8

Relatedness in Periphrasis:
A Paradigm-based Perspective

GERGANA POPOVA AND ANDREW SPENCER

1. Introduction

THIS PAPER AROSE FROM OUR INTEREST IN PERIPHRASES—complex linguistic expressions that serve as exponents of grammatical features and values and are often integrated in morphological paradigms.[1] Elsewhere, in co-operation with others, we have written in more detail about how such expressions can be defined, recognized, and distinguished from other linguistic phenomena (Brown et al., in press). We will recap some of this below. Our main interest here, however, is in how periphrases within a single grammatical system are related to each other and how their relatedness is best expressed. In particular, we are interested in those cases where it appears that a periphrase is embedded within another complex periphrastic expression.

We think that assuming certain relationships between periphrases (for example, seeing one periphrase as being derived from another, or one being embedded within another) can be linked to the distribution of forms within them. Here, we draw conclusions about relatedness on the basis of the distribution of various forms of an auxiliary verb within the paradigm. Our views of relatedness between periphrases depend to some extent on our views of the phenomenon itself and our views of paradigms. We explain these below.

The data included in this paper come primarily from Bulgarian. Apart from our familiarity with the language, it is a good source of material because of the unusually rich verbal paradigm which combines synthetic and analytic forms and in which many of the analytic forms appear to have similar composition. We will not be able to cover all the constructions that could be considered part of the verbal paradigm in the language. Rather, we have selected a

number of constructions that allow us to illustrate the point we wish to make.

Following the work of others, for example Sadler and Spencer (2001), Ackerman and Stump (2004), Bonami and Samvelian (2009), Bonami and Webelhuth (this volume) we build a case for treating periphrases within a realizational framework and in a fashion similar to the treatment of morphological exponence in Paradigm Function Morphology (Stump, 2001). In particular, we take in the lexical analyses proposed in Bonami and Samvelian (2009) and elaborated further in Bonami and Webelhuth (this volume). These analyses suggest (using as a formalism the framework of HPSG) an account of periphrases which accommodates both the place they have in morphological paradigms, and the variation and complexity of their syntactic behaviour.

In the next section, we sketch our understanding of the phenomenon of periphrasis. We then give a brief overview of the data that will be covered in the paper. Section 4 discusses the relationships between the constructions presented in the data section. Section 5 presents briefly our views on the morphology-syntax interface. The next three sections present the formal frameworks that are used in the account and the account itself.

2. Lexemes, paradigms, and periphrasis

Periphrases are syntactic expressions which bear significance for the morphological system of a language. This insight has led to detailed descriptions of periphrases as a phenomenon which, though not unified, occupies a place between morphology and syntax (see Brown et al. in press and references therein). We will not go here into the level of detail that is present in other discussions. We will only point out some important insights about the nature of periphrases which will be significant for the analyses we sketch later.

Periphrastic constructions are syntactic, in that they comprise at least two elements which have some degree of syntactic independence. Apart from the fact that they are not bound, elements in a periphrastic construction might behave as independent words in other ways, too: for example, they might inflect for certain morphosyntactic feature-values. As such, periphrases can be largely similar to syntactic structures that have no relevance to the morphological system in a given language.

Rather than any syntactic peculiarities, often what sets periphrases apart is their relevance to inflection. For example, in their entirety they can be seen to be equivalent to morphological (typically inflected) word forms. A periphrastic construction might be in a complementary distribution with inflected forms and would differ from them not in terms of lexical meaning but in terms of the morphosyntactic information it is associated with. As a

consequence, a periphrastic construction can be seen to occupy a cell in the paradigm of a lexeme which is otherwise populated with inflected forms. Such views of periphrasis often entail certain views of morphology (for example, where paradigms are determined by morphosyntactic features and their combinations rather than being viewed as simply the collection of inflected forms in a language). Amongst the morphological theories that are based on this understanding of paradigm are realizational theories of morphology, like the one we follow here. Realizational morphological theories assume that inflected forms of lexemes enter into paradigms that are defined by the morphosyntactic features available in a given language. The cells in the paradigm represent a pairing of forms and sets of morphosyntactic features. We illustrate this in (1) with the paradigm of the Latin verb *amāre* 'love' in the present tense:

(1)
	Active		Passive	
	Singular	Plural	Singular	Plural
1	amō	amāmus	amor	amāmur
2	amās	amātis	amāris	amāminī
3	amat	amant	amātur	amantur

In this fragment of the paradigm of the Latin verbal system forms can be said to be associated with sets of morphosyntactic features with particular values; some forms and their respective feature content are shown in (2):

(2) *amō* '(I) love'　　　　　　{Mood: Indicative, Tense: Present, Voice: Active, Person: 1, Number: Singular}

...

amantur '(they) are being loved'　　　　{Mood: Indicative, Tense: Present, Voice: Passive, Person: 3, Number: Plural}

We can use the cross-categorization of these feature-value sets to define cells in the paradigms of lexemes. This will open the possibility that no inflected form is associated with a particular set of features and values, which is nonetheless available in the language. Cells in the paradigm, though defined by the features available in a given language, may be empty if there are no inflected forms to fill them.

Alternatively, we might find that in some cases cells in the paradigms are filled in by syntactic constructions—periphrases (see example in (3) of the perfect tense forms of the Latin *amāre* 'love').

(3)
	Active		Passive	
	Singular	Plural	Singular	Plural
1	amāuī	amāuimus	amātus/a/um sum	amātī/ae/a sumus
2	amāuistī	amāuistis	amātus/a/um es	amātī/ae/a estis
3	amāuit	amāuērunt	amātus/a/um est	amātī/ae/a sunt

In the present tense, as we saw above, Latin verbs have inflected singular and plural forms for both active and passive voice. In the perfect tense, however, Latin verbs have inflected forms only for the active voice. The passive perfect tense cells in the Latin verbal paradigm can be said to be filled in by syntactic structures. These structures, though they comprise more than one syntactically autonomous form, in their entirety carry the same information load as the inflected forms in the rest of the paradigm. This is sketched in (4):

(4) *amāuī* '(I) have been loved'   {Mood: Indicative, Tense: Perfect, Voice: Active, Person: 1, Number: Singular}

     ...

     amātī/ae/a sunt '(they) have been loved' {Mood: Indicative, Tense: Perfect, Voice: Passive, Person: 3, Number: Plural}

In addition to expressing the same information as inflected forms, periphrases have the same distribution as inflected forms. Seeing periphrases as part of a morphological paradigm can have significant empirical advantages, as suggested in Sadler and Spencer (2001). For example, it can explain why certain morphosyntactic content is associated with the construction as a whole even though it is not associated with its parts. Integrating periphrases in the paradigm can also explain why we find in a language the periphrases we find, but not some other plausible ones: for example, ones that will fill in cells that are already filled by synthetic forms.[2]

In the Latin example shown above, we have a situation which is important for the identification of periphrases: we have feature intersection (see Sadler and Spencer 2001 and Ackerman and Stump 2004), in other words, a cross-classification by distinct sets of morphosyntactic properties. Thus, we have a morphologically realized property of active vs. passive voice and a morphologically realized property of present vs. perfect tense/aspect. But the value *passive* of the feature VOICE is expressed synthetically in some cells of the paradigm and periphrastically in others. The number of morphosyntactic features appropriate for a given lexical class in a language is usually larger than one and features often combine freely with each other. The Latin example in (4) shows a combination (intersection) of [VOICE:*active*, *passive*] and [TENSE:*present*, *perfect*] both of which can be expressed morphologically. The presence of morphological synthetic forms for most of the combination of these two features and their values creates the expectation that all of them will be expressed with morphological synthetic forms.

---

[2] See Kiparsky (2005) for alternative views on blocking.

The situation illustrated here with Latin is highly typical of periphrasis, however it is not the only possible one. There are also cases when a certain feature is expressed via both synthesis and periphrasis, but the particular values of that feature are associated either with synthesis, or with periphrasis, not with both. Most of the data we will be looking at in this paper are of this kind (we discuss the data further in Section 3). The same situation can (arguably) be found in French, where within the tense system, imparfait is always synthetic, whereas passé composé is always periphrastic (examples with *aimer* 'love' are given in (5)).

(5)		Imparfait		Passé composé	
		Singular	Plural	Singular	Plural
	1	*j'aimais*	*nous aimions*	*j'ai aimé*	*nous avons aimé*
	2	*tu aimais*	*vous aimiez*	*tu as aimé*	*vous avez aimé*
	3	*il/elle aimait*	*ils/elles aimaient*	*il/elle a aimé*	*ils/elles ont aimé*

Allowing for a value of a certain feature to be consistently associated with periphrasis is, of course, less restrictive than demanding that we see both synthetic and periphrastic forms for the same value. In this chapter, we will argue for a broadening of the understanding of periphrasis in just this way.

What we want to focus on is the relationship between periphrases which appear to have very similar structures. But before we come to our central point, we need to review some of the data we will be using.

3. A brief overview of the data

The data we use to exemplify our points come from Bulgarian, a South Slavic language which has preserved and even increased the complexity of the verbal system seen in Old Church Slavonic.[3] It is an inflecting language, but at the same time many of the features (morphosyntactic or morphosemantic) associated with its verbs are realized via syntactic constructions, rather than via morphological markers like affixes, so Bulgarian is especially rich in periphrases too. In this paper we cover mostly tense and mood constructions, but a full account of the verbal system will have to take into account the fact that each Bulgarian verb belongs to one of two aspects (perfective and imperfective), the systems of voice, reflexivity, etc.

According to traditional accounts, Bulgarian possesses three synthetic tenses and six periphrastic tenses. The full tense system in the indicative is illustrated briefly in Table 1 using the verb *mislja* 'think'.[4] For full descriptions

---

[3] In this respect, Bulgarian together with Macedonian are unique within the Slavic family.
[4] Bulgarian has no infinitive. All citation forms are in 1SG present tense. *Az* is the 1SG pronoun 'I'.

see, for instance Pašov (1966), Tilkov et al. (1983), Scatton (1984), Lindstedt (1985), Popov et al. (1998), Nicolova (2008) amongst others. A Paradigm Function treatment of Bulgarian synthetic morphology is available in Stump (2001).

Table 1: Categories of finite verb inflection

	Form	Gloss
present	*mislja* think.1SG.PRS	'I think'
aorist	*mislix* think.1SG.AOR	'I thought'
imperfect	*mislex* think.1SG.IPRF	'I was thinking'
perfect	(*az*) *săm mislil/la/o* be.1SG.PRS think.PTCP.M/F/N	'I have thought'
past perfect	*bjax mislil/la/o* be.1SG.IPRF think.PTCP.M/F/N	'I had thought'
future	*šte mislja* want.CL think.1SG.PRS	'I will think'
future-in-the-past	*štjax da mislja* want.1SG.IPRF DA think.1SG.PRS	'I would think'
future perfect	*šte săm/băda mislil/la/o* want.CL be.1SG.PRS think.PTCP.M/F/N	'I will have thought'
future-in-the-past perfect	*štjax da săm/băda mislil/la/o* want.1SG.IPRF DA be.1SG.PRS think.PTCP.M/F/N	'I would have thought'

A few comments on the data in Table 1 follow. Bulgarian has three synthetic tenses: a present tense and two past tenses (aorist and imperfect). The compound, or periphrastic, tenses are composed of either the *l*-participle of the main verb or the inflected form of the main verb, and an auxiliary verb. In the latter case both the main verb and the auxiliary agree with the subject. Participles also mark some inflectional distinctions, namely number (singular/plural) and, in the singular only, gender (masculine, feminine, neuter).

The auxiliaries that occur in the series of constructions above are based on the verbs BE and WANT. In the perfect tense, for example, we find the present tense form of the verb BE, whereas in the past perfect tense we see the imperfect BE. Most auxiliaries inflect, though in the future and in the future perfect we see the form *šte*. Historically, this is the 3SG present tense of WANT,

but in the contemporary language it is an invariant form and does not inflect (we gloss it as a clitic). In the future-in-the-past and the future-in-the-past perfect there is another form of WANT, the imperfect tense form. This form still inflects for person and number throughout the paradigm. In most cases the auxiliary and the verb are in a paratactic relationship, but in the future-in-the-past and in the future-in-the-past perfect the periphrase includes the element *da*. The status of *da* is controversial: it has been analysed as a particle-conjunction (Nicolova, 2008, 304), as an auxiliary (by Rudin, 1986 for example), and as a complementizer (see references cited in Rudin, 1986, 57). Some scholars have argued that *da* and the verb together form a lexical sign (Simov and Kolkovska, 2002) whilst identifying three uses of *da*: as a grammatical particle, as a conjunction, and as a modal particle. In Tilkov et al. (1983, 498f) *da* in compound tenses is defined as a grammatical particle (*formoobrazuvašta častica*). The syntactic behaviour of *da* across these different uses is similar, however. We gloss it throughout simply as DA.

As is already suggested by the outline above, periphrases are not associated with a single homogeneous type of syntactic construction — on the contrary, constructions we might want to define as periphrastic may have very different syntactic properties. This point is made particularly forcefully with respect to different languages by Bonami and Webelhuth (this volume) and we will come back to the model they construct to deal with this variability. However, we can still make the general observation that periphrasis (in the inflectional domain) is often associated with functional syntax and bears evidence of the forces of grammaticalization, i.e. one of the elements of the construction typically loses its status as a lexical word and becomes a function word. At the next stage, it loses its syntactic independence and morphologizes. This morphologization is sometimes taken as further evidence that a realizational account of periphrases is the right way to approach them (Bonami and Samvelian, 2009).

The periphrastic constructions we deal with here have undergone morphologization to a different degree. For example, the present tense form of the verb BE is a clitic. The imperfect tense form of BE, however, has not become a clitic. Apart from some of the auxiliaries, there are a number of other verbal clitics in the language. These interact with the auxiliaries in the clitic cluster. One of them is the negative particle *ne* 'not', the addition of which is the most productive way of negating in the language. In the paragraphs that follow we describe the syntactic behaviour associated with some of the constructions outlined above. Our focus will be on the future, perfect, and future perfect constructions, as well as on some of their negated forms. In the following observations, we lay the foundations for a more formal account in Sections 7 and 8.

The verb *săm* 'be' is the only fully inflected auxiliary clitic in the language (and its behaviour is the same regardless of whether it is used as an auxiliary in a compound tense or as a copula, see Franks and King, 2000, 49–51). As a clitic, the verb 'be' is not allowed to appear clause-initially (6a) and needs a host to the left (6b, c). The negative particle *ne* and the future tense particle *šte*, when present, precede the auxiliary 'be' (both these particles can also host the clitic cluster) (see 6d, e). On the other hand, the clitic verb 'be' precedes the dative and accusative clitic pronouns (6f) — unless we are dealing with the 3SG form of the clitic auxiliary, in which case the auxiliary comes at the end of the cluster (6g).

(6) a. * *Săm*   *dala*   *statiite*   *na*   *studenta.*
be.PRS.1SG   give.PTCP.F   papers.DEF   to   student.DEF
'(I) have given the papers to the student.'

   b. *Az săm*   *dala*   *statiite*   *na*   *studenta.*
I be.PRS.1SG   give.PTCP.F   papers.DEF   to   student.DEF
'I have given the papers to the student.'

   c. *Dala*   *săm*   *statiite*   *na*   *studenta.*
give.PTCP.F   be.PRS.1SG   papers.DEF   to   student.DEF
'(I) have given the papers to the student.'

   d. *Az ne*   *săm*   *dala*   *statiite*   *na*   *studenta.*
I CL.NEG   be.PRS.1SG   give.PTCP.F   papers.DEF   to   student.DEF
'I haven't given the papers to the student.'

   e. *Az ne*   *šte*   *săm*   *dala*   *statiite*   *na studenta.*
I CL.NEG CL.FUT be.PRS.1SG   give.PTCP.F papers.DEF to student.DEF
'I would not have given the papers to the student.'

   f. *Dala*   *săm*   *mu*   *gi.*
give.PTCP.F   be.PRS.1SG   CL.3SG   CL.3PL
'(I) have given them to him'.

   g. *Dala*   *mu*   *gi*   *e.*
give.PTCP.F   CL.3SG   CL.3PL   be.PRS.3SG
'(She) has given them to him'.

The morphologization of the auxiliary in the future tense construction is of a somewhat different nature. By way of a reminder, this construction comprises an inflected present tense form of the lexical verb and an invariant clitic particle *šte*, which historically arose from the 3SG present tense form of the verb *šta* 'want'.

In some respects *šte* appears to be more morphologized than *săm* — it does not inflect. It is a proclitic and is normally positioned on the left of the verb (7a). There are severe restrictions on what can intervene between the verb and *šte*. Not even adverbs like *veče* 'already' are particularly welcome in this position (7b). Other clitics, however, can intervene between *šte* and the verb, as long as the obey the restrictions on the clitic cluster (see (7c) — this is

elaborated on below). Unlike *săm*, however, *šte* can appear clause initially (7d) and cancels the need of other clitics for a host to the left.[5]

(7) a.   *Az   šte   dam   kartinite   na   studenta.*
          I   want.CL   give.1SG   paintings.DEF   to   student.DEF
          'I will give the paintings to the student.'

     b.   * *Utre   šte   veče   xodja   na   učilište*
          tomorrow   want.CL   already   walk.1SG   to   school
          '(I) will go to school tomorrow already.'

     c.   *Az   šte   mu   gi   dam.*
          I   want.CL   CL.3SG   CL.3PL   give.PRS.1SG
          'I will give them to him.'

     d.   *Šte   mu   gi   dam.*
          want.CL   CL.3SG   CL.3PL   give.PRS.1SG
          '(I) will give them to him.'

Both auxiliaries (i.e. *săm* and *šte*) enter the clitic cluster. This is a verb-adjacent cluster of auxiliaries and pronominals and the elements in the cluster are ordered according to the template shown in (8) (for more details see Avgustinova 1994; Franks and King 2000; Bošković 2001; Nicolova 2008 amongst others):

(8)   Bulgarian clitic cluster *Neg → Fut → Aux → Dat → Acc → 3sgPresAux*

Clitics in the cluster appear pre-verbally, if possible, and adjacent to the verb. They appear to the right of the verb if there is no available host to the left and if none of the clitics that can appear clause-initially are present (*ne* and *šte*, for example, can appear clause-initially).

The future tense has especially interesting properties with respect to negation. The default means of expressing negation in Bulgarian is by the addition of the negative proclitic *ne* which has already been mentioned briefly, because it, too, enters the clitic cluster. We can illustrate this default negation with the perfect tense, whose negated forms are given in (9) below:

(9)   Singular   Plural
     1   *ne săm mislil/a/o*   *ne sme mislili*
     2   *ne si mislil/a/o*   *ne ste mislili*
     3   *ne e mislil/a/o*   *ne sa mislili*

Forms with *ne* are, in principle, available for the future tense. They are felt to be old-fashioned and are less frequent, however. Much more commonly the

---

[5] As noted in Franks and King (2000, 61f), clusters like *šte mu gi*, even though they do not need further hosts, do not form a prosodic domain and cannot be separated from the verb.

negated future is realized by a special fused negated form of the verb *imam* 'have', i.e. the form *njama* (both sets of forms are illustrated in (10) below).[6]

(10)	Negated Future with *ne* (rare)	Negated Future with *njama*

	Singular	Plural	Singular	Plural
1	*ne šte mislja*	*ne šte mislim*	*njama da mislja*	*njama da mislim*
2	*ne šte misliš*	*ne šte mislite*	*njama da misliš*	*njama da mislite*
3	*ne šte misli*	*ne šte misljat*	*njama da misli*	*njama da misljat*

The negated forms of the future tense are of interest because they involve what looks like a subordinate clause as part of a periphrase. They also show that periphrases share many properties with other syntactic constructions in a given language. We will illustrate both of these points.

Many Bulgarian non-auxiliary verbs can take as complements subordinate clauses headed by the particle *da*. In many respects, the syntactic behaviour of an auxiliary verb taking a *da*-clause is the same as the syntactic behaviour of a non-auxiliary taking a *da*-clause. In all cases, the verb in the subordinate clause inflects as a finite verb. The matrix verb and the subordinate verb can share the same subject and indeed this is obligatory for auxiliaries. The agreement facts in such control constructions can be different for lexical matrix verbs and an auxiliary matrix verb like *njama*, as the auxiliary verb has become an invariant form and no longer agrees with its subject. It is also important to note that Bulgarian is a pro-drop language and subjects can be left unexpressed—in fact, in the subordinate clauses illustrated here, this is the usual pattern, as in (11) below.

(11)	a.	*(Az)*	*iskam*	*da xodja*	*na*	*učilište.*
		I	want.1SG	DA walk.1SG	to	school
		'(I) want to go to school.'

	b.	*(Az)*	*njama*	*da xodja*	*na*	*učilište.*
		I	have.not	DA walk.1SG	to	school
		'(I) will not go to school.'

Da-clauses have to obey the following restriction: only present tense verbs can appear in the *da* construction, whatever the tense of the matrix verb, see (12) for one example:

(12)	a.	*(Az) iskax*	*da xodja*	*na učilište i*	*xodex*	*vseki den.*
		(I)	want.PST	DA walk.1SG.PRS	to school and	walk.PST	every day
		'I wanted to go to school and I went every day.'

	b.	**(Az)	iskax*	*da xodex*	*na učilište i*	*xodex*	*vseki den.*
		(I)	want.PST	DA walk.1SG.PST	to school and	walk.PST	every day
		(intended) 'I wanted to go to school and went every day.'

---

[6] The future negative of *imam* 'have' itself is *njama da imam* or *ne šte imam*, see (Nicolova 2008: 304–5).

When a non-auxiliary verb takes a *da*-subordinate clause, the verbs in the main and subordinate clauses need not share an argument (as in 13a, b). Lexical verbs taking *da*-clauses also allow independent adverbial modification (see 13g). In sentences with the auxiliary *njama*, however, only one predication is possible and the auxiliary and the subordinate verb cannot have two different subjects (see 13c, d, e, f).

(13) a. *Iskam   da   xodiš   na   učilište.*
          want.1SG   DA   walk.2SG   to   school
          '(I) want (you) to go to school.'

     b. *Iskam   Ivan   da   xodi   na   učilište.*
          want.1SG   Ivan   DA   walk.3SG   to   school
          '(I) want Ivan to go to school.'

     c. *Njama   da   xodiš   na   učilište.*
          have.not   DA   walk.2SG   to   school
          '(You) will not go to school.'

     d. *Njama   da   xodi   na   učilište.*
          have.not   DA   walk.3SG   to   school
          '(He) will not go to school.'

     e. **Az njama   da   xodiš   na   učilište.*
          I have.not   DA   walk.2SG   to   school
          'I will (you) not go to school.'

     f. **Az njama   da   xodi   na   učilište.*
          have.not   DA   walk.3SG   to   school
          'I will (he) not go to school.'

     g. *Otdavna   iskam   da xodja   na   učilište   peša.*
          Long   want.1SG   DA go.1SG   to   school   on-foot
          'I have been wanting to go to school on foot for a long time.'

Syntactic material can be inserted between a lexical verb taking a *da*-clause and the *da*-clause itself quite easily. *Njama* tends to stay close to the subordinate verb, though it is possible for syntactic material to come between *njama* and the *da*-clause, as in (14a, b), adapted from Nicolova (2008: 305).

(14) a. *Njama   neprekăsnato   az   da   xodja   za   xljab.*
          have-not   incessantly   I   DA   go.1SG   for   bread
          'I will not be the one to go and buy bread all the time.'

     b. *Iskam   kolkoto   se   može   po-često   az da xodja   za xljab.*
          want.1SG   as-much   REFL   can   more-frequent   I   DA go.1SG for bread
          'I want to be the one to go and buy bread as often as possible.'

The restrictions on *njama* + *da*-clause sentences illustrated above are only to be expected for auxiliary verb constructions. It is important to mention,

however, that negated future *njama*-constructions are not unique in terms of their syntactic behaviour and that syntactic peculiarities are not confined to constructions we might want to call periphrastic. This point will be taken up again below.

To sum up the observations above, as can be expected, in the future tense construction the verb *njamam* 'not-have' has undergone grammaticalization: it does not inflect to agree with the subject and it prefers to stay close to the subordinate verb. It does not introduce an event into the semantic structure of the clause. It has not fused with the *da*-clause verb completely, however, and in some cases can be separated from it by other (substantial) syntactic material. In this respect, it resembles the behaviour of full lexical verbs taking *da*-clauses as complements.

The exceptional behaviour is a sign that *njama* (or rather, the construction it is part of) has acquired grammatical status. *Njama* itself contributes grammatical information when it appears (the future tense property). However, this exceptional behaviour is not unique to this construction. Indeed it is not unique to constructions we might wish to define as periphrastic. We briefly elaborate by giving an example of a construction which shares the restricted behaviour of a *njama da*-clause and is indeed a counterpart to it in many ways, but which we (and descriptive grammars) would not wish to call periphrastic.

The construction we have in mind here is one where the verb taking a *da*-clause is the verb *imam* 'have' (in other words, the non-negated counterpart of *njamam* 'not have'). In this construction, the verb *imam* also bears some of the hallmarks of a transition from lexical to grammatical status: it does not inflect to agree with the subject and does not introduce an event in the semantic interpretation of the clause. On the other hand, it does introduce a subtle modal nuance of coercion or effort, which we gloss as 'a lot' in the examples in (15) below. Traditional grammars often include this construction in discussions of compound tenses but recognise that it has not reached the status of a tense (i.e. has not fully grammaticalized). Unlike the future and the negated future tenses, this construction has not generalized across the whole verbal lexical class and has narrower semantics. For these reasons, the construction cannot be said to realize a functional or inflectional property and therefore it cannot be regarded as a periphrase. The construction is illustrated in (15) below:

(15) a. *Ima   da   xodja   za   xljab   dogodina.*
          have.3SG   DA   go.1SG   for   bread   next-year
          'I will have to go buy bread a lot next year.'

      b. *Deteto   ima   da   xodi   na   učilište   kato   svărši   vakancijata.*
          child.DEF   have.3SG DA   go.3SG   to   school   when   finish.3SG   holiday
          'The child will do a lot of going to school once the holiday is over.'

Given these similarities between periphrases and other constructions in the language, an account is needed that will allow a fairly flexible interface between morphology and syntax. We borrow ideas about what such an interface might look like from Bonami and Webelhuth (this volume). Our aim in this paper, however, is to look at the relationships between periphrastic constructions themselves. For example, the future-in-the-past construction illustrated in Table 1 is very similar to the future construction: the difference is that in the future-in-the-past the auxiliary appears in a different tense. There are also strong similarities between the future, the perfect, and the future perfect. It is such relationships that we want to discuss in the next section.

4. Relatedness in periphrasis

4.1. Periphrases as inflected constructions

A number of the forms listed in Section 3 appear to be constructions derived by 'inflecting' some element of another practically identical construction. Here are some examples: the past perfect is the same as the perfect except that the auxiliary verb *săm* 'be' is not in the present, but is in the imperfect tense. The paradigms of *mislja* 'think' are shown in (16) below.

(16)

	Perfect		Past perfect	
	Singular	Plural	Singular	Plural
1	*(az) săm mislil/a/o*	*(nie)sme mislili*	*bjax mislil/a/o*	*bjaxme mislili*
2	*(ti) si mislil/a/o*	*(vie) ste mislili*	*beše mislil/a/o*	*bjaxte mislili*
3	*(toj/tja/to) e mislil/a/o*	*(te) sa mislili*	*beše mislil/a/o*	*bjaxa mislili*

It is tempting to provide an analysis where the tense of the construction as a whole is linked to the tense of the auxiliary verb, for example along the lines of the analysis sketched in (17) below:

(17)   Aspect $\boxed{1}$
               Tense $\boxed{2}$

            BE   L-PART
         Tense $\boxed{2}$
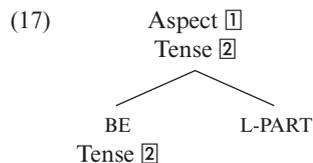
Whatever the semantic import here, an analysis along these lines will have to contend with a significant difficulty: the perfect tense construction and the past perfect construction exhibit different syntactic behaviour. Whereas the present tense auxiliary is a clitic and enters into the clitic cluster, the imperfect auxiliary is not a clitic and, indeed, can act as a host to the cluster.

There is a similar parallel between the forms of the future and the forms of the future-in-the-past. Whereas in the future, as mentioned already, the invariant particle *šte* is historically a present tense form of the verb *šta* 'want', in the future-in-the-past the verb *šta* appears in the imperfect tense. The imperfect form does not enter the clitic cluster in the same way as the present form. It conjugates for person and number and it requires the verb to be marked by *da*. The paradigm of the future-in-the-past for *mislja* 'think' is in (18) below:

(18) | | Singular | Plural |
|---|---|---|
| 1 | *štjax da mislja* | *štjaxme da mislim* |
| 2 | *šteše da misliš* | *štjaxte da mislite* |
| 3 | *šteše da misli* | *štjaxa da misljat* |

There is another aspect of these constructions that has to be taken into account: the fact that it is the auxiliary and not the lexical verb that carries the imperfect tense is, to some extent, arbitrary. We think the existence of dialectal variation is evidence of the rather stipulative nature of such inflectional changes. We will give here as an example the existence of alternative forms for the future-in-the-past tense typical of some western dialects of Bulgarian (for a fuller description see Xaralampiev 2001: 150). We list them in (19) below:

(19) | | Singular | Plural |
|---|---|---|
| 1 | *šte mislex* | *šte mislexme* |
| 2 | *šte misleše* | *šte mislexte* |
| 3 | *šte misleše* | *šte mislexa* |

Whereas in (18) it is the auxiliary *šta* that is in the imperfect tense (to contrast with the historically present tense of the invariant particle *šte* we find in the future tense), in (19) is it the lexical verb that is inflected (for imperfect tense, and also person and number) whereas the auxiliary verb is the invariant particle *šte*, just as in the future tense.

4.2. Periphrases as nested constructions

We come now to our main point. Even a brief look at the constructions we mentioned in Section 3 can lead to the observation that periphrases seem to be 'related' to each other, in the sense that we can see structural similarities between them.

For example, the future perfect tense in Bulgarian appears related to the future and the perfect. Below in (20) we put 1sɢ forms of the constructions side by side to aid the exposition; we illustrate some of the uses of the constructions in (21).

(20)   Future   Perfect   Future Perfect
       šte mislja   *(az)* **săm** *mislila*   *šte săm mislila*
       I shall think   I have thought   I shall have thought

(21)   a.   *Mnogo   păti   šte   mislja   za statijata.*
            Many   times   want.CL   think.1SG.PRS   of article
            'I will think about the article many times.'

       b.   *Mnogo   păti   săm   mislila   za statijata.*
            many   times   be.1SG.PRS think.PTCP.F   of article
            'I have thought about the article many times.'

       c.   *Do utre   šte   săm   mislila   veče   za statijata.*
            by   tomorrow   want.CL   be.1SG.PRS think.PTCP.F   already   of article
            'By tomorrow I will have thought about the article already.'

The relationship between the three constructions is both semantic and formal. Here we are interested in the formal relationship. Intuitively, the future perfect seems to be 'composed' of the future and the perfect, or we could say the future seems to be 'nested' in the perfect.

Such nesting is not surprising. Periphrases realize functional features on lexemes, e.g. tense, aspect, voice, mood, etc., for verbs. It is rare that only one feature will be realized in a given word form – more often than not a word form carries a whole set of such features. In inflected forms more than one feature often (though not always) means more than one morph. In periphrases, more than one feature seems to lead to constructions being 'blended' together, as it were. It is this 'blending' or 'nesting' of constructions resulting from the combining of morphosyntactic properties that we want to explore further in this paper.

Potentially, we could think of the future perfect as the future tense of the perfect construction (*šte* is added to *săm mislila*)[7], or we could think of the future perfect as the future tense of the auxiliary 'be' (itself periphrastic) being added to the *l*-participle of the lexical verb (*šte săm* is added to *mislila*). We could, of course, also conceive of the future perfect as being completely independent of any similar tense constructions and being built out of the future particle *šte*, the present tense of the auxiliary verb 'be' and the *l*-participle. Any approach will have to account for the fact that it is the *l*-participle (as in the perfect) rather than an inflected form of the main verb (as in the future) that appears in the future perfect.

---

[7] Note that since both *šte* and the present tense auxiliary are clitics, their position in the clause is determined by the rules for positioning clitics and the clitic cluster. Our discussion here does not relate directly to the structures that need to be assumed to account for the syntactic behaviour of cltic constructions.

We believe that Bulgarian gives us an opportunity to distinguish formally between these options on the basis of the alternative, doublet forms of the future perfect, based on alternative, doublet future forms of the auxiliary 'to be'.

In Bulgarian, the verb 'to be' has alternative forms, based on a different root. One of these forms, historically a present, came to be used as a future tense form: *băda*. Later, the future particle *šte*, already grammaticalized as an exponent of the future tense, was added to *băda* and in present-day Bulgarian *šte băda* is an alternative future form for the verb 'to be' (more details of the historical developments can be found in Mirčev 1978: 225 and also in Xaralampiev 2001: 151). Examples to illustrate the forms can be found in (22). Crucially, *băda* can no longer be used as an independent present tense form in main clauses.

(22) a.   *Toj*   *e*   *učitel.*
        he   be.3SG.PRS   teacher
        'He is a teacher.'

     b.   *Toj*   *šte*   *e*   *učitel.*
        he   want.CL   be.3SG.PRS   teacher
        'He will be a teacher.'

     c.   **Toj*   *băde*   *učitel.*
        he   be.3SG   teacher
        '(intended) He is a teacher.'

     d.   *Toj*   *šte*   *băde*   *učitel.*
        he   want.CL   be.3SG   teacher
        'He will be a teacher.'

This alternative future tense form of 'be' can be used in the formation of the future perfect, but *băda* cannot appear as present tense form in the perfect tense construction. We show this in (23) for 1SG forms:[8]

(23)   Future   Perfect   Future Perfect
      šte mislja   (*az*) *săm misila*   *šte săm mislila*
                    *băda mislila   *šte băda mislila*

---

[8] A construction consisting of a present tense perfect aspect form of *băda* 'be' and an *l*-participle of the lexical verb did exist, as we have already mentioned. But this was not the perfect tense construction; rather this was a future perfect construction (*bădešte predvaritelno*). The verb *băda* 'be' in that construction is thought of as a marker of futurity, rather than simply a present tense form. A construction where the simple form *băda* was replaced with the periphrastic *šta* + *băda* appeared as early as the Old Bulgarian period. A historical account of the development of the future perfect can be found in Ivanova-Mirčeva (1962), Ivanova-Mirčeva and Xaralampiev (1999), and Xaralampiev (2001), amongst others.

The availability of these alternative forms for the future perfect suggests that we should analyse the future perfect as being composed of the future tense form of the verb 'to be' (itself a periphrastic form) and the *l*-participle of the lexical verb. We show schematically the analysis we have in mind in (24):

(24) *   [šte   [săm   mislila]]

      [[šte   săm]$_{fut}$ mislila]

      [[šte   băda]$_{fut}$ mislila]

      [BE$_{fut}$ mislila]
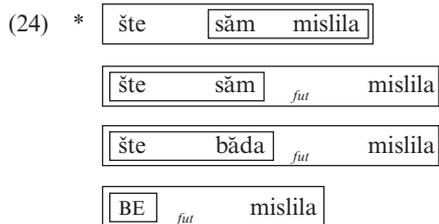
If our analysis is correct, then the future perfect tense construction has to be thought of as a periphrastic form embedded in another periphrastic form.[9] Both periphrastic forms are best thought of as non-compositional, even if for some of them a compositional analysis can be devised. Without belabouring the point, we will just mention here that the negation of the future perfect is in many ways similar to the negation of the future: negative forms following the productive pattern of adding *ne* to the whole construction are possible, but alternative negative forms with the negative future form of the auxiliary 'to be' exist and are indeed more frequent (see example (25) below).

(25) a. *Az*   *ne*   *šte*   *săm*   *mislila*   *po văprosa.*
          I   not   want.CL   be.1SG.PRS   think.PTCP   on matter.DEF
          'I will not have thought about the matter.'

      b. *Az*   *ne šte*   *băda*   *mislila*   *po văprosa.*
          I   not want.CL   be.1SG   think.PTCP   on matter.DEF
          'I will not have thought about the matter.'

      c. *Az*   *njama*   *da*   *săm*   *mislila*   *po văprosa.*
          I   have.not   DA   be.1SG.PRS   think.PTCP   on matter.DEF
          'I will not have thought about the matter.'

      d. *Az*   *njama*   *da*   *băda*   *mislila*   *po văprosa.*
          I   have.not   DA   be.1SG   think.PTCP   on matter.DEF
          'I will not have thought about the matter.'

As we mentioned before, *băda* can appear in the future tenses in combination with *šte*. Though it cannot be used as an independent verb in main clauses, it can appear in subordinate clauses (most frequently ones introduced with *da*) as in the examples in (26) below:[10]

---

[9] We have to emphasize once again that (25) is not meant to represent syntactic structure.

[10] In most cases, we have selected examples found on the internet, though similar examples appear in the literature, see Lindstedt (1985), for instance.

(26) a. *Kogato si   sama,   znači   da   bădeš   štastliva.*
          when   be.2SG.PRS alone.F it-means DA be.2SG.PRS happy.F
          'When you are on your own, you should be happy.'

      b. *Eto   četiri načina da bădeš   kreativen i   da misliš različno.*
          here-are four   ways   DA be.2SG.PRS creative   and DA think   differently
          'Here are four ways (for you) to be creative and to think differently.'

The verb *băda* can even appear as an exception to the present-tense only constraint in *da*-constructions and appear in the re-narrated form in subordinate clauses, see (27):

(27) *Deteto   se   polaga na baštinata si   riza   i   čorapi, za da*
      child.DET REFL   put   on father's   REFL   shirt and socks so DA
      bădelo   xrisimo   i poslušno.
      be.RENAR   timid   and obedient
      'You lay the child over his/her father's shirt and socks, to make it timid and obedient.'

Our claim here, however, is that *băda* cannot appear in the formation of the present perfect. This is true even if we try to embed a present perfect construction in a *da*-dependent clause, for example along the lines of (28):

(28) *Kogato   i   da   săm   xodil   na more, vinagi e   imalo*
      whenever and DA   be.1SG.PRS go.PTCP to sea   always be.3SG.PRS have.PTCP
      mnogo   turisti.
      lots   tourists
      'Whenever I have been at the beach, there have always been many tourists.'

It is not possible to substitute *băda* for *săm* as in (29):

(29) **Kogato   i   da băda   xodil   na   more, vinagi e*
      whenever and   DA be.1SG.PRS go.PTCP to   sea   always be.3SG.PRS
      imalo   mnogo   turisti.
      have.PTCP   lots   tourists
      [Intended] 'Whenever I have been at the beach, there have always been many tourists.'

4.3. The future-in-the-past and the future-in-the-past perfect

Before we try to sketch a formal analysis for some of the constructions we discuss here, we would like to point out that embedding periphrases within periphrases is not rare. We will give one more example with the future-in-the-past and the future-in-the-past perfect. In (30), we show the form of these constructions:

(30) Future-in-the-Past   Future-in-the-Past-Perfect   Future Perfect
      štjax da mislja   *štjax da săm mislila*   *šte săm mislila*
                              štjax da băda mislila   *šte băda mislila*

The future-in-the-past comprises the imperfect tense of the verb *šta* 'want' and an inflected present tense verb marked by *da*. In the future-in-the-past perfect, we see an imperfect tense form of the verb *šta* 'want' linked to the verb *săm* 'be' or the verb *băda* 'be' in the present and the *l*-participle of the main verb. For ease of reference and comparison, we also list in (30) the forms of the future perfect we have discussed so far.

Our focus here is on the future-in-the-past perfect. The complexity of the construction raises again the question of what constructions are nested within this one (if any). As before, one possibility is to assume that *săm mislila* and *băda mislila* are perfect forms of *mislja* 'think'. The problem here is the same one we saw before: *băda mislila* does not occur as a perfect construction. Another possibility, and this is the one we prefer, along with descriptive grammars (see for example Nicolova 2008, 316), is that *stjax da săm* and *stjax da băda* are future-in-the-past forms of 'be', just like *stjax da mislja* is a future-in-the-past form of *mislja* 'think'.[11] We re-iterate here an important point. When we say that the future-in-the-past of the verb 'be' is nested inside the future-in-the-past perfect construction, the justification for this claim is the grammatical form of the words found in that construction. Grammatical peculiarities of the future-in-the-past construction are inherited by the perfect future-in-the-past. At the same time, we automatically account for certain other properties, most notably the fact that the negation of the future-in-the-past perfect forms is expressed by the imperfect tense of *njama* 'not-have', just like the negation of the future-in-the-past. However, this does not mean that the syntactic structure will reflect such nesting relationships transparently. On the contrary, it is precisely here that we observe mismatches in the syntactic structure. We will return to this point below.

---

[11] Historical accounts like the one offered by Xaralampiev (2001: 152–153) suggest that at the outset the construction probably comprised the imperfect form of the verb *šta* 'want' followed by the infinitive of *băda* 'be' followed by the *l*-participle of the lexical verb. The form of *băda* (and *săm*) marked by *da* would have appeared with the demise of the infinitive. The infrequency of future-in-the-past perfect forms, however, means that their development needs to be reconstructed as they occur extremely rarely in the written documents that survive. Ivanova-Mirčeva and Xaralampiev (1999) note in passing that future-in-the-past perfect forms (*bădešte predvaritelno v minaloto*) are not attested in the Middle Bulgarian period. Xaralampiev (2001) puts forward the hypothesis that the future-in-the-past perfect form could move towards fixing the form of *šta* 'want' to be the same for all persons and both numbers, and inflecting the verb 'be' for the imperfect tense. Existing dialectal forms suggest that both *băda* 'be' and *šta* 'want' could be tensed in this construction, see example from Xaralampiev (2001: 153)

(1) *St'axa   b'axa   umr'ali   ot   glat'.*
     want.IPRF   be.IPRF   die.L-PTCP   from   hunger
     '(They) would have had died of hunger.'

5. Periphrasis and the morphology-syntax interface

We come back here to the point that periphrases are morphological and
syntactic. They are syntactic constructions that express grammatical features
and as such are often complementary to morphologically realized forms.
Their status as exponents of grammatical meaning leads to the grammaticali-
zation of some elements of the construction and to exceptional syntactic
behaviour. However, this exceptional syntactic behaviour is not uniform and
is not restricted to periphrases. Their syntactic properties, in other words, do
not necessarily warrant including them in the grammatical model on terms
very different from those that apply to other constructions. What sets them
apart from other structures, however, is the fact that periphrastic construc-
tions in their entirety serve to express morphosyntactic properties. Moreover,
periphrases are productive and general, and are not lexically restricted. This
means that, in some sense of paradigm, they are part of the paradigm of a
given lexical class in the language.

    As a result, we expect periphrases to meet two specific requirements. First,
they should be exponents of morphosyntactic properties (grammatical
features and their values). Second, the constructions themselves are part of
the syntax of the language and, other things being equal, we would have no
reason to think that their syntactic behaviour should be specified within the
morphological component of the grammar. These requirements are difficult
to reconcile and not surprisingly there have been different accounts of peri-
phrases. In some cases, authors adopt a syntax-centred approach in line with
a general trend towards incorporating morphology into syntax (see for exam-
ple, Migdalski 2006 and references therein). Such treatments sometimes go
hand in hand with attempts to render non-compositional interpretations
compositional (see for example Kiparsky 2005 for a reinterpretation of the
Latin perfect passive periphrasis, and Co Vet 2007 on French).

    In other accounts, periphrases are subsumed into the morphological
component. An early appeal for a morphological treatment is made in Sadler
and Spencer (2001), and Ackerman and Stump (2004) offer more detailed
proposals. In Sadler and Spencer's analysis of the Latin periphrastic perfect,
the point is made that the periphrasis realizes a cell in a purely morphological
paradigm. As we have seen, the cell is part of a morphological paradigm
because it lies on the intersection of morphological feature sets, i.e. feature
sets which are otherwise expressed synthetically, by the inflectional morphol-
ogy. They claim that the way to handle this is to permit some kind of reference
(better, 'referral') to a syntactic construction, namely, the predicate adjective
construction type on which the participial periphrasis is parasitic. A similar
claim is made in Spencer (2001) and Spencer (2003), the latter in respect of
the Bulgarian data we will discuss in more detail below. Ackerman and Stump,

by contrast, incorporate a rudimentary syntactic description into the morpho-logical component. The main problem with that approach is that it is entirely unclear how that syntactic proto-distinction is to be integrated with the syntax proper. An important advantage of a morphological realizational approach to periphrastic constructions is that it provides a way to deal with morpho-syntactic non-compositionality. This point is made particularly forcefully in Ackerman and Stump (2004). Morphosyntactic non-compositionality means that the features and values associated with the construction as a whole can-not be computed in a straightforward manner from the features and values associated with the elements of the construction. In thorough-going realiza-tional models of the kind advocated by authors such as Anderson, Aronoff, Corbett, Stump and others there is, in effect, a complete break between the formal morphosyntactic features and semantic interpretation. The interpreta-tion of the feature content of an expression is provided by some mapping which imposes a complex and possibly non-compositional relationship between form and function/meaning. Indeed, in the 'morphology-by-itself' model, some of the features which regulate morphology are 'morphomic', that is, they have no semantic interpretation even in principle. For a more detailed discussion of non-compositionality, see Spencer (this volume). Though we believe that some periphrases can be treated as compositional, the point remains that often this is problematic for a purely syntactic approach.[12]

In a sense, the achievement of Bonami and Samvelian (2009) is to com-bine the insights of both these approaches: they express the idea that the per-iphrase realizes the intersective morphological feature set by incorporating a rule of referral to a syntactic construction, as in Ackerman and Stump's account. However, they differ from Ackerman and Stump in providing a detailed characterisation of the precise syntactic structures concerned, ensuring that those structures are actually the feature descriptions provided independently by the syntax (of Persian, in their case). Thus, they fulfil the (implicit) desideratum of Sadler and Spencer that the form of the periphrase should (at least in canonical cases) be in some sense inheritable from the wider syntax of the language. We outline the essence of their approach and an application of the formal machinery to Bulgarian data below.

6. Explicit approaches to periphrasis

The approach we take as a reference point here combines a realizational approach to morphology with a lexical non-transformational syntactic frame-

---

[12] There is a discussion of the theoretical problems posed by periphrasis and different lexicalist treatments of it in Ackerman et al. (2011).

work for the treatment of syntactic phenomena. In particular, Bonami and Samvelian (2009) combine Paradigm Function Morphology (PFM) along the lines of Stump (2001) with some modifications, and Head-Driven Phrase Structure Grammar (HPSG) along the lines of Pollard and Sag (1994).

Given that Bonami and Samvelian (2009) assume at least two modes for constraining possible linguistic objects (PFM and HPSG), an explicit and principled integration between the two is needed. To achieve this, they posit the following interface between morphology and syntax:

(31)   A sign of type *word* meeting the description below is well-formed only if the PFM grammar licenses phonology [1] and arguments [2] as a realization of the features [4] for the lexeme [3].

$$
\begin{bmatrix}
\text{PHON} & \text{[1]} \\
\text{ARG-ST} & \text{[2]} \\
& \\
\text{HEAD} & \begin{bmatrix} \text{LID} & \text{[3]} \\ & \\ \text{MORSYN} & \text{[4]} \end{bmatrix}
\end{bmatrix}
$$

HPSG is a lexicalist non-derivational syntactic framework. Linguistic objects in HPSG are signs, and are modelled via typed feature-structures. Traditionally (and this is the practice we follow here), HPSG divides objects of types *sign* into two sub-types, namely *word* and *phrase*, though in construction-based approaches constructions can be signs as well. Types are usually assumed to be ordered in hierarchies (for example, *phrase* can have subtypes like *head-complement phrase*, or *head-selector* phrase). Different types have different features associated with them, for example objects of type *phrase* have a DAUGHTERS feature associated with them, whereas objects of type *word* do not. The usual practice is to use Attribute Value Matrices (AVMs) to express descriptions of grammatical objects.

In PFM, the morphosyntactic features and values appropriate for a particular language and their possible combinations are specified as given in the grammar. So, for example, the grammar of Bulgarian will declare that Bulgarian verbs have mood, tense, person, number, voice and aspect and that in the imperative only second-person forms are possible. All the possible feature-value sets appropriate for a particular class of lexemes define the paradigm. Different word forms of lexemes are generated by the paradigm function (PF), which takes the lexemic index and a set of morphosyntactic features appropriate for this lexeme and returns a word and the same set of morphosyntactic features. The paradigm function itself is often the composition of realization rules (RRs). The realization rules are also functions which take a phonological form (for example, the root of a lexeme) and a set of morphosyntactic features and return a (possibly different) phonological form

and the same set of morphosyntactic features. Realization rules are them-
selves organized in blocks (the order of the blocks is stipulative) to reflect
position classes of morphemes. There may be more than one rule in a block
in which case there is competition between them. This competition is resolved
according to a principle which Stump (2001) calls 'Pāṇinian Determinism'.
This is just the familiar principle underlying default inheritance: a more
specific rule overrides a more general one with which it is in competition.
Competing rules have to be formulated in such a way that they can be ordered
in a subset-superset relation (subsumption) so that it is clear which rule is able
to override which.

   Since Bonami and Samvelian (2009) adopt the format for RRs introduced
in Ackerman and Stump (2004), we illustrate the way RRs work in PFM
using that format and then show how such rules are encoded using HPSG
typed feature structures. We illustrate a RR for the 1SG aorist tense form of
the verb lexeme *nosja* 'carry', i.e. the rule that will derive the word form *nosix*.
This rule is formulated as in (32) using the feature structure representations
of Bonami and Samvelian (2009):

(32)

$$\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \end{bmatrix}, \sigma: \begin{bmatrix} \text{PER } 1 \\ \text{NUM } sg \\ \text{TENSE } aor \end{bmatrix} \rightarrow \begin{bmatrix} \text{PHON } X \oplus ix \\ \text{LID } Y \end{bmatrix}$$

In words, this realization rule takes an entity with the phonology X,[13] a lexeme
identity index Y (here CARRY) uniquely identifying the lexeme to be inflected,
and a complete set of morphosyntactic properties σ defining the paradigm
cell occupied by that inflected word form. If the set σ contains the subset of
morphosyntactic properties PERSON:1, NUMBER:*sg*, and TENSE:*aorist*, the
realization rule will return an entity (in this case a word) whose LID is still Y,
but whose phonology has had the string -*ix* appended to it. The string -*ix* then
serves as the realization of the property set PERSON:1, NUMBER:*sg*, and
TENSE:*aorist*.[14] It is important to note that the framework does not assume a
1:1 correspondence between form and (grammatical) meaning and is, indeed,
specifically designed to deal with phenomena that make such correspond-
ences problematic.

   Classical PFM adheres to the principle of A-morphous Morphology
enunciated in Anderson (1992), under which a realization rule serves solely

---

[13] In Stump (2001), by default, X denotes the root of the lexeme in reference to a Block I rule. In
our case X = nos-.

[14] We have simplified here, as we ignore the proper identification of rule blocks and any
morphophonological processes.

to manipulate the phonological form of the input. Such rules therefore do not introduce morphological objects (morphs), much less syntactic objects. To model periphrastic constructions, however, Bonami and Samvelian (2009) and Bonami and Webelhuth (this volume) expand the definition of RRs. The RR has access to an enriched representation of the input lexical form, allowing it to manipulate the valency of the input. In effect, inflectional RRs assume something of the character of derivational morphological rules. A RR that is allowed to manipulate valency (and one that realizes perfect tense, for example) might look as in (33):

(33)

$$
\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \\ \text{SUBJ } LIST_1 \\ \text{COMPS } LIST_2 \\ \text{XEXP <>} \end{bmatrix} , \sigma: \begin{bmatrix} \text{TENSE } perf \end{bmatrix} \rightarrow \begin{bmatrix} \text{PHON } X \\ \text{LID } Y \\ \text{SUBJ } LIST_1 \\ \text{COMPS } LIST_2 \oplus LIST_3 \\ \text{XEXP } LIST_3 \end{bmatrix}
$$

Here we follow the spirit of the proposals made in Bonami and Webelhuth (this volume). They introduce a feature XEXP appropriate for words (and possibly lexemes or other morphological entities) which contains the elements added by the realization rule. These elements can then be linked via an identity relation to the valents of the lexeme: complements, subjects or, potentially, any others. In the case of (33), we have added the XEXP list to the COMPS list.

Another adaptation of PFM necessary to model periphrases involves rules of referral. These are principally needed to model syncretism in morphological paradigms. In Bulgarian, for example, 2SG and 3SG forms of verbs within the aorist and imperfect paradigms are the same. To render this formally, one rule of referral might be allowed to say that the form that realizes 2SG imperfect, for example, is exactly the same as the one that realizes 3SG imperfect. For Stump (2001), referrals are defined at the level of individual RRs, in single rule blocks. Where an entire word form is referred to another word form wholesale, Stump defines the referral over all the rule blocks defining that word form (for instance, as in Stump 2001: 55–6). Stump (2001: 233) defines the periphrastic future tense in Sanskrit by means of a rule of referral (to a derivational form), and Bonami and Samvelian (2009) use referrals in a similar way to model periphrasis in the Persian verb system. We will illustrate this shortly when we come to the analysis of specific periphrastic constructions.

7. An account of the future, perfect, and future perfect

It will be overly ambitious here to try to account for all the periphrastic constructions we have mentioned (and the many we have not). This will necessitate a complete formal account of a large fragment of Bulgarian syntax in a framework that is compatible with the one we are using. What we want to do here, however, is to illustrate the observation that is driving our paper: namely, that periphrases can be 'nested' within other periphrases. An account for this nesting will allow a more economic explanation of why constructions have the composition that they do. Periphrases also exhibit paradigmatic effects, for example in the paradigms of Bulgarian verbs positive forms with *šta* 'want' generally have corresponding negative forms with *njamam* 'not-have'. The syntactic structure of the language does not necessarily mirror this nesting, however. For example, in negated future perfect forms we want to think of a construction like *njama da săm mislila* '(I) will not have thought' as comprising the negated future tense of *săm* 'be' and an *l*-participle. The negated future tense of *săm* 'be' is the periphrastic *njama da săm*. *Njama da săm* is not a syntactic constituent, however. The element *da* is hosting the clitic auxiliary verb *săm* and introducing a subordinate clause syntactically dependent on *njama* as its head.

What we need, then, is to be able to define well-formedness conditions on these constructions without commitment to the specifics of syntactic form and then map their elements to the appropriate syntactic frame. To achieve this, we deploy the mechanisms proposed by Bonami and Webelhuth (this volume). We wish to emphasize, however, that we do not make strong claims about the syntactic analyses we propose. They serve the purpose of making the point above, but without doubt a full analysis will have to consider the syntactic properties associated with the constructions we use as examples here very carefully.

The first point to make is that many of the forms we deal with are clitics and enter the clitic cluster. We sketched the properties of the Bulgarian clitic cluster earlier. For the purposes of our account here, we will assume that the clitic cluster is a syntactic constituent, though how this is ultimately cashed out will depend on details of the morphology-syntax interface that we cannot discuss here. We assume that the positioning of the cluster is stipulative and that it will depend on some interface with phonology which is well beyond the scope of what we wish to do here. The clitic cluster elements we are concerned with here are the present tense forms of the verb BE and the invariable *šte* future clitic. However, the cluster may also include a prosodically placed interrogative clitic *li*, as well as a variety of pronominal clitics. We assume that the clitics are ordered in the cluster and are then placed in the clause by virtue of independent principles which we ignore here. The principles governing cli-

tics need to know which elements are clitics, and also which elements can themselves host other clitics while still remaining part of the cluster (specifically the negation clitic *ne* and *šte*). We therefore assume two attributes with binary values for defining clitics, CLITIC and CLAUSE-INIT. The first marks out all the clitics, the second separates the clitics that can occupy the clause-initial position from those that cannot.

In (34) below we sketch the RR for the perfect tense form of verbs.

(34)   General perfect tense series realization rule

$$
\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \\ \text{COMPS LIST} \end{bmatrix}, \sigma : \begin{bmatrix} \text{TENSE } \boxed{1} \; pres \\ \text{ASP } perf \\ \text{AGR } \boxed{2} \\ \text{GEN } masc \end{bmatrix} \rightarrow \begin{bmatrix} \text{PHON} \quad \text{refer} \left(\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \end{bmatrix}, \sigma \; / \; [\text{V-FORM } ptcp] \right) \\ \text{LID } Y \\ \text{HEAD} \begin{bmatrix} \text{FORM } verb \\ \text{MORSYN } \sigma \\ \text{CLITIC} - \end{bmatrix} \\ \text{COMPS LIST} \oplus \boxed{3} \\ \text{XEXP } \boxed{3} \; \left\langle \begin{bmatrix} \text{LID } be \\ \\ \text{HEAD} \begin{bmatrix} \text{MORSYN} \begin{bmatrix} \text{TENSE } \boxed{1} \\ \text{AGR } \boxed{2} \end{bmatrix} \\ \text{CLITIC} + \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}
$$

The RR in (34) will define the perfect series of forms in (16). By replacing *pres* with *past* we obtain the past perfect forms.

The RR takes a lexeme and the set σ of morphosyntactic properties appropriate for that lexeme and returns a word. For this rule to apply, the set σ should contain the specification *present* for the feature TENSE, *perfect* for ASPECT, some PERSON and NUMBER features (which we have included here under the feature AGReement) and some specification for the feature GENder. Here we have assumed that *perfect* is an aspectual specification and that the tense specification is *present*. Nothing much hangs on this assumption, however. Whatever features and values are present in the set, σ will be passed on to the word that is the domain or range of the RR as a value for the feature MORSYN.

The RR outputs a word whose lexemic index is the same as that of the input. We mentioned earlier that Bonami and Samvelian (2009) allow rules of referral within RRs that constrain periphrastic constructions and allow rules of referral to take only the phonologies of forms. The phonology that appears in (34) above is governed by exactly such a rule of referral. The phonology of the word on the right-hand side of the RR in (34) is constrained to be the

same as the phonology of a form with the same lexemic index and morphosyntactic properties, except for the fact that its value for the feature V-FORM is an *l*-participle.

The output of the rule also introduces a feature XEXP, defining the form of the valents introduced by the RR. The word that is returned by the RR in (34) is a verbal form, but, following Bonami and Samvelian (2009) we assume that its representation is enriched by the addition of the perfect auxiliary to its COMPS list. Thus, the perfect tense RR introduces as an extra argument of the word the present tense form of the verb BE with the same AGReement features as the one in the set σ. The valent introduced in the XEXP list needs to be integrated into the lexeme that is the output of the RR in such a way that it reflects the syntactic structure of the relevant linguistic elements. In this case, the added valent, the present tense form of BE, is a clitic. The XEXP argument is appended to the value of COMPS and it is ultimately ordered by virtue of the clitic-specific linearisation principles.

We now need to show a RR that will produce the (periphrastic) future tense of lexemes. We will show first the future tense of an ordinary lexical verb (such as *mislja* 'think') and then turn briefly to the future tense of the verb *săm* 'be'.

(35) Perphrastic future schema

$$
\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \\ \text{COMPS } L \end{bmatrix}, \sigma : \begin{bmatrix} \text{TENSE } \textit{fut} \end{bmatrix} \;\rightarrow\;
\begin{bmatrix}
\text{PHON} \quad \texttt{refer}\!\left(\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \end{bmatrix}, \sigma\,/\,[\text{TENSE } \textit{pres}] \right) \\[2ex]
\text{LID } Y \\[1ex]
\text{HEAD} \quad \begin{bmatrix} \text{FORM } \textit{verb} \\ \text{MORSYN } \sigma \end{bmatrix} \\[2ex]
\text{COMPS } L \oplus \boxed{1} \\[2ex]
\text{XEXP} \quad \left\langle \boxed{1}\begin{bmatrix} \text{PHON } \textit{šte} \\ \text{HEAD} \begin{bmatrix} \text{FORM } \textit{aux} \\ \text{CLITIC } + \\ \text{CLITIC-INIT } + \end{bmatrix} \end{bmatrix}_{\textit{word}} \right\rangle
\end{bmatrix}
$$

The rule in (35) says that to realize a future form of a verb with a lexemic index Y (say THINK) what is needed is a word whose phonology is the same as the form of this verb with the same set of morphosyntactic features apart from the TENSE:*present* specification, and which on its XEXP list has a clitic whose phonology is *šte*. Thus, for *mislja* 'think' in the 1SG form, (35) will deliver the sequence *šte mislja*. The entry for this verb also stipulates that *šte*

will be added to the list of complements associated with the verb. *Šte* is a clitic that can appear in initial position in the clitic cluster and can enable the cluster to remain preverbal in clause-intitial position (hence the value + for the feature CLAUSE-INIT).

We would expect the grammar to generate in a similar way the future tense form of the verb *săm* 'be'. In principle, this form will be the same as the form of any other verb, and superficially this appears to be the case: the future of *săm* is *šte săm*. The only difference is that *săm* is itself a clitic, as indicated in (36) below.

(36)   Periphrastic future of *săm* 'be'

$$
\begin{bmatrix} \text{PHON } X \\ \text{LID Y} \\ \text{COMPS L} \end{bmatrix}, \sigma : \begin{bmatrix} \text{TENSE } \mathit{fut} \end{bmatrix} \;\rightarrow\;
\begin{bmatrix}
\text{PHON} & \texttt{refer}\left(\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \end{bmatrix}, \sigma / [\text{TENSE } \mathit{pres}] \right) \\[2ex]
\text{LID } Y \\
\text{HEAD} & \begin{bmatrix} \text{FORM } \mathit{verb} \\ \text{MORSYN } \sigma \\ \text{CLITIC +} \\ \text{CLITIC-INIT } - \end{bmatrix} \\[4ex]
\text{COMPS L} \oplus \boxed{1} \\[2ex]
\text{XEXP} & \left\langle \boxed{1} \begin{bmatrix} \text{PHON } \mathit{šte} \\ \text{HEAD} \begin{bmatrix} \text{FORM } \mathit{aux} \\ \text{CLITIC +} \\ \text{CLITIC-INIT +} \end{bmatrix} \end{bmatrix}_{\mathit{word}} \right\rangle
\end{bmatrix}
$$
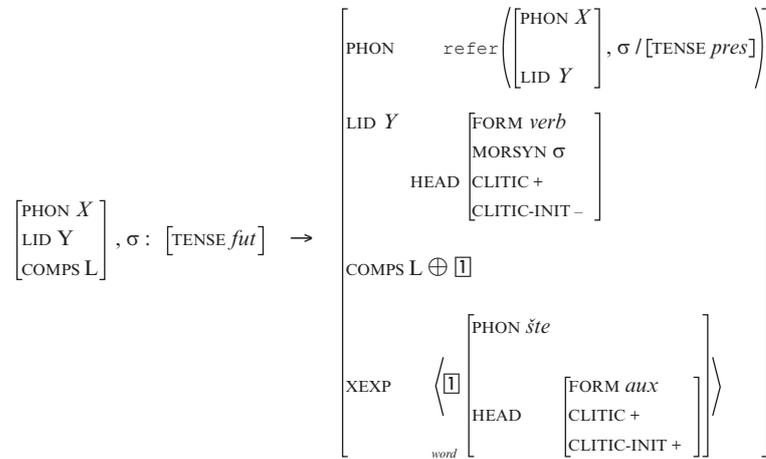
The fact that 'be' is itself a clitic does not prevent it from taking the future clitic as an added complement, of course, even when, as the copula, it is the sole verb lexeme of the clause, as in (37):

(37)   *Petăr   šte   e   na   teatăr   utre.*
       Peter   want.CL   be.3SG   at   theatre   tomorrow
       'Peter will be at the theatre tomorrow.'

We are now in a position to address the crucial point: sometimes one periphrastic construction appears within another one. Unlike in inflectional morphological structures, we do not assume here that to realize the future perfect we need to have two RRs ordered in consecutive blocks. The rules we have been showing so far do two things at the same time: they introduce an element in the XEXP list and also show how this element is integrated in the syntactic structure associated with the head verbal element. Here we want to have two RRs apply at the same time, as it were, but we want the elements on the XEXP list to be integrated in the syntactic structure simultaneously. The AVM in (38) is intended to capture this intuition:

(38)   Future perfect schema

$$
\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \\ \text{COMPS } L \end{bmatrix}, \quad \sigma : \begin{bmatrix} \text{ASP } perf \\ \text{TENSE } fut \\ \text{GEN } masc \end{bmatrix} \rightarrow
$$

$$
\begin{bmatrix}
\text{PHON } \texttt{refer} \; \left(\begin{bmatrix} \text{LID } Y \end{bmatrix}, \sigma / \begin{bmatrix} \text{V-FORM } ptcp \end{bmatrix} \right) \\[2ex]
\text{LID } Y \\[1ex]
\text{HEAD } \begin{bmatrix} \text{FORM } verb \\ \text{MORSYN } \sigma \\ \text{CLITIC } - \end{bmatrix} \\[2ex]
\text{COMPS } L \oplus \boxed{2} \\[2ex]
\text{XEXP} \quad \boxed{2} \left\langle
\begin{bmatrix}
\text{PHON } \texttt{refer} \, [sam] \\
\text{LID } Y \\
\text{HEAD } \begin{bmatrix} \text{FORM } verb \\ \text{CLITIC } + \\ \text{CLITIC-INIT } - \end{bmatrix} \\
\text{XEXP} \quad \left\langle \boxed{1} \; \begin{bmatrix} \text{PHON } \check{s}te \\ \text{HEAD } \begin{bmatrix} \text{FORM } aux \\ \text{CLITIC } + \\ \text{CLITIC-INIT } + \end{bmatrix} \end{bmatrix}_{word} \right\rangle
\end{bmatrix}
\oplus \boxed{1} \right\rangle
\end{bmatrix}
$$

The rule shown in (38) specifies that to realize the future perfect masculine of a verb (say, *mislja* 'think') we need a verb whose phonology is the same as the phonology of the *l*-participle of that verb (e.g. *mislil*) and whose XEXP list contains the future tense form of the verb *săm*, namely the words *šte săm*. We add those two elements on the XEXP of the verb *săm* to the elements on the list of XEXPs of the lexical verb (this is shown by the co-indexing of $\boxed{1}$). The amalgamated list of three word forms is co-indexed with the COMPS list of the lexical verb (this is shown with the co-indexing of $\boxed{2}$).

8. The negated future, the perfect and the negated future perfect

Another series of forms which illustrates the points above is the series of the negated future, the (negated) perfect and the negated future perfect. For ease of reference, the relevant 1SG forms of the verb *mislja* 'think' are presented again side by side in (39):

(39)   Negated future   (Negated) perfect   Negated future perfect
       njama da mislja   *(az) (ne) săm mislil*   *njama da săm mislil*

Again, we argue that from a morphosyntactic point of view the correct way to see the form of the negated future perfect is as the negated future form of the auxiliary BE and the *l*-participle of the main verb (this is in accordance with the understanding of traditional grammarians). Again, the argument can be made on the basis of the alternative future forms of the verb BE in the negated future perfect:

(40) Negated future   (Negated) perfect   Negated future perfect
      njama da mislja   *(az) (ne) săm mislil*   *njama da săm/băda mislil*
                             *(az) (ne) băda mislil

      njama da săm/băda

The challenge of these forms is that the syntactic structure is somewhat different from the morphological provenance suggested above: though in some sense the expression *njama da săm* in *njama da săm mislil* is a single construction (the negated future tense form of the auxiliary BE), in syntactic terms *da* and *săm* are clitics that are hosted by the *l*-participle *mislil*. *Njama* itself does not enter the clitic cluster, and hence can be separated from the clitics *da* and *săm* need to stay adjacent to the verb:

(41) a. *Utre*   *po*   *tova* *vreme* *njama*   *v*   *nikakăv* *slučaj*
          tomorrow   at   this   time   not-have.3SG   in   no   case
          da   *săm*   *dal*   *statijata*   *na*   *redaktora.*
          DA   be.1SG.   give.L-PTCP.M   article.F.DEF   to   editor.M.DEF
          'There is no way I will have given the article to the editor by this time
          tomorrow.'

      b. *Utre*   *po*   *tova* *vreme* *njama*   *v*   *nikakăv* *slučaj*
          tomorrow   at   this   time   not-have   in   no   case
          da   *săm*   *mu*   *ja*   *dal.*
          DA   be.1SG.   he.DAT.CL   she.ACC.CL   give.L-PTCP.M
          'There is no way I will have given the article to the editor by this time
          tomorrow.'

      c. **Utre*   *po*   *tova* *vreme* *njama*   *da*   *săm*   *v*
          tomorrow   at   this   time   not-have   DA   be.1SG.   in
          nikakăv   *slučaj*   *dal*   *statijata*   *na redaktora.*
          no   case   give.L-PTCP.M   article.F.DEF   to editor.M.DEF
          'There is no way I will have given the article to the editor by this time
          tomorrow.'

A proper syntactic analysis of *da* is beyond what we hope to achieve here, but we sketch a possibility below.

We will assume that verbs are furnished with an appropriate value of the binary feature [DA-FORM ±]. The lexical rule changes the value of this feature but also introduces a clitic *da* into the COMPS list of a verb. This is something of an oversimplification, but it serves the purpose of indicating that the
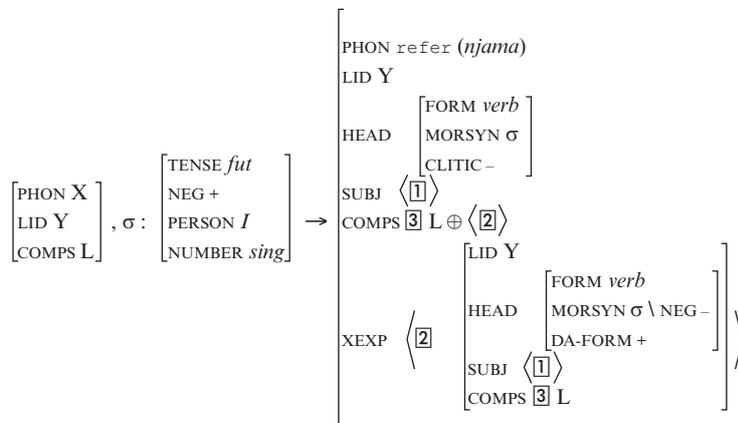
da element functions as a marker of a subordinate clause that has much the same status with respect to a periphrastic expression as an auxiliary verb, namely, the whole clause serves as the partial exponent of a set of grammatical properties, in this case negation and future tense.

(42) *da*-introduction schema

$$
\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \\ \text{HEAD} \begin{bmatrix} \text{FORM } verb \\ \text{DA-FORM } - \end{bmatrix} \\ \text{COMPS } L \end{bmatrix} \rightarrow \begin{bmatrix} \text{PHON } X \\ \text{LID } Y \\ \text{HEAD} \begin{bmatrix} \text{FORM } verb \\ \text{DA-FORM } + \end{bmatrix} \\ \text{COMPS } L \oplus \begin{bmatrix} \text{PHON } \langle da \rangle \\ \text{LID DA} \\ \text{HEAD} \begin{bmatrix} \text{CLITIC } + \\ \text{CLAUSE-INIT } + \end{bmatrix} \end{bmatrix} \end{bmatrix}
$$
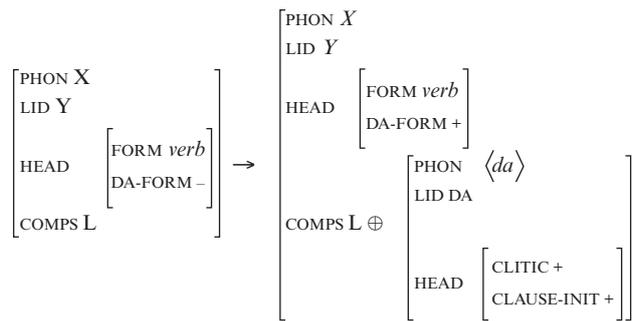
The realization rule for the negative future introduces a *da*-form of a verb into the XEXP list and this verbal form is linked to the COMPS list of the impersonal verb *njama*:

(43) Periphrastic negated future

$$
\begin{bmatrix} \text{PHON } X \\ \text{LID } Y \\ \text{COMPS } L \end{bmatrix}, \sigma : \begin{bmatrix} \text{TENSE } fut \\ \text{NEG } + \\ \text{PERSON } I \\ \text{NUMBER } sing \end{bmatrix} \rightarrow \begin{bmatrix} \text{PHON refer } (njama) \\ \text{LID } Y \\ \text{HEAD} \begin{bmatrix} \text{FORM } verb \\ \text{MORSYN } \sigma \\ \text{CLITIC } - \end{bmatrix} \\ \text{SUBJ } \langle \boxed{1} \rangle \\ \text{COMPS } \boxed{3} L \oplus \langle \boxed{2} \rangle \\ \text{XEXP } \left\langle \boxed{2} \begin{bmatrix} \text{LID } Y \\ \text{HEAD} \begin{bmatrix} \text{FORM } verb \\ \text{MORSYN } \sigma \setminus \text{NEG } - \\ \text{DA-FORM } + \end{bmatrix} \\ \text{SUBJ } \langle \boxed{1} \rangle \\ \text{COMPS } \boxed{3} L \end{bmatrix} \right\rangle \end{bmatrix}
$$

Finally, we come to the negated future perfect, where we have periphrasis nesting. The 'morphological' structure is the negated future form of BE in construction with the l-participle of the lexical verb. We can think of the RR here as a composition of two rules. We can represent this schematically in the following way:

$$RR_{neg,fut,perf}\,(mislja) = RR_{perf}\,(mislja) \circ RR_{neg,fut}\,(be)$$

Syntactically, however, we have a different structure: *da* and *săm* are clitics and are hosted by the *l*-participle. We assume that the syntactic structure is something along the lines of (44):

(44)

njama   *da săm mislil*

with *da săm* entering into a clitic cluster hosted by *mislil*.

Our account here captures the essence of this intuition. We sketch our proposal below:

(45)   Periphrastic negated future perfect schema

$$
\begin{bmatrix} \text{PHON X} \\ \text{LID Y} \\ \text{COMPS L} \end{bmatrix}, \sigma : \begin{bmatrix} \text{TENSE } fut \\ \text{NEG +} \\ \text{PERSON } 1 \\ \text{NUMBER } sing \end{bmatrix} \rightarrow
$$

$$
\begin{bmatrix} \text{PHON refer } (njama) \\ \text{LID Y} \\ \text{HEAD} \begin{bmatrix} \text{FORM } verb \\ \text{MORSYN } \sigma \\ \text{CLITIC } - \end{bmatrix} \\ \text{SUBJ } \boxed{1} \\ \text{COMPS } \boxed{3} \text{ L} \oplus \boxed{2} \oplus \boxed{4} \\ \text{XEXP} \left\langle \boxed{2} \begin{bmatrix} \text{LID BE} \\ \text{HEAD} \begin{bmatrix} \text{FORM } verb \\ \text{MORSYN } \sigma \setminus \text{NEG} - \\ \text{DA-FORM +} \\ \text{CLITIC +} \\ \text{CLAUSE-INIT } - \end{bmatrix} \\ \text{SUBJ } <> \\ \text{COMPS } <> \oplus <da> \end{bmatrix}, \boxed{4} \begin{bmatrix} \text{PHON } \langle mislil \rangle \\ \text{LID BE} \\ \text{HEAD} \begin{bmatrix} \text{FORM L-PART} \\ \text{CLITIC } - \end{bmatrix} \\ \text{SUBJ } \boxed{1} \\ \text{COMPS } \boxed{3} \text{ L} \oplus \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix}
$$

We assume that *njama* (which is represented as a phonological string associated with a form of the verb *mislja*) heads the syntactic structure. The RR rule adds two elements on the XEXP list: the verb *be* in its *da*-marked form, and the participial form *mislil*. The rule also shows the syntactic relationships between these elements and the head: the verb participle form *mislil* is on the COMPS list of *njama* (this is shown by the co-indexing of $\boxed{4}$).

9. Conclusions

Periphrastic constructions are special because of their dual nature — like syntactic constructions they consist of independent elements, but like morphological forms they express grammatical features. Some scholars have taken a purely syntactic approach to them, and others have approached them as forms that realize morphosyntactic features, just like inflected forms do.

One of the hallmarks of periphrases is feature intersection: periphrastic constructions realize some of the set of feature/values that are realized by morphologically inflected forms. Another hallmark is the fact that periphrastic constructions realize features in their entirety and as a result they may realize features that do not appear on any of the constituents, or features whose values are different from those of the constituents. This morphosyntactic non-compositionality is what makes approaches like Paradigm Function Morphology especially well suited to accounting for periphrasis.

What we have tried to show in this paper is that in languages, where there is a rich variety of periphrastic constructions, we find similarities across constructions that suggest we need to account for their relatedness. If we think of periphrases as parts of the paradigm, then we might understand their relatedness in terms of the application of cross-categorisation of morphological features. Such a view helps predict some of the properties of periphrases: the presence or absence of alternative forms, for example.

The morphological composition which results from the way that morphosyntactic properties are combined, and the syntactic structure of periphrases nested within other periphrases are not always isomorphous, however. We have tried to show periphrases that illustrate such deviations from isomorphism and have suggested an account for them. We conclude that nested periphrases constitute a complex interaction between morphology and syntax.

References

Ackerman, Farrell and Gregory Stump. 2004. Paradigms and periphrastic expressions: a study in realization-based lexicalism. *Projecting Morphology*, ed. by Louisa Sadler and Andrew Spencer, 111–158. Stanford, CA: CSLI Publications.

—— and Gert Webelhuth. 2011. Lexicalism, periphrasis and implicative morphology. *Non-transformational Theories of Grammar*, ed. by Robert Borsley and Kersti Börjars, 325–358, Oxford: Blackwell Publishers.

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Avgustinova, Tanya. 1994. On Bulgarian verbal clitics. *Journal of Slavic Linguistics*, 2.29– 47.

Bonami, Olivier and Pollett Samvelian. 2009. Inflectional periphrasis in Persian. In *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar*, 361–180. University of Göttingen, Stanford, CA: CSLI Publications.

Bonami, Olivier and Gert Webelhuth. This volume. The phrase-structural diversity of periphrasis: a lexicalist account.

Bošković, Željko. 2001. *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Amsterdam: Elsevier.

Brown, Dunstan, Greville G. Corbett, Marina Chumakina, Andrew Spencer, and Gergana Popova. In press. Defining periphrasis: Key notions. Morphology 22.

Co Vet. 2007. The Descriptive inadequacy of Reichenbach's tense system: a new proposal. *Tense, Mood and Aspect: Theoretical and descriptive Issues*, ed. by Louis de Saussure, Jacques Moeschler and Genoveva Puskas, 7–26. Amsterdam:Rodopi.

Franks, Steven and Tracy Holloway King. 2000. *A Handbook of Slavic Clitics*. Oxford: Oxford University Press.

Ivanova-Mirčeva, Dora 1962. *Razvoj na bădešte vreme (futurum) v bălgarskija ezik ot X do XVIII vek*. Sofia: Izdatelstvo na Bălgarskata Akademija na Naukite. Conclusions reprinted in *Xristomatija po istorija na bălgarskija ezik*.

—— and Ivan Xaralampiev. 1999. *Istorija na bălgarskija ezik*. Faber, Veliko Tărnovo.

Kiparsky, Paul. 2005. Blocking and periphrasis in inflectional paradigms. *Yearbook of Morphology* 2004, ed. by Geert Booij and Jaap van Marle, 113–35. Dordrecht: Springer.

Lindstedt, Jouko. 1985. *On the Semantics of Tense and Aspect in Bulgarian*. PhD thesis, University of Helsinki.

Migdalski, Krzystof 2006. *The Syntax of Compound Tenses in Slavic*. PhD thesis, Netherlands Graduate School of Linguistics, Utrecht.

Mirčev, Kiril 1978. *Istoričeska gramatika na bălgarskija ezik*. 3rd edition. Sofia: Nauka i izkustvo.

Nicolova, Ruselina 2008. *Bălgarska gramatika: Morfologija*. Sofia: Universitetsko izdatelstvo 'Sv. Kliment Oxridski'.

Pašov, Petăr 1966. *Bălgarskijat glagol. Klasifikacija. Vidoobrazuvane. Slovoobrazuvane.*, volume 1. Sofia: Nauka i izkustvo.

Pollard, Carl and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Popov, Dimităr, Kiril Simov, and Svetlomira Vidinska. 1998. *Rečnik za pravogovor, pravopis i punktuacija*. Sofia: Atlantis.

Rudin, Catherine. 1986. *Aspects of Bulgarian Syntax: Complementizers and WH constructions*. Columbus, OH: Slavica Publishers, Inc.

Sadler, Louisa and Andrew Spencer 2001. Syntax as an exponent of morphological features. *Yearbook of Morphology 2000*, ed. by Geert Booij and Jaap van Marle, 71–96. Dordrecht: Kluwer.

Scatton, Ernest, A. 1984. *A Reference Grammar of Modern Bulgarian*. Columbus, OH: Slavica Publishers, Inc.

Simov, Kiril and Sia Kolkovska. 2002. Interpretacija na *da*-konsktrukciite v Oporna frazova gramatika. In *Proceedings of the 6th National Slavistic Readings: Slavic studies at the beginning of the 21st century. Traditions and expectations*, Sofia.

Spencer, Andrew. 2001. The paradigm-based model of morphosyntax. *Transactions of the Philological Society*, 99.279–313.

—— 2003. Periphrastic paradigms in Bulgarian. *Syntactic Structures and Morphological Information*, ed. by Uwe Junghanns and Luka Szucsich, 249–282. Berlin: Mouton de Gruyter.

—— This volume. Sentence negation and periphrasis.

Stump, Gregory, T. 2001. *Paradigm Function Morphology. A Theory of Inflectional Form*. Cambridge: Cambridge University Press.

Tilkov, Dimităr, Stojan Stojanov and Konstantin Popov (eds). 1983. *Gramatika na săvremennija bălgarski knižoven ezik. Morfologija*. Sofia: Izdatelstvo na Bălgarskata Akademija na naukite.

Xaralampiev, Ivan. 2001. *Istoričeska gramatika na bălgarskija ezik*. Veliko Tărnovo: Faber.