level. The cognitive structure is described in (7). At the second structural level, the first word is the CU_rheme (abrupt pitch movements) and the second word is the CU_theme (slow pitch movements). They are annotated by A+R and P+T, respectively, in (7).

The nuclear accent of the utterance is identified on the last word by applying the NIR_NE rule, the intonational phrase having a non-emphasized contour. The first constituent does not generate emphasis because its peak does not reach the top level during the accented syllable *dri*. The top level is reached during the last non-accented syllable *ving*. In the cognitive view, the last constituent has the nuclear position having the CU_predicate function. It is labelled by N in (7).

When the number of constituents increases, rise-fall-rise contours are structured by nested prosodic relations as in the case of the contour illustrated in Fig. 3 corresponding to the sentence (2b). At the lowest level, the utterance has two partitions. The first one is that of the main clause *I thought* and the subject *she (was)* of the subordinated clause. They are related as the CU_argument and the CU_predicate constituents. The CU_argument is non-prominent because the high target tone is followed by a tonal fall down to an intermediate high level of the CU_predicate. This marks the latter constituent as the local nuclear element (NIR_NE) that projects its CU_rheme function to the whole left lower-level CU. The cognitive structure of the contour is described in (8).
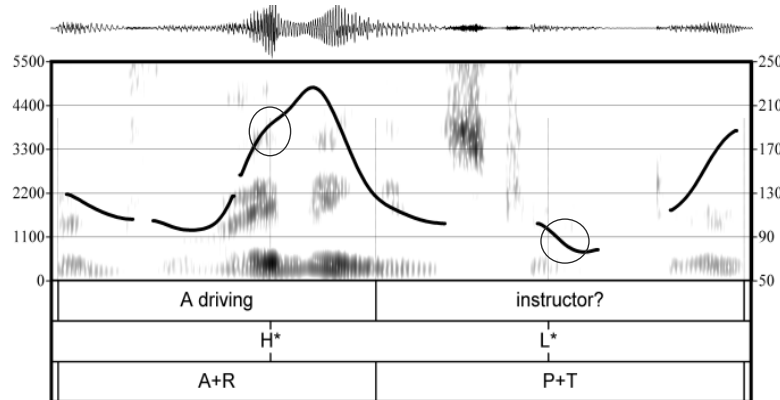
Figure 2. The contour of one utterance of the English echo question *A driving instructor*?

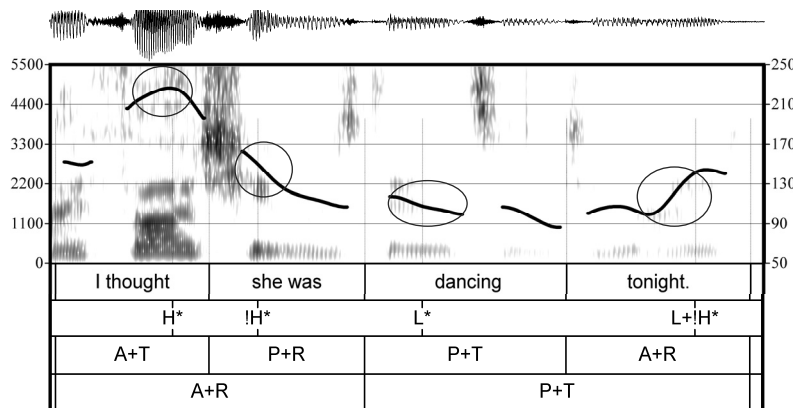(7) A driving^{A+R}/ instructor$^{P+T}_{N}$?

Figure 3. The contour of one utterance of the broad focus statement
*I thought she was **dancing**$_F$ tonight*

(8) [[I thought^{A+T} / she was $^{P+R}$]$_{A+R}$ /[dancing$^{P+T}_{N}$ /tonight^{A+R}]$_{P+T}$

The righ lower-level CU is that of the verb *dancing* and the adverb *tonight* as the CU_predicate and CU_argument element, respectively. The CU_argument *tonight* has no prominence because the contour does not fall after the target tone is reached. Thus, the verb *dancing* bears the nuclear element at this level (NIR_NE).

At the global level, the group *I thought she was* is the CU_argument and CU_rheme constituent which is in contrast with the last group *dancing tonight* with the CU_predicate and CU_theme functions. None of constituents has prominence and the global nuclear element is the global CU_predicate *dancing* (NIR_NE).

At the semantic level, the focus event can be related to the group *she was dancing* because a set of alternatives may be assigned to it. In the cognitive view, we can understand why the first constituent *(I) thought* does not bear the primary accent contrary

to Ladd's interpretation based on the phonological reasons. In all types of sentences the primary accent is carried by the global nuclear element and it results after the utterance partitioning and the nucleus identification on a cognitive basis.

3.2 Utterances with descending phrase-final contours

Ladd (2008) has a special interest in describing different sentence stress patterns of the same syntactic group *five francs* in different semantic contexts. We propose a cognitive description of patterns in the analysed contexts. One of them is that produced by the broad focus statement (3b). Its descending contour is represented in Fig. 4 where the word *five* has a high level pitch movement and the word *francs* has a wide range falling movement. This leads to the CU_argument function of the former element and the CU_predicate function of the latter one. The sentence structure is described in (9). The constant high level marks the word *five* as the CU_rheme element while the slow falling pitch movement during the word *francs* marks it as CU_theme element. The phonological normal stress rule identifies the sentence stress on the last constituent *francs* because the falling pitch movement makes it acoustically prominent and it is in the righmost sentence-position.
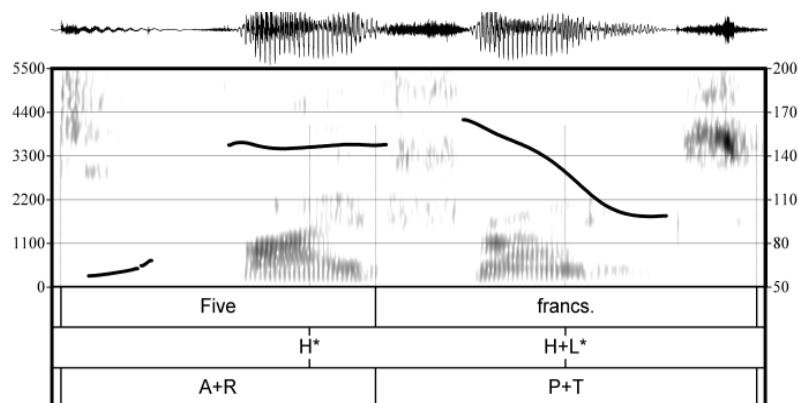
Figure 4. The contour of one utterance of the broad focus statement *Five francs*$_F$

(9) Five^{A+R} / **francs** $_N$$^{P+T}$

The same decision results after applying NIR_NE rule to the analysised contour. The descending contour has no emphasis because the word *five* is non-prominent. The CU_argument *five* reaches a high target tone but the tone is not followed by a falling pitch variation. The CU_predicate constituent *francs* begins with a tonal step up to a little higher level and then the falling pitch variation follows. Based on NIR_NE rule, the last constituent bears the nuclear function having the CU_predicate function.

(Ladd 2008) considers the phonological Normal Stress rule has problems when it has to decide the nuclear position of the numeral *five* in the case it carries new information. In this particular case, the sentence stress must be on *five* despite the fact that

the phonological rule associates the normal stress with the wide pitch range and the righmost position of the word *francs*. There is no phonological rule to associate a word in a non-final-sentence position to the nuclear function within broad focus statements. In the light the FTA theory the prominence that applies the nuclear function to the word *five* can occur only in the contrastive focus context exemplified in Ladd (2008) by the sentence in (10).

The contour of one utterance of the second clause of the sentence (10) is presented in Figure 5 and we use it to describe the sentence stress pattern of the group *five francs* with the primary accent on the word *five*. We consider that the sentence (10) is uttered with a neutral intonation despite the contrastive context suggested by the text. Thus, the sentence in (10) does not generate a narrow focus statement with non-neutral intonation. The cognitive analysis has to be applied in both broad and narrow focus statement interpretations.

Figure 5. The contour of the clause *I gave him **five** francs* in the context of the narrow focus statement *I didn't give him three francs, I gave him **five**$_F$ francs.*

(10)   I didn't give him three francs,
          $[$(I gave him)$_{P+T}$ /(**five** $_N$$^{A+R}$ / francs^{P+T})$_{A+R}$]$_{A+R}$
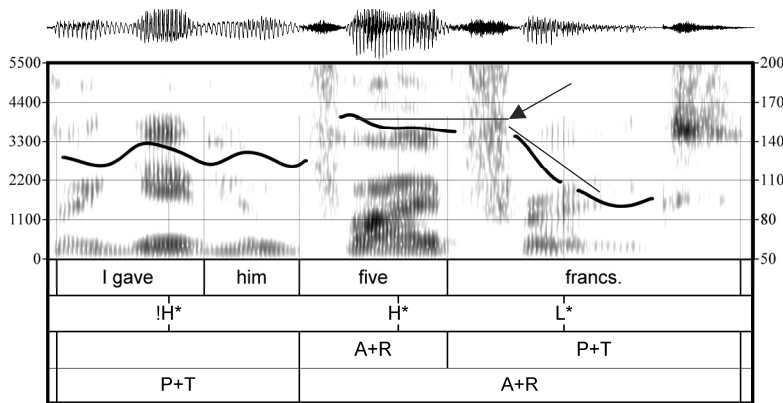
At the global level, the verbal group *I gave him* is the CU_predicate and CU_theme element and the noun phrase *five francs* is the CU_argument and CU_rheme element. At the local level, the word *five* is prominent having the highest tonal target followed by a falling pitch variation during the word *francs*. This leads to the nuclear function of the former constituent with CU_argument function (NIR_E rule).

The group *five francs* represented by the nuclear element *five* is the prominent CU_argument of the intonational phrase that bears the global nuclear function (NIR_E rule). *Five* is annotated by *N* in (10). We claim that the normal stress of the answer *five francs* of the question (3a) is generated in the same manner as the group *five francs* in the case of sentence (10) when it is uttered with neutral intonation. We consider that broad focus statements include all statements that do not introduce a new information element marked by narrow focus. Thus, normal stress can be identified in both cases of statement (3b), with the normal stress on *francs* and on *five*, by using NIR_NE and NIR_E rules, respectively.

In (Ladd 2008), the case of contrastive focus on *francs* is also analysed, in the context of the sentence (11). The second intonational phrase produced by one utterance of the sentence is represented in Figure 6 where it can be observed that the word *five* with the CU_predicate and CU_thematic element (P+T label) is related to the verbal group *I gave him* with the CU_argument and CU_rheme functions (A+R label). The CU_argument constituent is non-prominent because it is followed by a tonal step up to the level where the word *five* begins its falling pitch movement. Thus, the numeral bears the local nuclear function having CU_predicate function (NIR_NE rule).

At the second clause level, the group *I gave him five* is in contrast with the focus word *francs* which is the global CU_argument and CU_rhematic element (A+R label). The latter constituent is prominent, its highest target tone being followed by a falling pitch variation. The word *francs* is marked by emphasis which gives it the global nuclear function at the second clause level. Sentence (11) is uttered as a contrastive focus statement with non-neutral intonation.
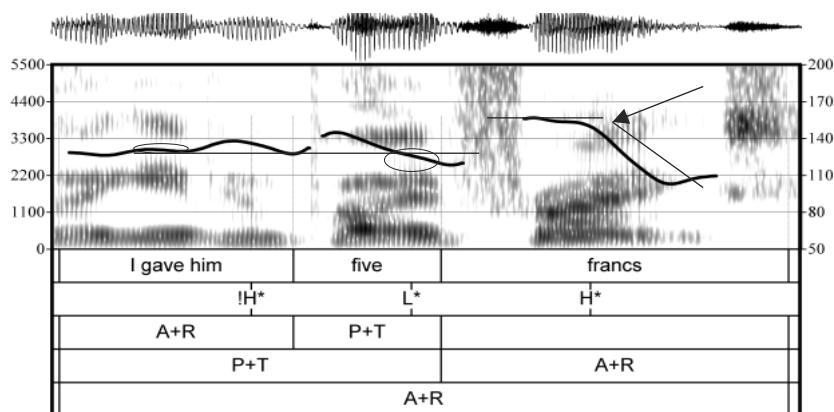

Figure 6. The contour of the clause *I gave him five **francs**$_F$* in the context of the contrastive focus statement *I didn't give him five pounds, I gave him five **francs**$_F$*

(11)   I didn't give him five pounds,
        $[[(\text{I gave him})^{A+R} /\text{five}^{P+T}]_{P+T} /\text{francs }_N{}^{A+R}]_{A+R}$

We conclude that the sentence stress pattern in this contrastive focus context may be characterized as follows. The numeral *five* is grouped to the left with the verbal phrase and it bears the local nuclear element in the resulted group. The noun *francs* is the global CU_argument and nuclear element marked by emphasis (prominence).

Ladd (2008: 214) introduces another sentence stress pattern for the group *five francs* within the double narrow focus statement produced "in relatively unusual circumstances ". It is exemplified by the utterance of the sentence in (12). The second intonational phrase of this utterance is presented in Figure 7 where it can be seen that both focus words *five* and *francs* have CU_argument functions and begin their falling pitch movements at very high tones. The word *five* is the CU_argument element within the embedded phrase where the verbal group *I gave him* is the CU_predicate. Its high

target tone near the top level and the following falling pitch movement produce emphasis at this local level and mark it as the local nuclear function (NIR_E rule). It is labelled by *n* in (12).

At the global level, the word *five* is the CU_predicate because its contour suddenly falls under the top level of the word *francs*. Thus, the latter constituent is a CU_argument marked by prominence because the top-level target tone is followed by falling pitch movement. This explains its global nuclear function (NIR_E rule), and it is labelled by *N* in (12).

The prominence pattern of the clause *I gave him five francs* in the double focus context differs from that of the sentence (11) by the local high prominence of the word *five*. The two prominence patterns have in common the global prominence of the word *francs* that marks it with global focus function.
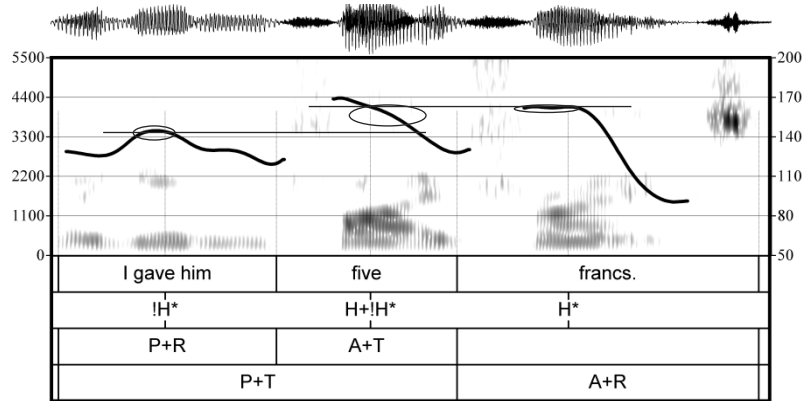

Figure 7. The contour of the clause *I gave him **five francs*** in the context of the narrow focus statement *I didn't give him seven euros, I gave him **five**$_f$**francs**$_F$*

(12)   I didn't give him seven euros,
          [[I gave him $^{P+R}$/ five $_n$ $^{A+T}$]$_{P+T}$/ francs $_N$ $^{A+R}$]$_{A+R}$

The prominence pattern of the sentence (12) in the double focus context differs from that of sentence (11) by the CU_argument and nuclear functions of both words *five* and *francs* that are in agreement with their contrastive focus functions, at the semantic level.

The last prominence pattern of the clause *I gave him five francs* is related to the context of sentence (13) where "the phrase *five francs* is contrasted as a unit" to the word *a dollar* from the first clause Ladd (2008: 214). The group *five francs* taken as a unit, bears new information in the second clause leading to its focus function. Figure 8 represents the contour proposed by us to represent the sentence (13) because in database of Ladd's book the related utterance does not structuraly differ from that represented in Figure 6.

In Figure 8 we can see that the word *five* is grouped to the right with the word *francs* within the low-level CU. In their unit, the former constituent is the CU_argument having the higher target tone and the latter one, is the CU_predicate having the lower

target tone. In this low-level unit, the word *francs* is the nuclear element (NIR_NE). The high target tone of *five* does not generate prominence because it is followed by a tonal step up the a little higher tone and then it falls until the low boundary tone.

At the global level, the group *five francs* is the CU_predicate and CU_theme constituent, and the verbal phrase *I gave him* is the global CU_argument and CU_rheme element. The verbal phrase has no prominence and group *five francs* bears the global nuclear function (NIR_NE).

The sentence in (13) is a contrastive focus statement with new information but it is uttered as a broad focus statement. The last accent of the intonational phrase has a low type and marks *francs* as the global focus word in the same manner as in the case of sentence (9). Thus, the last utterance differs to the utterances represented in Figures 6 and 7 where the word *five* is grouped to the left with the verbal phrase in a lower-level group. In the utterance illustrated in Figure 8 the word *five* is related to the right with the word *francs* under the global focus unit.
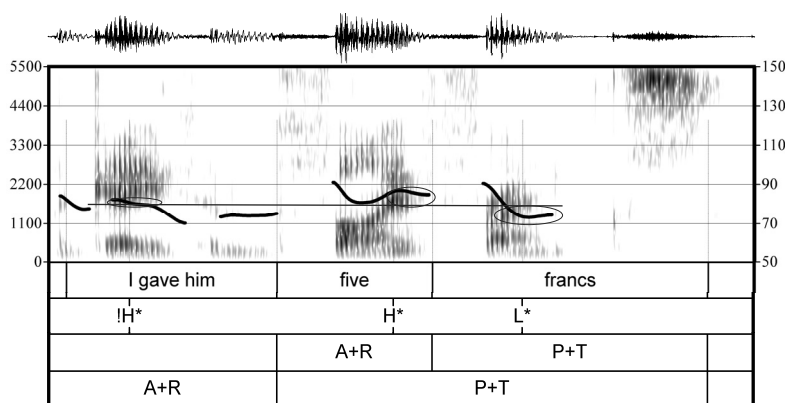

Figure 8. The contour of the clause *I gave him (five francs)*$_F$ proposed in the context of the contrastive focus statement *I didn't give him a dollar, I gave him (five francs)*$_F$.

(13)   I didn't give him a dollar,
         [(I gave him)$_{A+R}$ /(**five** $^{A+R}$/ francs^{P+T}$_N$)$_{P+T}$]$_{P+T}$

In the first part of section 3 one *yes-no* question and two broad focus statements are presented, the first broad focus statements having ascending phrase-final contour and the second one, descending phrase-final contour. In the former case, we explain why the high target tone of the sentence-initial constituent does not produce emphasis and why the global nucleus is deduced by the NIR_NE rule on the global CU_predicate constituent of the utterance. In the latter case, the group *five francs* in the broad focus statement is presented. In the second part of section 3 the four prominence patterns of the clause *I gave him five francs* in the four narrow focus contexts are characterized by using the perspective and the categories of the cognitive model presented in section 2. The intonational phrase of the utterances related to the four contexts has different structural and nuclear patterns, presented in Figures 5-8 and described in (10)-(13). We consider that the words *five* and *francs* act as a unit only in the contour illustrated in Figure 8

corresponding to neutral intonation. Only in this case the two words are merged into a low level relation at the cortical level and their related prosodic words are the two parts of the same peak, at the prosodic level.

4. Conclusions

The paper proposes the cognitive model of information structure as a common basis in analysing prominence patterns of statements in broad focus and narrow focus contexts. The model formulates rules for the nuclear element identification and gives answers to the Ladd's question: "On what basis is a single word selected to bear the main accent". The rules formulated within the model for the nuclear accent identification legitimates nuclear elements to project their cognitive functions to the whole phrase/unit which they belong to.

The prominence of nuclear accents is a result of the competition between neurons which evocate speech items of utterances at the cortical level during their integration as perceptual objects. This neuronal mechanism gives a cognitive meaning to the nuclear element of phrases. Focus events are linguistic (semantic) events implemented by constituents with nuclear functions at the pragmatic level.

The paper proposes the cognitive perspective for the utterance partitioning description. The semantic information structure analysis has to use the cognitive description of utterances and then it may assign semantic functions to constituents. Halliday (1967) proposes the background-focus structure for describing the structure of prosodic phrases (intonation units) and Steedman (2000) also adopted it, but semantic focus constituents correspond to cognitive nuclear elements; this explains why focus category cannot be used in descripting the structure of phrases. Focus (nucleus) labels only signals the sentence/phrase accent and other cognitive categories are involved in phrase structure desriptions (CU_argument, CU_predicate, CU_theme and CU_rheme) because prosodic phrases (relations) have cognitive meaning. After the cognitive description, semantic category labels may be assigned to certain constituents with semantic functions at the information structure level.

References

Bocci, G. Bianchi, V. & Crushina, S. 2020. Focus in *wh*-questions. Evidence from Italian. *Natural Language & Linguistic Theory* 39 (2): 405-455.

Bolinger, D. 1965. *Forms of English: Accent, Morpheme, Order: A Theory of Pitch Accent in English.* Cambridge, MA: Harvard University Press.

Chomsky, N. & Halle, M. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Cinque, G. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24: 239-297.

Halliday, M. A. K.1967. *Intonation and Grammar in British English*. The Hague: Mouton.

von Heusinger, K. 2002. Information structure and the partition of sentence. *Prague Linguistic Circle Papers: Travaux du Cercle Linguistique de Prague Nouvelle série, 4*. viii (376): 275-305.

Hurford, J. R. 2003. The neural basis of predicate-argument structure. *Behavioral and Brain Sciences* 26 (3): 261-282.

Jitcă, D. 2019. An information packaging view on Romanian intonational contours. IEEE Proceedings of the 11th Conference on Speech Technology and Human-Computer Dialogue (SpeD 2021), Timişoara, Romania, October 10-12, 2019.

Jitcă, D. 2020. Information packaging correlates of semantic information structure categories. *Bucharest Working Papers in Linguistics* XXII (1): 5-27.

Jitcă, D. 2022. Cognitive features for the sentence stress patterns description. *Bucharest Working Papers in Linguistics* XXIV (2): 23-40.

Katz, J. & Selkirk, E. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language* 87: 771-816.

Quilty-Dunn, J. 2020. Concepts and predication from perception to cognition. *Philosophical Issues* 30 (1): 273-292.

Ladd, D. R. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press.

Steedman M. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry* 34: 649-689.

Zacks, J. M. 2020. Event perception and memory. *Annual Review of Psychology* 71: 165-191.

Zubizarreta, M. L. 1998. *Focus and Word Order*. MIT Press, Cambridge.