

# *Defining 'periphrasis': key notions*

**Dunstan Brown, Marina Chumakina,  
Greville Corbett, Gergana Popova &  
Andrew Spencer**

**Morphology**

ISSN 1871-5621

Morphology

DOI 10.1007/s11525-012-9201-5



**Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.**

## Defining ‘periphrasis’: key notions

Dunstan Brown · Marina Chumakina · Greville Corbett ·  
Gergana Popova · Andrew Spencer

Received: 27 June 2010 / Accepted: 12 January 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** We examine the notion of ‘(inflectional) periphrasis’ within the framework of Canonical Typology, and argue that the canonical approach allows us to define a logically coherent notion of periphrasis. We propose a set of canonical criteria for inflectional morphology and a set of canonical criteria for functional syntax, that is, syntactic constructions which include functional elements and which express grammatical features. We argue that canonical periphrasis is exemplified in our theoretical space of possibilities whenever a cell in a (canonically morphological) inflectional paradigm (‘feature intersection’) is expressed by a multiword construction which respects the canonical properties of functional syntax. We compare our canonically-based approach with the approach of other authors, notably, Ackerman & Stump (2004), who argue for three sufficient conditions for a construction to be regarded as periphrastic: feature intersection, non-compositionality and distributed exponence. We argue that non-compositionality and distributed exponence, while sometimes diagnostic of periphrasis on a language-particular basis, do not constitute canonical properties of periphrasis. We also examine crucial but neglected syntactic aspects of periphrastic constructions: recursion of periphrases and headedness of periphrastic constructions. The approach we propose allows us to distinguish between constructions in actual languages which approximate the ideal of canonical periphrasis to various degrees without committing us to a categorical distinction between periphrastic and non-periphrastic constructions. At the

---

D. Brown · M. Chumakina (✉) · G. Corbett  
University of Surrey, Surrey, UK  
e-mail: m.tchoumakina@surrey.ac.uk

G. Popova  
Goldsmiths, University of London, London, UK

A. Spencer  
University of Essex, Essex, UK

same time we can capture the intuition that there is in some languages a distinct identifiable set of multiword constructions whose principal role is to realize grammatical features.

**Keywords** Periphrasis · Canonical Typology · Inflectional morphology · Syntax

## 1 Introduction

Grammatical meanings can be expressed in two ways. One is by inflectional morphology where the meaning is expressed in the word itself, such as aspect in Russian, illustrated in (1).

- (1) bolta-l-i  
 chat[IPFV]-PST-PL  
 ‘were chatting’

The imperfective in (1) conveys the idea that the action was going on for some time. These meanings tend to be obligatory: every verb in Russian must be either perfective or imperfective. Another type of situation is when the grammatical distinction is expressed by separate words, such as in Lao (2).

- (2) kamlang2<sup>1</sup>                      son3-siaw3  
 PROG                                  chat  
 ‘were chatting’ (Enfield 2007, p. 209)

In (2) the preverbal modifier *kamlang2* means ‘to be in the process of (V)-ing’. The progressive modifier in Lao is not obligatory, it “does not occur often in texts, and is limited to situations in which the ongoing or extended nature of the action is critical to the current framing of discourse” (Enfield 2007, p. 209). This type of expression must be accounted for by syntax. There is, however, a third possibility: a language can have inflectional ways of realizing some grammatical features, but there are also instances of syntactic structures fulfilling this function. Thus, present and past tense in Russian are realized within the verb, as (3) and (4) show.

- (3) my        bolta-em  
 we        chat-PRS.1PL  
 ‘We chat/we are chatting.’
- (4) my        bolta-l-i  
 we        chat-PST-PL  
 ‘We chatted/we were chatting.’

However, for imperfective verbs the corresponding future in (5) is realized as two words.

<sup>1</sup> Numbers here denote tones.

- (5) my            bud-em            bolta-t'  
       we            be.FUT-1PL        chat-INF  
       'We will chat/ we will be chatting.'

Given the expectation arising from (3) and (4) that a grammatical distinction will be realized by a single inflected word, we have an instance of periphrasis in (5).

Of course, this expectation can come about for a number of reasons, and periphrasis is a term which describes a space of related phenomena. Existing approaches to the problem typically try to address periphrasis in terms of the binary question of whether it is syntax or morphology. In this paper we argue that, logically, it can be both, and we take on the challenge of demonstrating that this represents a coherent claim. We do this by adopting a canonical approach and defining the dimensions relevant for periphrasis. A significant motivation for doing this is that it is a prerequisite for empirical research. This paper makes use of some of the distinctions of Ackerman and Stump (2004), but takes issue with the need for non-compositionality and distributed exponence as criterial, while suggesting that recursion and headedness are relevant. We claim that this follows naturally from basing our definitions on canonical syntax and canonical morphology together, because deviations in terms of non-compositionality and distributed exponence, are themselves deviations from canonical syntax and canonical morphology. Ackerman and Stump's (2004, p. 111) stated aim is to develop, "... an explanatory account of the special characteristics of periphrastic expressions." They do this by applying an inferential-realizational framework. Our aim, in contrast, is to map out the logical space of possibilities. But there are significant points of overlap. The criteria of Ackerman and Stump (2004) are like those of canonical typology, because they describe three dimensions that are not dependent on each other. As Ackerman and Stump claim, the criteria are sufficient but not necessary. As our focus is typological, we argue that it is appropriate to look for several dimensions. Establishing a canonical typology of periphrasis in this way is a necessary step towards determining which of these dimensions are allotted the greatest role cross-linguistically. This is important, because we have descriptions of data relevant to periphrasis in a range of languages, and these descriptions have led to various generalizations. However, we do not know how representative the particular languages are, and how much of the space of possibilities they cover. It may be that properties which occur together in the languages which have shaped our thinking about periphrasis are no more than coincidences in those particular languages. Furthermore, different traditions have produced specific terminology, which may limit our ideas in unfortunate ways. For many linguists, for instance, the term periphrasis is used only of verbal forms but, as Haspelmath (2000) points out, there is no good reason for this restriction, since similar issues arise with other parts of speech.

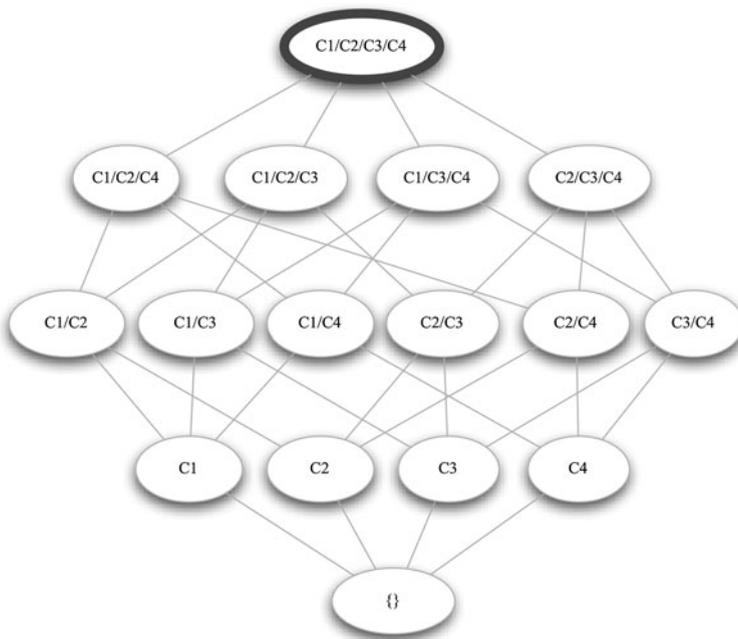
### 1.1 Canonical Typology

In Canonical Typology we take definitions to their logical end point, enabling us to build theoretical spaces of possibilities. We generalize from what we have already

observed to what could exist in principle. An analogy is the system of cardinal vowels, where from vowels of different degrees of openness and frontness phoneticians invoke a potential vowel that is maximally close and maximally front. This serves as an anchoring point for the vowel space, irrespective of whether we find such an extreme vowel in a given language. Unlike the method of classical typology, in Canonical Typology we set out the theoretical space, and only then ask how this space is populated with real instances.

The basic method involves giving *criteria* for a linguistic phenomenon. Each of these defines a dimension which is canonical at one end, and non-canonical at the other. For instance, as part of our claim that periphrasis is canonical syntax and canonical inflectional morphology, we define in Sect. 2 four criteria for canonical inflection. For these criteria the choices are binary ('canonical' or 'not canonical'), although it is also possible to have further points along the dimensions defined by the criteria. Fig. 1 represents the 16 possible types defined by the four criteria as a Boolean lattice.

The first point to note is that the canonical ideal (where all of the criteria line up) is likely to be rare or even non-existent, because it defines one point. In Fig. 1, for example, we have represented instances where the canonical values hold by giving the labels for the criteria (e.g. C1, C2 etc.). There is one point at the top of the lattice (the canonical ideal C1/C2/C3/C4) where all criteria have the canonical value. Decreasing canonicity is defined by greater distance from the canonical ideal. As Fig. 1 illustrates, however, it is possible for some constructions to be equally



**Fig. 1** Lattice of possible types for four criteria

non-canonical. But the canonical method still allows us to rank certain types of construction relative to others. So in Fig. 1, for instance, there are four types of construction which are one step removed from the canonical ideal (but which cannot be ranked among each other), while there are six types which are two steps removed from the canonical ideal (and therefore less canonical than the four above them). This goes on until we reach the bottom point where the construction is as far removed from the ideal as it can be.

We must be careful to distinguish 'canonical' from 'prototypical'. Work on prototypes from its earliest instantiations (Rosch 1973) has emphasized the role of (psychological) salience as the focus for categorial organization. A construction can be salient because it is high frequency, or because it stands out in contrast with other phenomena. Canonical typology is a linguist's analytical method which relies on the use of logical distinctions. If we consider again Fig. 1, a prototype of a phenomenon may actually be one step or two steps down from the ideal, because a particular combination of properties is privileged (either because it occurs frequently in languages or because it stands out).

The canonical approach has been worked out particularly for syntax and for inflectional morphology. In syntax, agreement has figured large, for instance in Corbett (2003, 2006), Comrie (2003), Evans (2003), Polinsky (2003), Seifart (2005, pp. 156–174) and Suthar (2006, pp. 178–198); inflectional morphology has been treated by Baerman et al. (2005, pp. 27–35), Corbett (2007), Nikolaeva and Spencer (2008), Spencer (2007), Stump (2005, 2006), Stump and Finkel (2008) and Thornton (2008). A working bibliography of this growing body of research can be found at <http://www.surrey.ac.uk/LIS/SMG/CanonicalTypology/index.htm>. A potential value of the canonical approach for a typology of periphrasis is that it encourages us to set out a wide space of theoretical possibilities. Within this, we can then get a sense of how representative the currently better studied languages are and of the extent to which the languages so far investigated fill out the space. The approach encourages us to separate out the different elements of periphrasis, which may be overlapping only coincidentally in certain familiar languages. The canonical approach also allows us to handle gradient phenomena in a principled way. And there is a nice practical point: when we say that the examples nearest to canonical are those which are 'best', 'clearest', 'indisputable', the last implies that in defining a particular use of a term we should be able to assume it covers the canonical core; in the ideal scenario, differences in use of terms can be specified in terms of how far out from the canonical point different researchers allow given terms to apply. This can help break the terminological logjams created by conflicting traditions.

## 1.2 Periphrasis

As we saw with examples (3)–(5), periphrasis, at its core, is the situation where we find two (or more) words even though we had a reasonable, morphology-based, expectation of finding only one. This state of affairs can be seen clearly in the Latin examples of periphrasis below. Latin verbs have inflected forms both for the active and the passive voice. Table 1 shows the present tense active and passive forms of

the verb *amāre* ‘love’, and similar inflected paradigms can be shown for the imperfect and simple future tenses.

**Table 1** Active and passive forms of the Latin verb *amāre* ‘love’ in the present tense

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

In the perfect tense (and also in the pluperfect and the future perfect), however, only the active has inflected forms. When passive forms are needed, the language avails itself of syntactic constructions. This is shown in Table 2 for the perfect tense.

**Table 2** Active and passive forms of the Latin verb *amāre* ‘love’ in the perfect tense

	Active		Passive	
	Singular	Plural	Singular	Plural
1	amāvī	amāvīmus	amātus/a/um sum	amātī/ae/a sumus
2	amāvīstī	amāvīstis	amātus/a/um es	amātī/ae/a estis
3	amāvīt	amāvērunt	amātus/a/um est	amātī/ae/a sunt

In the perfect tense the passive construction is composed of passive perfect participle and the present form of the verb *esse* ‘be’ (for the pluperfect the same participle combines with the imperfect tense form of ‘be’ whereas for the future perfect the participle combines with the future simple form of ‘be’).

We think that periphrasis has a morphological part because the inflectional system provides the expectation that we would have found a single form. If in a given language we find a rich array of tenses, each marked inflectionally, and a single tense formed by a combination of two words, say an auxiliary and a participle, then we appear to have a clear instance of periphrasis. In Latin, as we can see, there are three tenses where both active and passive forms are realized by inflectional morphology which in a highly inflected language raises the expectation that inflected forms will realize the passive throughout the tense system. But periphrasis also has a syntactic part because the words that comprise a periphrastic construction are defined in terms of syntax; for example, their relative order may change, other words may intervene between them—as happens in Latin. The particular cells filled by periphrastic forms are not random: as Corbett (forthcoming) demonstrates, periphrastic patterns are always either externally or internally motivated. In the former case, the periphrastic cells make up a natural class; if the latter is true, periphrasis follows the existing morphomic patterns of the language.

In what follows we will explore the properties of periphrasis as a consequence of this dual morphological and syntactic nature. We will show that periphrastic constructions are not all of a kind, but differ from each other in various ways, and this

variation can be described as a relative distance from being canonically morphological and canonically syntactic and the tension between these two. Our criteria will allow us to map out the possibilities so that we can distinguish periphrastic constructions from serial verb constructions, compounds and idioms.

For the rest of the paper we discuss periphrasis in relation to inflectional morphology. We explain this choice in the next section. In Sect. 3 we give our understanding of some phenomena that are linked to periphrasis in crucial ways. In Sect. 4 we set out the criteria for canonical periphrasis. Sect. 5 focuses on those aspects of periphrasis that relate to its morphological nature, and in Sect. 6 we turn our attention to the syntactic side of canonical periphrasis. Next we argue that the tension between morphology and syntax makes periphrastic constructions rather heterogeneous—this is our subject in Sect. 7. We conclude in Sect. 8.

## 2 Inflection

We are concerned in this paper almost exclusively with inflectional morphology. In the next section, for example, we contrast canonical phrases with canonical inflected words. We would like to make it clear here why we do not refer to complex entities which can be said to belong to derivational morphology.

We are interested in syntactic constructions that serve as exponents of grammatical features and interact with inflectional paradigms. It is such constructions, in our opinion, that are most clearly periphrastic. However, we do not deny the existence of lexical stock expansion by multiword expressions (derivational periphrasis). Indeed, in some respects, compounding is exactly that. However, we wish to distinguish (canonical) periphrasis from compounding, since they have particular properties which mean they are not comparable phenomena.

As is well known, the inflection-derivation distinction is extremely problematical (see for example Stump 1998 and references there). However, the distinction regularly proves useful as a rule of thumb or a guideline for descriptive purposes. We will therefore provide a rough-and-ready characterization of canonical inflection. We will adopt a strong position here, namely, that canonical inflection is canonical morphology and that derivational morphology is therefore non-canonical (with respect to inflection).

Some relevant canonical properties of inflection are:

Criterion 1 (Canonical Inflection): obligatory > not obligatory

'Obligatory' means that the inflected form has to be used if the context for its featural interpretation is present. For instance, if (*ceteris paribus*) we wished to refer to a grouping of cats then we have to use the plural of the noun. Although there exists a syntactic expression which is apparently synonymous to the plural form *cats*, this cannot be substituted for the inflected plural *salva grammaticalitate*: *In the garden there were cats* ~ *In the garden there was more than one cat* appears to work, but it won't generalize: *You should feed the cats* ≠ *You should feed more than one cat*/\**You should feed the more than one cat*.<sup>2</sup>

<sup>2</sup> See further Corbett (2008, pp. 10–11) on this criterion and the history of the notion.

Criterion 2 (Canonical Inflection): expresses contextual feature > expresses inherent feature

Contextual inflectional features (Booij 1996) are those that are deployed for the sake of agreement or government. Other features, the so-called inherent features such as number on nouns, tense on verbs and so on, tend to have default semantic interpretations independently of their syntactic context. These are not canonical as inflectional features because they serve more than just a purely syntactic role. Moreover, it can be difficult to distinguish them from derivational features. (See Sect. 5.1 for further discussion of features.)

Criterion 3 (Canonical Inflection): creates a word form of a lexeme > creates a new lexeme

Although inflection and derivation are often opposed, as though at opposite ends of a scale, it is actually difficult to characterize the difference between the two notions. Moreover, the characterization of derivation is tangential to our aims here, so we will offer only brief remarks (for fuller discussion see Corbett 2010b). Derivation is canonically morphology that creates new lexemes by adding a semantic predicate to the representation of the base lexeme. Inflection which is canonical with respect to criterion 2 does not do this, because it serves to realize contextual features. However, both criteria 2 and 3 are required to map out the space of possibilities, because inflection which is non-canonical with respect to criterion 2 (i.e. realizes inherent features) may still be canonical with respect to criterion 3. For example, number inflection expressed on the noun creates word forms of lexemes while involving the expression of an inherent feature. Canonical derivation is like inflection in that it is (canonically) completely transparent, regular and productive (so that both are expressed by means of regular affixation, for instance). Derivation is not, however, obligatory in the way that inflection is. If the expression *driver* is an instance of derivational morphology then *driv-er* and *drive* are distinct lexemes such that *driv-er* means something along the lines of 'person who drives'. However, the existence of a lexeme does not preclude the use of a synonymous word or expression, so we can say either *Here's the driver of the car* or *Here's the person who drives the car*.

Derivation can be contrasted with lexical compounding. In canonical derivation there is only one base lexeme and the derivational process is realized by a morphological operation that adds an affix. In compounding we have a multiword combination, and both words are canonical. We have not investigated the notion of canonical compounding and it is therefore unclear to us whether a canonical compound should be endocentric or not. Arguably, a canonical compound is very similar to a canonical phrase since both consist of two canonical words. Indeed, this is exactly the pattern we find with compounds in many languages of Asia, where the compound is distinguished from the phrase principally in that the compound is used onomastically, that is, to name an entity, property or event while a phrase is used to express a predication.

In addition to criterion 3 canonical inflection is ordered into paradigms.

Criterion 4 (Canonical Inflection): paradigmatic contrast > no paradigmatic contrast

This criterion in essence states that canonical inflection will involve paradigms. It would be wrong to assume that this is really the same thing as criterion 3, because the paradigmatic contrast can arise where words have been bleached of their lexical semantics and obligatorily express a grammatical feature, as we explain in Sect. 5.1.

Having outlined relevant properties of inflection which distinguish it from derivation we outline in the next section the mapping between syntactic structure and (inflectional) morphological structure and introduce an additional criterion for functional syntactic structure.

### 3 The canonical mapping between syntax and inflectional morphology

It is our argument that canonical periphrasis is in essence a situation where both canonical syntax and canonical morphology are present at the same time. But when we set out the key criteria for canonical syntactic and canonical morphological structures, they contradict each other in important ways, as is to be expected. For example, canonical morphology involves a rigid ordering of its elements, whereas canonical syntax involves a greater degree of independence for its elements. However, a key assumption in our approach is that syntax and morphology are separate components of grammar. Canonical periphrasis itself exists in the universe of mappings between syntax and morphology, and so it is necessary for us to define the mapping in canonical terms (Table 3).

**Table 3** Canonical mapping of syntax and inflectional morphology

Syntax	Morphology
Creates syntactic structure on the basis of word class	Inserts words into syntactic structure
Creates syntactic structure on the basis of morphosyntactic features	Spells out morphosyntactic features
Distributes morphosyntactic features	

We assume that one of the roles of syntax is to distribute morphosyntactic features among parts of syntactic structures, typically under agreement or government. We also assume that syntax is blind to purely formal (morphological) properties. Morphology, on the other hand, defines word forms for insertion into syntactic structures and for spelling out morphosyntactic features. This constitutes a key difference between morphology and syntax. In the canonical case we do not expect morphology to distribute features across syntactic nodes, and we do not expect syntax to realize morphosyntactic features. Note that the term 'morphosyntactic feature' does not encompass grammatical relations such as 'subject/object-of', or morphosemantic relations such as 'Agent/External Argument-of', though it does include properties such as 'nominative case' or 'subject agreement'. (For further discussion of the relationships between syntactic structures, lexical structures and morphological organization see Ackerman et al. 2011.)

### 3.1 Canonical syntactic and morphological structures compared

When we talk about canonical syntax and morphology, we have in mind structures that represent clear and easily identifiable instances of syntactic phrases and inflected words. We take inflected word as a starting point and establish the criteria that will define the canonical instances of it. As both syntactic and inflectional structures involve the combination of their respective elements, there are points at which they overlap. Establishing the criteria for these two basic types of structure allows us to identify a third type, canonical functional syntax, which we shall use as a basis for our definition of canonical periphrasis in Sect. 4.

In relation to criterion 4, which we introduced in Sect. 2, canonical syntactic structure contrasts with canonical inflectional morphology in that it is not arranged into paradigms. In Table 4 we set out three further criteria (5–7) which identify dimensions along which the two structural ideals may differ.

**Table 4** Canonical structures of syntax and inflectional morphology

Criterion	Canonical syntax	Canonical morphology
Criterion 5 consists of two elements	yes > no	yes > no
Criterion 6 elements are in rigid order	not rigid > rigid	rigid > not rigid
Criterion 7 elements themselves may bear inflection	yes > no	no > yes

According to criterion 5 a canonical syntactic structure consists of two elements. This property is shared with canonical inflectional morphology. Syntax is typically conceived as dealing in phrase structure, and it follows from criterion 5 that the canonical syntactic structure is binary branching.<sup>3</sup> While the nature of the elements in which it deals may differ, canonical morphology also combines at least two elements. This is not the same thing as binarity in syntax, but our purpose here is to show that there are typically two elements involved in both components, whether branching syntactic structure, or addition of a formative to a base. Criterion 6 treats the elements of a canonical syntactic structure as having some independence of placement within larger structures like sentences, i.e. it is possible for syntactic material to intervene between them. This contrasts with canonical inflectional morphology, of course, where the elements are typically rigidly determined in their order within larger structures and we can find no syntactic context in which something intervenes between them. Finally, elements of syntax can bear inflection, whereas the inflectional elements of canonical morphology are not themselves expected to bear inflection (criterion 7). Real life structures (morphological and syntactic) may depart in various ways from the canonical criteria. For example, in morphology we might have phenomena like distributed exponence, multiple exponence, etc. In syntax we might have structures where the order of elements is quite rigid. A departure from canonical syntactic properties may be interpreted as convergence on canonical properties of morphological structures. For example, a

<sup>3</sup> Criterion 5 need not be understood in terms of phrase structure. For example, it could also be conceptualized in terms of function application in categorial grammar (for which see Bach 1988, p. 21).

restricted freedom of movement of one or both elements in a syntactic structure may make it appear more 'morphological' than others. The dimensions set out in Table 4 allow for a number of possibilities somewhere between canonical syntax and canonical morphology. Of course, when we apply these criteria to a particular construction, we obtain different results for the different elements of the construction under scrutiny.

### 3.2 Canonical functional syntax

We are arguing that canonical periphrasis is both canonical syntax and canonical morphology. In order to advance this claim, however, we need to introduce the notion of canonical functional syntax. Canonical functional syntax is canonical syntax, but where the relationship with the semantics is not quite canonical (Table 5).

**Table 5** Canonical semantics and interpretation

Criteria	Dimension
Criterion 8 semantics of words	all lexical > restricted
Criterion 9 interpretation of the whole	lexical > lexical and grammatical

As criterion 8 indicates, canonically all syntactic elements which require semantic interpretation are members of an open lexical class and have lexical meaning (*black cat*, or Russian *zlaja sobaka* 'fierce dog'). Naturally, the headed nature of syntax means that one of the elements of a syntactic structure, in virtue of being the head of a phrase, may contribute some grammatical information to the interpretation of the combination as a whole. But this element also contributes its lexical meaning to the interpretation of the whole. In contrast, with functional syntax particular words (rather than parts of words) will belong to the non-canonical end of criterion 8 in that they have restricted or conventionalized semantics. Linguists typically associate abstract or conventionalized semantics with bound inflectional marking (see Plank 1994, p. 1672). But being bound is not a prerequisite for conventionalized or abstract semantics. Just as the contrast between lexical and restricted semantics can appear in morphology, typically opposing derivational morphology to inflectional morphology respectively, it can do so in syntax. Criteria 8 and 9 are therefore independent of the criteria for canonical syntax and morphology, but define, in general terms, canonical expectations for the role of semantics in sentence interpretation. Criterion 9 states that in the canonical instance we expect semantic interpretation to be based on the lexical semantics with little or no role being played by the semantics of grammatical elements (be they parts of words, and therefore morphological, or independent words, and therefore syntactic).

What we have termed canonical functional syntax is a violation of this expectation. In a situation in which the non-canonical parts of the dimensions defined by criteria 8 and 9 are in evidence, semantics of the individual words will be restricted and conventionalized in some way, so that some words will be what linguists would generally consider to be function words, and the interpretation of the whole sentence

will involve grammaticalized semantics as well as lexical semantics (either the grammaticalized semantics associated with function words, or that associated with morphological elements such as affixes, or both). Again, it needs to be emphasized that canonical functional syntax *is* canonical syntax, but where the role of semantics is restricted or conventionalized in certain words. As noted, this means that non-compositionality is not even sufficient for defining periphrasis, because both the conventionalized and lexical semantics of canonical functional syntax could in principle be composed in a straightforward way. Compositionality can be understood either in terms of the relation between syntax and semantics as linguistic structure or in terms of a 'bottom up' process in which the semantics of the whole is computed on the basis of functions on the semantics of daughter elements (Nerbonne 1996, p. 473). Given a suitable abstract grammatical meaning for a function word, as with similar morphosyntax, the meaning of the whole could be readily computable in line with this standard view of compositionality.

With criterion 4 we argued that paradigmatic contrast is a canonical property of morphology, and this appears to be a vital part of what it means to be a periphrasis. But paradigmatic contrast is not a necessary element of canonical syntax. In fact, being paradigmatic can be an accidental property of certain syntactic structures. In raising structures, for example, the number of possible verbs involved is limited and the other elements occupy clearly identifiable positions (e.g. *X believes Y to be*; *X expects Y to be*), such that we could view these as paradigmatic. But these are purely by-products of a system which is not inherently reliant on paradigmatic contrast in the way that morphology is. This also means that paradigmatic contrast can be present without this impinging on the canonicity of the syntax involved, again indicating that canonical periphrasis can be both canonical syntax and canonical morphology.

If we take the definition of periphrasis to its logical conclusion, we will expect a canonical syntactic structure as described above to be part of a canonical morphological paradigm filled by canonical morphological structures (Corbett 2007). We rarely find the canonical extreme, of course. We elaborate on the characteristics of canonical periphrasis in the next section.

#### 4 Canonical periphrasis

As we saw in the introductory section, periphrasis has often been identified when there is a syntactic construction that fills a cell in an inflectional paradigm which is otherwise realized by morphology.<sup>4</sup> We take this as the basis for defining canonical

<sup>4</sup> Another way in which various authors seem to use periphrasis is as a syntactic construction that is synonymous with a single word in the same language, for example *be able to* can be defined as a periphrasis in relation to *can* (see Westney 1995). This opens up the possibility that we apply the notion of periphrasis when we have two phrasal elements when we expect (on semantic grounds and by analogy with other lexemes in the language) a single lexeme. One example are the Japanese periphrastic denominal verbs which can be formed by taking a nominal word and adding the verb *suru* 'do' to it, for example *denwa suru* (telephone do) 'telephone'. These are described in Poser (1992). We will restrict ourselves here to periphrasis as part of the grammatical/inflectional system of a language since we believe this is where the canonical instances lie. Indeed some scholars (Börjars et al. 1997) argue specifically that examples like *be able to* should not be considered periphrastic.

periphrasis. Since an inflectional paradigm realizes grammatical features, and functional syntax realizes grammatical features, the construction must be an instance of functional syntax. In other words, the phenomenon which logically exemplifies periphrasis to the greatest degree will have the following properties:

- a periphrastic construction realizes a (canonical) grammatical feature
- a periphrastic construction will occupy a cell in an otherwise inflected paradigm
- a periphrastic construction (like canonical syntax and canonical morphology) will exhibit a transparent relation between form and meaning
- a periphrastic construction is a canonical functional syntactic construction

These characteristics of canonical periphrasis tie with the criteria for canonical inflection and canonical functional syntax we identified before: the first bullet point refers to criteria 1–3, the second one refers to criterion 4. Of course, we expect there to be one realization for the paradigmatic cell. But overabundance, where realizations are in competition, is an attested possibility (Thornton 2011). Given this non-canonical phenomenon in morphology and the fact that periphrasis is both syntax and morphology together, we must allow for the logical possibility that periphrasis and synthesis can be in competition for the realization of a cell. Definiteness marking in the Scandinavian languages shows that this possibility can actually occur. In this instance a suffixed form of the noun competes with the article and noun, and conditions imposed by the syntax determine which is chosen (see Delsing 1993; Börjars 1998; Julien 2005; Hankamer and Mikkelsen 2002, 2005, 2008; and Heck et al. 2008).<sup>5</sup> The next bullet point, namely a transparent relation between form and meaning, describes an important characteristic of both canonical syntax and canonical morphology, and the last point says that canonical periphrasis must satisfy criteria 8 and 9 and contain functional elements.

Note that at first blush canonical periphrastic constructions may not be the ones we would most readily identify as representative of the phenomenon. There are exemplars readily available (in familiar languages or often-cited articles) which come to mind more readily than their properties merit. Constructions which are idiosyncratic in a given language and therefore look like good candidates for periphrasis are not necessarily exemplars of canonical periphrasis. What is important is how a given construction matches up against a set of (hopefully uncontroversial) criteria. Canonical periphrasis is canonical syntax but within the paradigmatic organization of canonical morphology. Given this 'mixed' nature of periphrasis we would expect that some properties of periphrastic constructions follow from their syntactic nature and some from their morphological nature. This is what we will attempt to show in the sections that follow.

## 5 Periphrasis as morphology

Often in languages we find expressed certain abstract, non-referential meanings, for example, location of a proposition in time. Such general meanings may participate

<sup>5</sup> We thank an anonymous reviewer for pointing this out.

in systematic contrasts with other general meanings (anterior to the utterance time, simultaneous to the utterance time, posterior to the utterance time). The expression of these meanings may be obligatory. Such meanings may be related to a systematic variation in the shape of lexemes. For example, in many languages location relative to the utterance time may be expressed by what traditionally has been called *tense* and some class of lexemes in the language (often verbs) will have tense morphology, i.e. a systematic variation in form to express past, present and future tenses.

Languages typically express not one but a number of abstract grammatical meanings like these (often referred to as grammatical features with different values) which are related in complex ways to the shapes of the different lexemes in the language. For example, in many languages verb morphology will contribute to the expression of tense, person, number, voice etc., nouns will have number, case, etc. Because features have different values (for example number can have the values *singular* and *plural*) and these values are generally incompatible (in the sense that no lexeme can be associated with two values of the same feature simultaneously) we can map a logical space of feature/value combinations for each lexeme. This logical space is defined by the intersection of the features found in a given language and is often expressed as the paradigmatic space of a language.<sup>6</sup> A formal expression of this paradigmatic space can be found in Stump (2001, pp. 32ff), where each cell in the paradigm of a lexeme is expressed as a pairing between a form (a word-form) and a set of morphosyntactic properties realized by the cell (property-set).

Periphrasis, as defined by us here will be part of this logical space, i.e. periphrasis must express a grammatical feature and be defined by the intersection of grammatical features in a language. While this property is predicted by Ackerman and Stump's (2004, p. 120) Periphrastic Realization Hypothesis, where the set of realization rules includes statements about both synthetic and periphrastic forms, other theoretical accounts might also entail it. We discuss periphrasis as an expression of a grammatical feature and as feature intersection separately in the next sections. Of course, one presupposes the other, i.e. to be part of the feature intersection of a language periphrasis must, by definition, express a grammatical feature. However, periphrasis as a syntactic construction which expresses a grammatical feature covers a wider range of data than periphrasis as part of some feature intersection. So we will discuss these in separate subsections.

### 5.1 Periphrasis as an expression of a grammatical feature

Making explicit the fact that a periphrastic construction expresses a grammatical feature helps distinguish periphrasis from idioms like *kick the bucket*. Such idioms

<sup>6</sup> This way of defining paradigms is not universally accepted. For some linguists paradigms are simply the set of inflected forms of a language. We have a different understanding of paradigms here, which are defined by features and their values (Haspelmath 2000 writes about this distinction in the context of periphrasis). Of course, there is a link between the two notions of paradigms, in that the definition of features and values in a language may be related to what inflected forms are present in it. Paradigms as defined by us here are a logical extension of the paradigms understood as a set of inflected forms. The distinction is crucial for periphrasis. If paradigms are defined strictly by inflection then syntactic constructions in paradigms will be a contradiction in terms.

have lexical meanings, with varying degrees of predictability, but are not part of the grammatical system. Periphrasis *is* a part of the system, indeed periphrastic expressions canonically form part of a morphological paradigm. There is no paradigm for which idioms form a particular combination of feature values. Thus a construction can be considered periphrastic only if the meaning it expresses is inflectional (grammatical), i.e. an identifiable regular meaning that can be ascribed to the construction independently of the lexical meaning of its parts (see Spencer 2003 for discussion). Of course, to say that periphrasis expresses a grammatical feature presupposes that we have a clear understanding of that concept. For detailed discussion of the problem surrounding this see Corbett (2010a).

In our conception of canonical periphrasis we require that it express a grammatical feature. Grammatical features are not themselves all of a kind, however, so we need to make further distinctions here. We take the canonical grammatical feature to be one which realizes contextual inflection, as follows from criterion 2. Another way to make a similar distinction is to say that features which have a role in both morphology and syntax, i.e. morphosyntactic features, are more canonical and provide therefore better examples of periphrasis. At least some of the values of a morphosyntactic feature must be subject to syntactic constraints (agreement or government). Typical examples are gender and person. Morphosyntactic features can be contrasted to morphosemantic features. Morphosemantic features are semantically charged and are reflected in morphology, but are not relevant in syntax; tense and aspect are often of this type (compare Stump 2005, p. 52).

Contextual features are all morphosyntactic, while inherent features can be both morphosyntactic (like gender in nouns which control agreement) and morphosemantic (like tense in verbs). The most canonical will be periphrasis expressing contextual features, further away from the canonical centre are inherent morphosyntactic features, yet further are the instances of periphrastic expression of inherent morphosemantic features, since the justification for the feature will have no syntactic backing. Here again we must flag the difference between canonicity and prototypes: verbal tense is a prototypical but not canonical instance of periphrasis.

The canonical typology approach allows for the canonical instances to be very rare or non-existent. It is precisely the case with the type of features realized by periphrasis: typically, the features are morphosemantic. It is very rarely that periphrasis is used to realize contextual features. One such instance is the realization of noun case in Nenets. We discuss this in detail in Sect. 5.2. in connection with another canonical property of periphrasis, that of feature intersection. Here we only want to point out that the feature realized periphrastically is at least partially contextual. In Nenets, four case values: dative, locative, ablative and prosecutive have periphrastic expression in the dual. These cases mostly have semantic usage, but not entirely. There are verbs in Nenets that govern some of them. Thus, 'want' governs the dative, 'be afraid' governs the ablative, 'think' governs prolativative, etc. (We give examples of the Nenets periphrasis in the next section, Table 6.)

A relevant issue in the definition of periphrasis is whether it is sufficient to say that it is a syntactic construction that expresses a grammatical feature. It will mean admitting within the realm of canonical periphrasis potentially problematic forms like the English *going to* construction or the English construction of an (in)definite

article and a noun. A canonical approach as the one we undertake here does not require a definitive yes or no answer. We think that canonical instances of periphrases should be constrained further to be part of morphological paradigms, i.e. part of sets of forms found in opposition where variation of form is linked to more than one feature at the same time. Typically, where we find such variation of form at least some of the forms will be inflected.

We can make further distinctions within this general picture. Suppose, for example, that we have a feature system which realizes its values either in morphology or in syntax in a completely transparent fashion but assigning each feature value to exactly one overt exponent and vice versa, as in (6).

- (6) F:a da  
 F:b du  
 G:c bi  
 G:d bo

Suppose, too, that {da, du} are always in some strictly defined linear order with respect to {bi, bo}, whether they are function words or affixes. Now contrast this with a feature system such as that in (7) in which one of the values of each feature is left unexpressed:

- (7) F:b du  
 G:d bo

The system in (6) is more canonical in the sense that there is a very regular and transparent relationship between form and meaning/function. The system in (7), however, allows us to make some inferences as to the grammaticalization of the elements involved. The reason for this claim is the following. We regularly encounter difficulties in deciding whether or not an incipiently grammatical construction represents a grammatical opposition (yet) or whether it's still a lexical opposition with bleached semantics. For instance, we might find that a language is beginning to use a verb *finish* to signal perfective aspect, a potential grammaticalization pathway (Bybee et al. 1994; Heine and Kuteva 2002, p. 138). A criterion that is very helpful in determining whether such a grammaticalization has taken place is whether the exponent is obligatory in contexts which demand its use. For instance, suppose we have a language with no verbal morphology but which often uses the verb FINISH in contexts which imply no more than a perfective interpretation of the verb. Compare these expressions in (8).

- (8) a. Mary write letter  
 b. Mary finish write letter

Typically, if FINISH is simply a lexical verb whose meaning is getting bleached then (8b) will be vague: it may refer to a perfective event or to an imperfective event.<sup>7</sup>

<sup>7</sup> The vagueness of this hypothetical example is in line with what happens in English (but with the appropriate inflections). So 'Mary regularly finishes writing letters on Wednesdays' or 'Mary was finishing writing the letter when John disturbed her' show that the verb *finish* is not restricted to perfective uses in English.

However, a signal that FINISH has been grammaticalized as an aspect marker would be if (8b) could only be interpreted perfectly. Then we would have to set up a grammatical feature, say, [Perfective:{yes, no}] such that FINISH is the exponent of [Perfective: yes] and there is no exponent of [Perfective: no], similar to our examples in (7). Because FINISH has now become obligatory in perfective aspect contexts we can no longer regard it (in this use) as simply a verb with the meaning 'finish'.

An example from Hungarian illustrates this point. In Hungarian, a construction with the verb *fog* 'catch, hold' expresses the future tense. Compare (9) and (10).

(9) lát-ni fog-ja  
 see-INF hold-3SG.DEF  
 'He will see (it).' (Rounds 2001, p. 50)

(10) fog-ni fog-om  
 hold-INF hold-1SG.DEF  
 'I will hold (it).' (Edith Moravcsik, p.c.)

There are indications that this construction has undergone some stages of grammaticalization: (10) demonstrates that when used in this construction, the lexical semantics of *fog* gets bleached. *Fog* also loses some of its morpho-syntactic possibilities: as an expression of the future, it can be only used in the present indicative, whereas when used as an independent verb, it has all tense-mood forms. Examples (9) and (10) can only be interpreted as future tense. However, present tense forms in Hungarian are ambiguous between present and future interpretation, so we cannot talk about a grammatical opposition here, and neither can we talk about the obligatoriness of the future (criterion 1). The canonical approach allows us to see where periphrasis in Hungarian is in relation to the canonical centre, since we compare it not to instances from other languages (which can be equally non-canonical), but to a logical construct.

Where we find syntactic constructions in otherwise morphological paradigms, we talk about feature intersection—this is a phenomenon we discuss in more detail below. It has been elaborated in the work of Sadler and Spencer (2001) and particularly Ackerman and Stump (2004). Ackerman and Stump (2004) also single out non-compositionality and distributed exponence as sufficient conditions for genuinely periphrastic constructions (non-compositionality is the subject of our Sect. 5.3; we discuss distributed exponence in Sect. 6.3). The presence of syntactic constructions in morphological paradigms makes them similar in some respect to word-forms, for example the construction as a whole is usually associated with a feature-value set analogous to the feature-value sets associated with the word-forms of lexemes. The fact that this grammatical information is carried by the construction as a whole allows form-function mismatches similar to the ones found in inflected word forms. For example, a word-form may be associated with some grammatical information which cannot be 'localized' onto a part of that word form (which sometimes leads to suggestions of zero morphs), or some information might be repeated redundantly on more than one element of the word form (multiple

exponence). Such mismatches could well be more typical of periphrastic forms than of other syntactic constructions, or may indeed occur exclusively in periphrastic constructions. It is important to emphasize though that non-compositionality and/or distributed exponence are not hallmarks of *canonical* periphrases as conceived by us here.

## 5.2 Periphrasis and feature intersection

That periphrasis should be defined as part of a paradigm where features intersect is implied strongly in the work of Matthews (1974). For instance, when talking of the Latin passive forms like *amātus sum* he writes: “There is a paradigm, it is implied, with both single words and groups of words among its members.” (1974, p. 171). More importantly, he notes that the traditional treatment of the verbal paradigm in Latin (including periphrasis) “is restricted to the forms which fill a gap (e.g. for ‘Future Infinitive’) in the intersections of the traditional categories or ‘accidentia’ ...” (1974, p. 172).

As we have indicated in Sect. 5.1, Nenets provides a clear example in the nominal domain of a paradigm where features intersect.

**Table 6** Nenets *ti* ‘reindeer’

	singular	dual	plural
nominative	ti	tex <sup>o</sup> h	tiq
accusative	tim	tex <sup>o</sup> h	ti
genitive	tih	tex <sup>o</sup> h	tiq
dative	tenəh	tex <sup>o</sup> h n’ah	tex <sup>o</sup> q
locative	tex <sup>o</sup> na	tex <sup>o</sup> h n’ana	tex <sup>o</sup> qna
ablative	texəd <sup>o</sup>	tex <sup>o</sup> h n’ad <sup>o</sup>	texət <sup>o</sup>
prosecutive	tew <sup>o</sup> na	tex <sup>o</sup> h n’amna	teqm <sup>o</sup>

This table is based on Salminen (1997, pp. 119–120), and is also cited in Ackerman (2000, p. 3). The meaning expressed by the given forms is regular and obligatory in that a certain case-number value must be assigned to every word form. It is expressed synthetically for most cells of the paradigm, except for the four cells in the dual (dative, locative, ablative, prosecutive) where the case form consists of two words (*tex<sup>o</sup>h n’ah*, *tex<sup>o</sup>h n’ana*, etc.), and looks like a syntactic phrase. But the function of this phrase is comparable with the function of the synthetic forms in the other parts of the paradigm: the meaning of the phrase is inflectional and the periphrastic construction is used in the same syntactic environments as respective cases in the singular and plural. Inevitably, the facts of the matter are somewhat more complex, in that the periphrastic construction shows signs of being only partially grammaticalized. In particular, the forms with the postposition are not available in all contexts, sometimes a quantification by the numeral ‘two’ (which takes the singular) is the only grammatical choice. Nikolaeva (forthcoming) claims that the use of the dual correlates with definiteness and/or discourse givenness. We

need to stress that periphrasis is the only possibility for the oblique dual cells in the paradigm to be filled, and that this complication concerns the usage of the dual in Nenets. The data as shown in Table 6 serve to illustrate our point, but in the nominal domain. We are very grateful to Irina Nikolaeva for sharing her data and for discussion of their significance. For full details see Nikolaeva (forthcoming).

Of course, just because we can fill certain cells in a table with periphrastic forms does not mean that we have feature intersection. We therefore need to be clear what we mean by 'feature intersection'. The notion only arises when we have features that can combine to define a cell in the paradigm space. We crucially need to rely on one simple notion. Consider the situation in which a feature  $F$  with a given set of values,  $f_1, f_2, \dots, f_x$  can combine with another feature  $G$  with its own set of values  $g_1, g_2, \dots, g_y$  so that each inflected form in the paradigm has feature value specification like  $\{f_1, g_1\}, \{f_2, g_1\}, \{f_1, g_2\}$ , etc. Assuming no restrictions on combinations, these features define a simple space with  $x$  times  $y$  cells. Synthetic and periphrastic forms can be distributed in this space in a number of different ways. Some features might be wholly associated with synthetic forms. To establish this, we consider the realization of any feature value and ask whether there is some cell in the paradigm where that feature value is expressed morphologically (synthetically). Let us call any such feature value an  $m$ -feature value (following Sadler and Spencer 2001). The normal expectation is that if a feature value is an  $m$ -feature value then it will be expressed morphologically for every cell in the paradigm, in other words if a certain value of a feature is realized morphologically in some cells of the paradigm, it creates the expectation for other cells to be expressed morphologically. If for a given collection of intersecting features every value of a certain feature is realized morphologically, we call those features 'm-features'. Where all interacting features are  $m$ -features we have a wholly synthetic paradigm.

In the Tundra Nenets paradigm, the intersecting features are Number and Case. We can conceive of the Case feature as associated with the set of (atomic) values  $\{nominative, accusative, genitive, dative, locative, ablative, prosecutive\}$ . The Number feature is associated with the set of (atomic) values  $\{singular, dual, plural\}$ . When we consider the Case values we find that for each value it is true that a synthetic form exists (for some value of Number). This happens throughout the singular and plural, including for the forms associated with Case: *dative*, Case: *locative*, Case: *ablative*, Case: *prosecutive*. When we consider the Number values we find, too, that for some Case value each of those Number values has a synthetic expression. Crucially, this is true of all three number values in the nominative, accusative and genitive forms. This means that both Case and Number values are  $m$ -feature values (following Sadler and Spencer 2001; Spencer 2008). Because all feature values of case and number are  $m$ -feature values, the expectation is that all possible legitimate combinations of Case/Number will be expressed synthetically. This expectation is not realized, however, for the periphrastic dual forms.

Consider further the Tundra Nenets example. Suppose for the sake of the argument that all case forms in the dual were expressed analytically. Then we would have a value of the Number feature which never receives a synthetic expression. It is possible that we would still want to talk about a periphrastic dual subparadigm, but we could not do so on the basis of feature intersection. Under this scenario, the

feature value Number: *dual* is no longer an m-feature value and we do not have an intersection of genuine morphological features. Such a system would be open to an alternative analysis under which dual number was not part of the paradigm as such, but rather was a syntactically expressed grammatical construction (note that it is a construction which does express a grammatical feature), in much the same way that definiteness in English is a syntactic construction, expressed by the definite article. Our aim here is to point out that a syntactic dual subparadigm which is not associated with feature intersection will be, in the very least, a less canonical instance of periphrasis.

Finally, suppose we decide to adopt a more articulated analysis of the Tundra Nenets case system and treat Case as a feature with an internal value structure, namely Case: {DirectCase, ObliqueCase}, where the two values of Case are defined as DirectCase: {*nominative, accusative, genitive*} and ObliqueCase: {*dative, locative, ablative, prosecutive*}. Will this allow us to factor out a portion of the paradigm and undermine the analysis of periphrasis as feature intersection? The answer is 'no'. Both DirectCase and ObliqueCase receive morphological expression through the singular and plural forms, so they are m-feature values.

We can run a similar argument with respect to the Latin passive perfect periphrasis illustrated in Table 2 above. Here the features that interest us are Voice with values {*active, passive*} and Aspect with values {*perfect, imperfect*}. The Person, Number, Tense and Mood features are irrelevant to our computation because these are expressed synthetically throughout the paradigm. We are therefore effectively dealing with a four-celled paradigm. The values of Voice are m-feature values because they both receive synthetic expression (in the Aspect: *imperfect* forms at the very least). Similarly, the Aspect values are m-feature values because Aspect: *imperfect* is always expressed synthetically, and Aspect: *perfect* is synthetic in the Voice: *active* subparadigm.

However, consider a hypothetical *Latin'* in which the passive is periphrastic throughout the whole of its paradigm. This situation would be the same as the hypothetical situation in which Tundra Nenets lacked a morphological dual number. On this hypothetical scenario there would be no combination of Voice: *passive* with Aspect which received purely morphological expression. In fact, there turns out to be no feature in the *Latin'* conjugation that intersects with the passive subparadigm in such a way as to give a synthetic form. Therefore, Voice: *passive* would not be an m-feature value in *Latin'* and we would not be able to appeal to the criterion of feature intersection to define the passive paradigm as periphrastic. In our terms, the passive periphrasis in *Latin'* is in the very least a far less canonical example of periphrasis than the passive perfect periphrasis in Latin proper.

### 5.3 Non-compositionality

Ackerman and Stump (2004, p. 142) argue that (morphosyntactic) non-compositionality is one of the most reliable criteria for diagnosing the paradigmatic role of analytic combinations. Their criterion is quoted below:

Non-Compositionality: If the morphosyntactic property set associated with an analytic combination C is not the composition of the property sets associated with its parts, then C is a periphrase. (Ackerman and Stump 2004, p. 142)

For Ackerman and Stump (2004) non-compositionality is a sufficient condition for defining periphrasis, but not a necessary one. Spencer (ms., p. 3) is in agreement about it not being a necessary condition, but also argues that it is not sufficient. Spencer (ms., pp. 6–7) divides non-compositionality into two different types: *idiomaticity* and *feature-clash*. In *idiomaticity* the whole construction is associated with some morphosyntactic content which is not part of the morphosyntactic information of any of the elements of the construction. A good example of this is the English *be + V-ing* construction. The overall construction is associated with a value for the feature aspect, but none of the elements of this construction has an aspectual feature associated with it.

However, in all putative cases of such *idiomaticity* it is, in principle at least, open to the linguist to propose some kind of non-idiomatic analysis. For instance, we might say that the verb 'be' just has as one of its many meanings that of aspect: progressive. (Alternatively, and with marginally more plausibility, one might say that the *-ing* form of a verb bore that feature value.) The difficulty here mirrors the debates which surround the issue of *idiomaticity* in general (and particularly the problem of 'semi-idiomaticity'). For instance, the non-idiomatic components of many idioms seem to be open to modification in a way which suggests that they may bear some (minimal) meaning of the appropriate kind. If we say 'the genie is out of the bottle and there's no way we can entice him back in again' it might appear as if *genie* is referential, denoting 'some undesirable situation which cannot be reversed' or some such. (See, for instance, Nunberg et al. 1994 and the papers in Everaert et al. (eds) 1995 for discussion.)

*Idiomaticity* contrasts with *feature-clash*, where morphosyntactic features are defined on the elements of the construction, but the value associated with the overall construction is incompatible with values on elements of the construction, or where the values of the individual elements are incompatible with each other. Feature clash in this sense can be illustrated with the example of non-compositionality given by Ackerman and Stump (2004, p. 143). They refer to a negative past II verbal form in Eastern dialects of Mari.

- (11) kolen            oməl  
       die.GERUND    be.PRS.NEG.1SG  
       'I didn't die'.

The overall construction is associated with the tense value *past II* and though the form *oməl* is a tensed form of the verb 'be', the value of the tense feature here is *present* so the tense values associated with the construction and with the element(s) of it are incompatible. (Further examples and discussion can be found in Spencer ms.)

Finally, we also find instances in which a periphrasis contains elements whose feature values clash with each other. This can be illustrated with the following example from Czech (Eva Hajičová and Jarmila Panevová, p.c.):

- (12) by-l-a            jste            velmi        laskav-á  
       be-PST-F.SG    AUX.2PL    very        kind-F.SG  
       ‘You were very kind’ (polite address to a woman)

Czech data including the example above is discussed in Corbett (2010a) and Corbett (2006, pp. 86–87). Similar data for Bulgarian is discussed in Popova (2010).

It may well be that non-compositionality is a property of syntactic constructions associated with periphrasis, and is not a property of syntactic constructions in general. It is not, however, a relevant criterion for defining canonical periphrasis. In fact in our understanding both canonical morphology and canonical functional syntax display a transparent relationship between meaning and form (for morphology in particular see Corbett 2007). However, it appears that morphosyntactic non-compositionality, though not canonical, is more readily associated with a range of related phenomena within the purview of morphology, most clearly in instances of dependency where the meaning of word forms cannot be inferred on the basis of their component parts, as they mean entirely the opposite of what one would expect. So syntactic structures that are morphosyntactically non-compositional appear to be less ‘syntactic’ and more ‘morphological’. It is for this reason that when periphrases are (morphosyntactically) non-compositional we think of them as more obviously periphrastic and this could well be an important consideration when debating how periphrases are to be treated in grammatical theory, but it need not be a defining property of canonical periphrasis. Morphosyntactic non-compositionality is also relevant to headedness in periphrasis. We return to headedness in Sect. 6.2.

## 6 Periphrasis as syntax

Periphrasis is not a way of creating new names for things, which is what sets it apart from compounding, though like compounding it relies on the combination of two or more elements. Instead, like inflection, periphrasis is canonically part of the system of grammatical contrasts. Though the grammatical meaning associated with a periphrastic construction may be expressed by one or more elements of the construction or indeed by none of them, the lexical meaning is usually associated with only one of the elements. This means that in practice a periphrasis would consist of one lexical word and one or more function words.

We can see the relationship between canonical syntax and canonical periphrasis by considering the following hypothetical situation. Suppose we have a language with two intersecting features,  $F:\{f_1, f_2\}$ ,  $G:\{g_1, g_2\}$  such that the cell  $[F:f_1, G:g_2]$  is expressed periphrastically. For concreteness, we might imagine a language with morphologically expressed verbal features like tense, mood, aspect, agreement, and affirmative/negative polarity, such that Polarity: *negative* in combination with one other feature value (say, Tense: *past*) is expressed by an analytic construction. There

are three principal ways in which negation could be expressed periphrastically in such a case:

- (i) Negation is expressed by an uninflecting particle in combination with a verb which is fully inflected for all the other features. The particle has exactly the same syntax as any other similar adverbial particle in the given language.
- (ii) Negation is expressed by means of an auxiliary element which attracts all of the other inflections and which combines with a bare form of the verb (or perhaps some special 'converb' form). The auxiliary has exactly the same syntax as any other auxiliary verb in the language.
- (iii) Negation is expressed by means of an auxiliary element which attracts some but not all of the inflections and which combines with a form of the verb which is inflected for the remaining features. The auxiliary has exactly the same syntax as any other auxiliary verb in the language.

This is represented schematically below:

- (i) Neg                    V-infl
- (ii) NegAux-infl    V
- (iii) NegAux-infl1 V-infl2

We can now ask which of the three possibilities is canonical with respect to the others. This question requires us to consider whether a language typically fits within the canonical head- or dependent-marking types. For instance, if a language is overwhelmingly head-marking, then a periphrastic construction which is realized as dependent-marking will be non-canonical with respect to the rest of the grammar, and vice versa, a head-marking periphrase will be non-canonical with respect to an overwhelmingly dependent-marking grammar. It is beyond the scope of this paper to develop a fully-fledged account of this so we will limit ourselves to illustrative remarks.

The type (iii) construction is canonical neither as syntax nor as morphology. As morphology it is not canonical because the Polarity: *negative* property should be expressed morphologically in this language but in the type (iii) construction it is expressed as a separate syntactic terminal. However, in addition, some of the morphology that should appear on the lexical head now appears on the functional head. Ex hypothesi, the language under consideration is head-marking and expresses its verbal properties on the verbal lexical head. Therefore, if some of those properties are instead expressed on a functional head the morphology-syntax interface will be non-canonical (for this grammatical system).

In terms of canonical syntax type (i) is the most canonical way of realizing the Polarity: *negative* value. This is because it involves simple compositional concatenation of two syntactic terminals. For a language which predominantly expresses grammatical features as head-marking of lexical elements type (i) will therefore be the canonical form of the periphrase. However, we might be dealing with a language in which grammatical properties are marked predominantly on functional heads. In such a language type (ii) negation will be the canonical type of periphrase.

The reason the toy example above is important is that in practice we often find languages which have periphrastic constructions having the properties illustrated in (i–iii) above. Specifically, it is quite common for languages to have periphrastic negation of the kinds illustrated above (Honda 1996; Kahrel 1996; Spencer 2008). Now, if the negation is only ever expressed by means of an uninflecting adverb-like particle, as is broadly speaking the case in languages such as Italian, then we would be very unwilling to speak of a periphrastic construction (at best, we could say that such a construction was a highly non-canonical periphrase). On the other hand, if we met a language with intersective morphological negation and with an idiosyncratic syntactically expressed construction such as that described in (iii) above we would probably take the idiosyncratic nature of the construction as evidence of a periphrasis. This is no doubt the correct response: the idiosyncrasy of the construction aligns it more with (non-canonical) morphological modes of expression and therefore makes it easier to distinguish such a construction from a purely non-periphrastic syntactic construction. But that does not mean that such a construction is a canonical periphrasis (though it may well be thought of as in some sense prototypical).

The way negation is expressed in Nenets comes close to what is described in (iii). The negative construction consists of the auxiliary which realizes person, present and past tense and also contains information on whether the verb belongs to the subjective, subject-object, or reflexive conjugation. The imperfective aspect is, however, realized by the lexical verb:

- (13) jam                    n'i-w°                    p'ire-mb'u  
       soup.NOM.SG    NEG-SG.OBJ.1SG    boil-IPF  
       'I am not making soup.' (Irina Nikolaeva, p.c. See also Nikolaeva forthcoming).

The negation in Nenets, however, is always expressed periphrastically and therefore we cannot talk about morphological intersectivity here. Another important aspect of (8) is that Nenets is a head final language, and so the negative auxiliary does not behave as a syntactic head.

Another possibility of marking negation is presented by the Australian (Tangkic) language Kayardild. The Kayardild verb marks negation synthetically in some cells of the paradigm but not in the past (Table 7):

**Table 7** Partial paradigm of a Kayardild verb (based on Evans 1995, p. 255)

Function	Positive	Negative
imperative	-TH.a	-na
actual	-TH.a	-TH.arri
immediate	-TH.i	-nang.ki
potential	-TH.u(ru)	-nang.ku(ru)
past	-TH.arra	–

The element–TH in Table 7 represents either of the two ‘thematics’ which signal conjugation membership (Evans 1995, pp. 253–254)

In the past negative sentences a negative noun phrase is used, as in (14).

- (14) ngada            kurri-jarra        **warirra-na**            **dangka-walath-ina**  
 1SG.NOM        see-PST            nothing-M\_ABL        person-lots-M\_ABL  
 'I saw no-one, I saw no groups of people' (Evans 1995, p. 375)

It is important to notice here that *kurrijarra warirrana dangkawalathina* represents an ordinary VP: there is a verb *kurrijarra* 'see' plus an object expressed by a noun phrase *warirrana dangkawalathina* 'no people'. We can see it is a noun phrase due to the case agreement (M\_ABL in the glosses). If *warirrana* were a particle or a negative auxiliary constituting one syntactic unit with a verb, we would be more inclined to consider this as an instance of periphrasis. This example shows the importance of defining periphrasis in terms of functional syntax: it is the fact that non-synthetic negation in Kayardild is computed on the basis of the semantics of the object noun phrase that prevents us from considering it as anywhere near the canonical definition of periphrasis.

Languages have other constructions that typically deviate from canonical syntax and closely resemble periphrasis. One example is serial verb constructions, where we have two (or more) verb forms that function as a single predicate and describe what is conceptualized as a single event (see, for instance, the contributions to Aikhenvald and Dixon (eds) 2006 for recent discussion). The property of combining a function word with a lexical word, however, helps distinguish (canonical) periphrasis from a serial verb construction, because a serial verb construction is expressed by means of two (or more) independent lexical verbs (otherwise we would want to call one of them an auxiliary verb, i.e. a function word, and we would have a good candidate for a periphrasis). Typically, serial verb constructions do not express a grammatical feature, which is another way of distinguishing between them and periphrases.

Slightly more tricky is the case of light verb constructions, such as *do a dance*, *have a bath*, *render assistance*, or Japanese *benkyoo/tenisu suru* 'to study/play tennis'. The light verb is certainly a non-canonical lexeme because it has no semantics. For instance, *to do a dance* simply means the same as *to dance* (whereas *to perform a dance* is lexically opposed to an expression such as *to choreograph a dance*). On the other hand, a light verb is not really a function word because it does not serve to realize a grammatical property or opposition. In effect, a light verb is a part of a multiword lexeme, an instance of derivational periphrasis, which we have set aside for the purposes of this paper.

The aspectual verbs found in many languages of South Asia are sometimes referred to as light verbs, though they have a different function from the examples just given. In these constructions, a verb with a primary meaning such as 'give', 'take', 'hit', 'stand' and so on may be combined with a lexical verb to convey properties such as (a)telicity, iterativity, habituality, mirativity and many others. Here we are dealing with constructions which might make very good candidates for periphrasis, provided that the meanings expressed are sufficiently grammaticalized to be incorporated into a set of obligatory oppositions.

## 6.1 Recursion

By ‘recursion’ we mean nothing more than the following: given a phrase  $W$  consisting of a word  $x$  and a head  $w$ ,  $[x w]$ , this phrasal element can itself function as the head of a phrase of the form  $[y [x w]]$  or as the non-head of a phrase of the form  $[ [x w] z]$  (where linear order is irrelevant throughout). The exact details will, of course, depend on the theory of phrase structure employed. Neither Ackerman and Stump (2004) nor Ackerman and Webelhuth (1998) discuss the question of recursion of periphrases. The matter is raised in Spencer (2001). That paper argues for a syntactic perspective on periphrasis which is somewhat different from the morphological perspective presented in Ackerman and Stump (2004), in the sense that the syntax is left the option of generating the construction underlying the periphrasis and a rule of referral refers directly to that construction. In that respect it is not really the morphology that provides the periphrasis. The Ackerman and Stump (2004) model is different, in that it is the realization rules themselves that generate the periphrastic construction. Recursion is a property linked with syntax and so its presence in periphrastic structures is an indication that periphrasis can be both syntactic and morphological at once. While this is an important point from the perspective of Canonical Typology, it also raises the theoretical question of how, while allowing for its morphological nature, we can account for the fact that the periphrastic construction will often share many if not all the properties of a regularly formed syntactic construction (see, for instance, the Sadler and Spencer (2001) analysis of the Latin periphrasis).

Languages differ in terms of how much recursion is allowed in periphrastic forms. As we saw in (5) the Russian imperfective future is expressed by the auxiliary *byť* ‘be’ plus the infinitive. We give another example in (15).

- (15)            [w            x]  
           on   bude-t        čita-t’  
           he   be.FUT-3SG   read-INF  
           ‘He will read.’

However, if the future of the verb *byť* is required, the form is just *budet* instead of an expected *\*budet byť*, as in (16).

- (16)            [w            x]  
           on   bude-t        sčastliv.  
           he   be.FUT-3SG   happy[M.SG]  
           ‘He will be happy.’

Thus here even the most minimal recursion is excluded.

In Bulgarian, in contrast, we can find structures which involve recursion in periphrastic constructions. The perfect series of tenses is formed by combining the present tense or the imperfect past tense of the auxiliary verb *săm* ‘be’ with the l-participle form of the lexical verb.

- (17) a. *az sām čel*  
 I be.PRS.SG.1 read.L\_PART  
 'I have read.' (Present perfect)
- b. *toj e čel*  
 he be.PRS.SG.3 read.L\_PART  
 'he has read.' (Present perfect)
- c. *az bjax čel*  
 I be.PST.SG.1 read.L\_PART  
 'I had read.' (Past perfect/pluperfect)
- d. *toj beše čel*  
 he be.PST.SG.3 read.L\_PART  
 'He had read.' (Past perfect/pluperfect)<sup>8</sup>

When used as the copular verb the auxiliary in Bulgarian inflects in the normal way. This contrasts with the Russian future tense examples in (16) where a periphrastic form of the verb 'to be' is ungrammatical.

Bulgarian also has a set of evidential constructions, called 'renarrated mood' (*preizkazno naklonenie*) (see Scatton 1984, pp. 330–332). These evidentials are formed in much the same way as the perfect indicative constructions in (17) by combining *sām* with the 1-participle. Indeed, for all but the 3rd person forms the perfect indicative and the corresponding renarrated form are homophonous, as illustrated in (18a). In the third person renarrated form the auxiliary is dropped, providing a contrast between (18b) and (17b).

- (18) a. *az sām čel/bil*  
 I be.PRS.SG.1 read.L\_PART/be.L\_PART  
 'I read/was (reportedly).' (Renarrated)
- b. *toj čel/bil*  
 he read.L\_PART/be.L\_PART  
 'He read/was (reportedly)'. (Renarrated)

The Bulgarian constructions with the 1-participle exhibit recursion in several ways. First, we can have the renarrated mood of the perfect series. To form these we take the auxiliary verb and put it into the renarrated form. Thus, the corresponding renarrated mood of (17a) or (17c) replaces the auxiliary with its periphrastic renarrated counterpart, as in (19).

- (19) *az sām bil čel*  
 I be.PRS.SG.1 be.L\_PART read.L\_PART  
 'I have/had read (reportedly).'

Schematically, the structure of these expressions is as shown in (20).

<sup>8</sup> The present perfect is traditionally called the 'past indefinite', while the past perfect/pluperfect is referred to as the 'past anterior'.

- (20) a. Perfect = L-Participle + Auxiliary (present or past).  
 b. Renarrated Past (Aorist) = L-Participle + Auxiliary (present or past).  
 c. Renarrated Perfect = L-Participle + Renarrated Past of Auxiliary.<sup>9</sup>

In other words, we have a 'nested' periphrasis, in which the renarrated mood is expressed by means of a periphrastic form of the perfect auxiliary.

Bulgarian is celebrated for a further elaboration of the renarrated mood, so-called emphatic or double renarration. In this construction the auxiliary verb of the renarrated mood is further renarrated. Thus, the present and imperfect indicative forms 'I am/was reading' are *četa/četjax*. The renarrated form for the present and the imperfect is formed using the imperfective l-participle, as in (21).

- (21) *četjal*                      *săm*  
 read.IPFV.L\_PART      be.PRS.SG.1  
 'I am/was reading (reportedly).' (Renarrated present or imperfect)

The form in (21) can be further renarrated, by putting the auxiliary *săm* into the renarrated form *bil săm*, to give (22).

- (22) *bil*                      *săm*                      *četjal*  
 be.L\_PART              be.PRS.SG.1              read.IPFV.L\_PART  
 'I am/was reading (reportedly).' (Emphatic renarration)

Schematically, this gives us the structures in (23).

- (23) a. Renarrated Present/Imperfect = Imperfective L-Participle + Auxiliary (present)  
 b. Emphatic renarrated Present/Imperfect = Imperfective L-Participle + Renarrated Auxiliary

The emphatic renarrated form of the Aorist (past tense) form is likewise formed by taking the renarrated Aorist and renarrating its auxiliary, as in (24).

- (24) Emphatic renarrated Past(Aorist) = L-Participle + Renarrated Auxiliary (present)

This is homophonous with the renarrated form of the perfect in (19), schematically represented in (20c). Similar data are discussed in Bonami and Samvelian (2009), Popova and Spencer (forthcoming) and Spencer (2003) amongst others.

Whatever view is taken of periphrases like the ones we discuss here, the grammatical model needs to be able to deal also with complex periphrases like the perfect of the passive, or the progressive of the perfect, etc. (where perfect, passive and progressive are all periphrastic in a given language). A syntactic view of such complex periphrases might call upon syntactic recursion, whereas a morphological view will have to see such complex structures as the result of the cross-categorization of different morphological (morphosyntactic and/or morphosemantic)

<sup>9</sup> In general, the word order doesn't reflect the implied syntactic constituent structure because the present tense forms of *săm* are second position clitics.

features. As we argue, both should in principle be compatible, because canonical periphrasis can be both canonical syntax and canonical morphology.

## 6.2 Headedness

As we said above, if a construction is a canonical periphrastic construction we will expect that it will fill in a cell of the paradigm of a lexeme at the same time as having canonical syntax. In our understanding, in canonical syntactic constructions one word form can be singled out as the head: both in terms of determining the syntactic category of the whole and in terms of being the morphosyntactic locus of the constructions, i.e. having those morphosyntactic categories that are associated with the whole construction (the notion head is discussed in Corbett et al. 1993; Hudson 1987; Matthews 2007; Zwicky 1985).

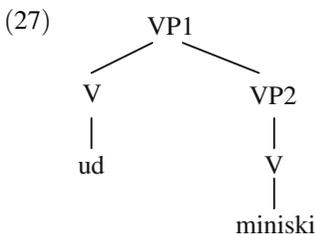
In the first discussion of periphrastic constructions within a realizational morphology, Stump (2001, pp. 231–235) when analyzing the periphrastic future paradigm of Sanskrit  $\text{D}\bar{\text{a}}$  'give' does not deal with headedness. In a later work in the framework of realizational morphology, heads appear in the principle that realizes periphrastic constructions (25b) which is presented together with the principle which defines headedness for synthetic constructions (25a). These are repeated from Ackerman and Stump 2004, p. 122, ex. 11 in the original source). (The notation  $\langle L, \sigma \rangle$  is used to refer to a lexeme,  $L$ , and the associated morphosyntax associated with the realization ( $\sigma$ ).

- (25) a. Synthetic Realization Principle  $\langle \dots \rangle$ :  
 Where the realization  $w$  of  $\langle L, \sigma \rangle$  is a synthetic member of category  $X$ ,  $w$  may be inserted as the head of  $XP$
- b. Periphrastic Realization Principle:  
 Where the realization of  $w_1 w_2$  of  $\langle L, \sigma \rangle$  is periphrastic and  $w_1 w_2$  belong to the respective categories  $X$  and  $Y$ ,  $w_1$  and  $w_2$  may be inserted as the heads of the respective phrases  $XP$  and  $YP$

Ackerman and Stump (2004) discuss examples like the one from Udmurt in (26) below (example 113 in the original source).

- (26) ton ud miniski  
 you not.2sg go  
 'You are not going.'

Presumed structure:



Here ‘head’ seems to mean simply ‘syntactic terminal’. This is under the assumption that all terminals are heads of some phrase or other, which is disputed by Toivonen (2003), who argues for ‘non-projecting words’.

In the same paper, in a discussion of Western Mari periphrasis, Ackerman and Stump (2004, p. 136) offer the following as one of the rules that realize Western Mari verbs with negative polarity:

- (28)  $RR_I$ , {POL: negative},  $v(\langle X, \sigma \rangle) = \langle \underline{Y}, Z \rangle, \sigma \rangle$ ,  
 where Y is the realization of  $\langle AK, \sigma \rangle$ ,  
 $Nar_I(\langle X, \sigma' \rangle) = \langle Z, \sigma' \rangle$ , and  $\sigma' = \sigma / \{POL: aff, TNS: 1^{st} \text{ past}\}$

This rule specifies that, where a verbal lexeme X is paired with a set of morpho-syntactic properties ( $v(\langle X, \sigma \rangle)$ ) which contains the property *pol*(arity) with value *neg* ({POL: negative}) the form of this verbal lexeme is periphrastic ( $\langle \underline{Y}, Z \rangle, \sigma \rangle$ ), with the head being a form of the negative auxiliary lexeme AK ( $\langle AK, \sigma \rangle$ ). The set of morphosyntactic properties  $\sigma$  appropriate for the verbal lexeme X are realized by an extension of  $\sigma$  which includes the specification that polarity is affirmative and the tense is first person past ( $Nar_I(\langle X, \sigma' \rangle) = \langle Z, \sigma' \rangle$ , and  $\sigma' = \sigma / \{POL: aff, TNS: 1^{st} \text{ past}\}$ ). The underlining represents the head of the periphrase (i.e. ‘head’ in a different sense from that of (25b)). This rule implicitly suggests that the head (Y) has the same set of morphosyntactic properties as those associated with the whole verbal lexeme. In other words, the notion of head that seems to be at stake here is the one that Zwicky (1985) defines as ‘morphosyntactic locus’, i.e. that element in a phrase that carries the morphosyntactic properties associated with the whole.

Further in the paper, Ackerman and Stump (2004, p. 136) introduce the following (language specific) rule of periphrastic syntax (example 20 in the original source). (Here  $XP[\sigma]$  refers to the c-structure phrase associated with the morphosyntactic properties realized by the paradigm cell  $\langle R, \sigma \rangle$ .)

- (29) Rule of periphrastic syntax:  
 Where  $\underline{Y}, Z$  (or  $Z, \underline{Y}$ ) is the realization of a cell  $\langle R, \sigma \rangle$  in a form-paradigm such that R belongs to category X, then in c-structure, Y heads  $XP[\sigma]$  and Z heads an XP complement of Y.

Here ‘head’ is taken to mean ‘structural head of a syntactic phrase’ as well as ‘locus of inflection’. Whether the negative auxiliary is a structural head will presumably depend on a detailed analysis of the syntax of Western Mari. It also depends on theoretical commitments. In LFG it is conceivable that the negative auxiliary and the lexical verb are actually co-heads (Bresnan 2001). In HPSG a variety of views have been espoused about the headedness of auxiliaries (mainly in well-studied European languages). Not all analyses treat auxiliaries as heads and lexical verbs as heading complements to those heads. The Western Mari example is similar to our hypothetical periphrastic type (ii) described in Sect. 6, in other words the Mari negative construction is of the type ‘(fully) inflecting auxiliary + converb’, in which the converb form of the lexical verb has little inflection. This is in other words an example which closely resembles our functional syntax criterion and, therefore, comes close to being canonically periphrastic.

Even in examples where auxiliaries are the locus of inflection, however, some of the properties of the construction as a whole are determined by the converb. The converb is normally the semantic head of the construction (see Anderson 2006)—in synthetic forms roots carry the semantic meaning in much the same way. The converb might indicate the syntactic patterns in which the construction as a whole fits. For example, the verb 'give' is ditransitive. The aspectual construction *be+giving* is also ditransitive. Note that Mari differs in one significant respect from Udihe (South Tungusic). Both languages have a negative auxiliary which bears most of the inflectional material of the clause. However, in Mari, one of the tenses (Past II) is expressed synthetically. This means that we can treat the negative periphrastic construction as an instance of feature intersectivity. This is not possible for Udihe, and so it is more difficult to motivate a periphrastic analysis for that language. What is important here is that we can't rely only on the nature of the construction to determine whether it is periphrastic or not. To talk about periphrasis we also need to look at the way grammatical distinctions are expressed in the language as a whole. Indeed, a related point is made by Ackerman and Stump (2004), when they argue that lexicality cannot be reliably determined on the basis of surface exponence.

It is also important to note that functional syntax is not the only option languages use in periphrasis. In many other auxiliary construction types inflectional formatives are distributed in complex ways across auxiliaries and lexical verbs.<sup>10</sup> This is true particularly when we look at negation (see Spencer ms.). When we factor in the kinds of variation described by Anderson (2006) in what constitutes an inflectional head, it becomes extremely difficult to see in what sense a periphrasis has to be headed as a universal or necessary property. If anything, lack of a clearly defined inflectional head may well be more characteristic of periphrasis—which may well mean that the typical periphrastic construction is not canonical in this respect.

### 6.3 Distributed exponence

Distributed exponence is defined as follows by Ackerman and Stump:

If the morphosyntactic property set associated with an analytical combination C has its exponents distributed among C's parts, then C is a periphrase. (Ackerman and Stump 2004, p. 147).

Ackerman and Stump (2004, p. 147) state that "(p)eriphrases often exhibit distributed exponence." Actually, we have encountered few, if any, convincing examples of this. It is certainly the case that features are often distributed across periphrastic constructions, but we doubt that this is always even a mild signal of a possible periphrasis, leave alone a sufficient condition. For instance, in the Latin perfect passive periphrasis we have distributed exponence: the auxiliary verb ('be') expresses

<sup>10</sup> A further problem arises when we have strings of auxiliaries. If the perfect and passive in English are both periphrases, then *has(n't) been written* is difficult to analyse: on the one hand *been* appears to be the head of the passive periphrasis, but on the other hand it's the head of the complement of the perfect auxiliary. It can't be the head of the periphrasis because there are two. Structurally speaking it is a head, because it takes a complement; in terms of the periphrasis as a whole we probably want to say that *has(n't)* is the head because that is the locus of tense/agreement/negation inflection.

person/number features while the passive participle expresses number/gender features. However, we can hardly appeal to this as evidence of periphrasis. The perfect passive derives from a constructionally identical syntactic formation. The perfect passive participle is an adjective and the auxiliary is homophonous with the copular verb. The features are distributed in the same way in combinations of copular verbs and predicate adjectives, and so the ‘distributed exponence’ of such constructions is completely homologous to the same pattern of exponence found in the periphrasis. Indeed, Sadler and Spencer (2001) use this fact to argue that the morphology of Latin has to have access to the syntax of the language so as to refer the morphological periphrastic construction to the normal syntactic construction. It also indicates that distributed exponence cannot even be sufficient to determine what is periphrasis.

What Ackerman and Stump have described is not a case of distributed exponence proving periphrasis but rather a case of syntactically unmotivated exponence proving periphrasis. The reason why the Udmurt negative future is striking is because it illustrates a pattern of distribution which goes against the grain of what is generally found in Uralic morphosyntax, where we would normally expect person and number to be cumulated together in a single affix. It is the fact that the construction is at variance with normal agreement syntax that is the ‘smoking gun’ suggesting periphrasis (Spencer ms., pp. 8ff). Interestingly, we can see this ‘smoking gun’ effect in a periphrastic construction in another language, though here the whole point is that the periphrasis *fails* to exhibit distributed exponence.

Spencer (2008) illustrates the notion of ‘feature intersection’ by considering negation in Japanese. The basic facts are these. Japanese verbs distinguish tense (nonpast/past), status or politeness (plain/polite) and polarity (affirmative/negative). There is no agreement of any kind in the language. The system is illustrated in Tables 8 and 9:

**Table 8** Japanese conjugation: *da* ‘be’ (copula/auxiliary)

	Plain	Polite
Affirmative		
Nonpast	da	desu
Past	dat-ta	desi-ta
Negative		
Nonpast	de wa nai	ari-mase-n
Past	de wa nakat-ta	<b>ari-mase-n desi-ta</b>

**Table 9** Japanese conjugation: *tabe-* ‘eat’

	Plain	Polite
Affirmative		
Nonpast	tabe-ru	tabe-masu
Past	tabe-ta	tabe-masi-ta
Negative		
Nonpast	tabe-nai	tabe-mase-n
Past	tabe-nakat-ta	<b>tabe-mase-n desi-ta</b>

Only two marked features can appear on a Japanese verb form. When we need to express three marked features there is therefore a problem, which is solved by means of a periphrasis. As can be seen from the table, the properties Negative, Polite, Past are expressed by taking the Negative Polite form of the lexical verb and the Polite Past (Affirmative) form of the auxiliary. The auxiliary itself forms its Negative Polite Past by using a suppletive stem. The point of this example is that the properties 'Polite' and 'Past' are expressed redundantly within the same construction. Now, extended or multiple exponence of this kind, that is, the exact opposite of distributed exponence, is hardly surprising from a cross-linguistic perspective. Indeed, Ackerman and Stump argue that such exponence is characteristic of normal, non-periphrastic syntax. But in Japanese grammar such multiple exponence is unique. Spencer (ms.) argues that it is that uniqueness which serves as an additional indication that we are dealing with a periphrasis.

Since multiple exponence is a key point in previous discussion on periphrasis it is worth considering it further. Consider the case of negation in Nenets (Gregory Stump, p.c.). Schematically, the example has the following structure: there are four tense paradigms, Present, Past, Future and Habitual. There is a negation auxiliary, Aux, with three forms, Present, Past, and Habitual, and three connegative converbs (Cvb) of a lexical verb: Bare, Future, Habitual. The Habitual will play no role in our argument, so we will ignore those forms. Thus, we obtain the paradigm schema in (30).

(30)	Pres	Past	Fut
	Aux1	Aux2	Aux1
	Cvb-bare	Cvb-bare	Cvb-FUT

We can assume the following feature sets: [ $\pm$ Past,  $\pm$ Fut]. These generate the following paradigm, where 'Subj' refers to a type of irrealis tense/mood form (Irina Nikolaeva, p.c.):

Pres	[−Past, −Fut]
Past	[+Past, −Fut]
Fut	[−Past, +Fut]
Subj	[+Past, +Fut]

The two Aux and Cvb forms are associated with the following feature specifications, where '0' means 'not specified for'. (We are not introducing an extra value here: rather the feature could simply be omitted.)

Aux1	[−Past, 0Fut]
Aux2	[+Past, 0Fut]
Cvb1 (=Cvb-bare)	[0Past, −Fut]
Cvb2 (=Cvb-FUT)	[0Past, +Fut]

Thus, exponence is distributed in this construction in the sense that the auxiliaries distinguish only values of the [Past] feature while the connegative forms distinguish only values of the [Fut] feature.

There is no reason to expect that such patterning will be restricted to negation constructions. One can easily imagine a language in which there is a Present/Past distinction and a Perfect/Non-Perfect distinction, and all forms of the paradigm are expressed by means of an auxiliary (say BE), as in (31) and (32).

(31)		NonPf	Pf
	Pres	is going	is gone
	Past	was going	was gone

Then we would have the paradigms:

(32)	Pres NonPf	Past NonPf	PresPf	PastPf
	is going	was going	is gone	was gone

Here, *is/was* play the role of Aux1 and Aux2 and *going/gone* play the role of Cvb1 and Cvb2 in the following way:

Aux1	[-Past, 0Perf]	is
Aux2	[+Past, 0Perf]	was
Cvb1	[0Past, -Perf]	going
Cvb2	[0Past, +Perf]	gone

The example is exactly homologous to Nenets except that all four possibilities are realized rather than just three. However, we suggest that the real reason why these constructions look like periphrases is not because exponence is distributed but because the exponents are realizations of grammatical categories such as negation, tense, aspect and because all but one component is clearly a function word. We can construct a hypothetical example which is exactly homologous to the fictitious examples (31)–(32) but in which the grammatical properties are distributed across two lexical items. Consider a language in which nouns inflect for various features, including number, but not definiteness, and in which adjectives inflect for just definiteness but not number. This gives us a paradigm along the lines of (33):

(33)	–def	+def	–def	+def
	–pl	–pl	+pl	+pl
	tall	tall.DEF	tall	tall.DEF
	tree	tree	tree.PL	tree.PL
	‘a tall tree’	‘the tall tree’	‘tall trees’	‘the tall trees’

Clearly, this patterning is homologous to (but more general than) the Nenets negation paradigm. It exhibits exactly the same kind of distributed exponence, but distributed over full lexical items. The problem is that there is no item whose paradigm combines the two features; rather we have different features realized in different places.

Assuming that an example such as (33) is not completely impossible in a human language this means that distributed exponence cannot be a sufficient criterion for periphrasis. The reason that the Nenets looks like a good example of periphrasis is because it involves precisely the sort of function word that typically takes part in periphrasis. It is difficult to reproduce the kind of example shown in (33) for the VP or clausal domain simply because it is difficult to construct a sequence of lexical verbs which simultaneously belong to one clause (by definition, usually, two lexical verbs means two clauses). Perhaps, however, one could look to serial verb constructions for parallel instances of distributed exponence in the absence of periphrasis.

## 7 The heterogeneous nature of periphrastic constructions

We have given here a view of periphrasis as the interaction of two sets of criteria: canonical morphology and canonical (functional) syntax. Canonical periphrasis should simultaneously possess canonical morphological properties and canonical syntactic properties. Even if it should turn out that these demands were impossible to reconcile and that canonically periphrastic constructions practically did not exist, this would not be a problem. We have the canonical point from which to calibrate the examples we find. Given our four criteria (listed again for convenience in (34) below) we can ask to what extent various kinds of construction conform to these criteria. We will consider constructed examples, because we are dealing with a logical space of possibilities, rather than trying to reach decisions about specific constructions in specific languages. Whether our hypothetical examples are attested or not, or even whether they could ever be attested is not relevant to the logic of the enterprise.

(34) Criteria for canonical periphrasis (repeated here for convenience)

- (i) a periphrastic construction is a canonical functional syntactic construction
- (ii) a periphrastic construction realizes a (canonical) grammatical feature
- (iii) a periphrastic construction (like canonical syntax and canonical morphology) will exhibit a transparent relation between form and meaning
- (iv) a periphrastic construction will occupy a cell in an otherwise inflected paradigm

For the reader's convenience we will imagine that a language grammaticalizes two tense properties, with the labels past/present tense and present/past perfect tense (the reason for this choice lies in the way the two will interact). Thus, we have translation equivalents of (35):

- (35)
- |       |                        |                |
|-------|------------------------|----------------|
| (i)   | present (non-perfect): | Mary sleeps    |
| (ii)  | past (non-perfect):    | Mary slept     |
| (iii) | present perfect:       | Mary has slept |
| (iv)  | past perfect:          | Mary had slept |

Now let the perfect tenses be parallel to the non-perfect tenses, in the sense that the same morphosyntactic relation holds between present ~ past in the perfect as in the

non-perfect (as in English, in fact). That is, the present perfect has to be regarded as the present tense form of the perfect and the past perfect (pluperfect) has to be regarded as the past tense form of the perfect. We therefore begin by assuming a feature set as (36) below:

- (36) [Perfect:{yes:[Tense:{present, past}], no:[Tense:{present, past}]}

The examples above therefore have the following feature characterizations:

- (37) (i) present non-perfect: [Perfect:no:[Tense:present]]  
 (ii) past non-perfect: [Perfect:no:[Tense:past]]  
 (iii) present perfect: [Perfect:yes:[Tense:present]]  
 (iv) past perfect: [Perfect:yes:[Tense:past]]

(Note that our hypothetical constructions already deviate from canonicity in that they express semantically interpretable features and not contextual inflection. We return shortly to the question of what the feature system looks like.)

If the four feature-value pairings were expressed by completely regular morphology then we might find in a language the suffix system in (38), where 'X' stands for the verb root/stem:

- (38) (i) [Perfect:no:[Tense:present]] X-hi-du  
 (ii) [Perfect:no:[Tense:past]] X-hi-di  
 (iii) [Perfect:yes:[Tense:present]] X-ha-du  
 (iv) [Perfect:yes:[Tense:past]] X-ha-di

Clearly, each of the four pairings occupies a cell in a morphologically inflected paradigm, and to this extent they each appear to respect (34iv). However, what this example illustrates is that the periphrasis criteria (34i) and (34iv) are required to occur together for a periphrasis to be recognizable: (34iv) only makes sense to the extent that at least one cell is occupied by a construction that satisfies (34i), and this is the force of the 'otherwise' in our characterization of (34iv). Indeed, this is the whole point of treating periphrasis as special (distinct from 'pure' syntax and 'pure' morphology). Example (38) is so far from respecting any of the periphrasis criteria that it would be perverse to call it periphrastic. (We will return to a modified version of this example, however.)

In the best case we would find that three of the cells of the paradigm were expressed in a purely inflectional (morphological) fashion and the fourth in a syntactic fashion.

- (39) (i) [Perfect:no:[Tense:present]] X-hi-du  
 (ii) [Perfect:no:[Tense:past]] X-hi-di  
 (iii) Perfect:yes:[Tense:present]] X-ha-du  
 (iv) [Perfect:yes:[Tense:past]] hav don X

Ideally, we would expect the syntactic construction to align with the canonical ends of criteria 5–7 (for canonical syntax). On the other hand, their semantics are likely to be non-canonical from the point of view of criteria 8–9. For example, it would

normally be the case that only X (in 39iv a free form rather than a root) bears a canonical lexical meaning. The rest of the elements in a periphrastic construction are likely to be function words and therefore non-canonical. These function words may behave like members of one of the major lexical classes (for example auxiliaries may behave syntactically like the rest of the verbs in the languages), or they might have exceptional syntactic properties (somewhat like the auxiliaries in English). Constructions will also vary depending on whether they exhibit feature clash or idiomaticity or not. The elements of the construction may occur elsewhere in the language, or one or more of them might occur only in the periphrasis. In other words, real world periphrastic construction that are clearly part of an otherwise morphological paradigm can be discussed in terms of the degree to which they depart from the criteria for being canonical syntax (composed of canonical words). A more important issue that arises in this scenario is whether the periphrastic construction occurs only within the morphological paradigm, or is a more general construction that happens to express a grammatical meaning that coincides with the meaning of the rest of the paradigm. The account of periphrasis in individual languages might be therefore quite complex, however, if there is feature intersection, there is a case to be made for periphrasis.

Now let us consider a set of situations in which there is no feature intersection, and so there is much less certainty about whether we are dealing with a periphrasis. We will construct a variety of hypothetical language types and evaluate the extent to which each type can be considered canonically periphrastic according to our criteria. Note that we will not be proposing any metric, we are simply illustrating the applicability of the criteria to various construction types. Consider a language in which the features are expressed by means of grammatical words which are represented as syntactic terminals, as seen in (40), where 'V' stands for the bare verb:

- (40) (i) [Perfect:no:[Tense:present]] du hi V  
 (ii) [Perfect:no:[Tense:past]] di hi V  
 (iii) [Perfect:yes:[Tense:present]] du ha V  
 (iv) [Perfect:yes:[Tense:past]] di ha V

Let's assume that the syntactic construction exemplified here is more-or-less canonical. Here we have feature values expressed by syntax but no intersection with morphologically expressed features (m-features). Is this construction type periphrastic? To the extent that the features expressed are canonical grammatical features we would have to say 'yes', though this construction type is not particularly close to the canonical type. What this means in practice is that such a system might very well be described as periphrastic by descriptive linguists who have a wide interpretation for the term 'periphrasis'.

Now let's modify the example slightly. So far we have assumed a system in which all values of all features receive distinct, overt expression. Both for morphology and for syntax such systems, in which there is a transparent correspondence between form and meaning, are more canonical. Suppose now that the [Perfect:no] and [Tense:present] feature values are expressed by zero exponence. This means that our syntactic construction in (40) will now look like (41):

- (41) (i) [Perfect:no:[Tense:present]] V  
 (ii) [Perfect:no:[Tense:past]] di V  
 (iii) [Perfect:yes:[Tense:present]] ha V  
 (iv) [Perfect:yes:[Tense:past]] di ha V

We can now easily speak about a default value of each feature. We would claim that the construction system represented in (41) is a (slightly) better instance of a periphrasis than that illustrated in (40). The reason for this is that a syntactic construction is now in opposition to a word form, rather than another syntactic construction. The interpretation of the word form and of the syntactic construction becomes dependent on this contrast, i.e. we have the beginnings of a paradigmatic contrast.

Consider now a slightly different system from (40) and (41), that of (42):

- (42) (i) [Perfect:no] V  
 (ii) [Perfect:yes:[Tense:present]] ha V  
 (iii) [Perfect:yes:[Tense:past]] di ha V

Here, there is no non-perfect past tense form. Let us suppose that the present perfect is vague as to a simple past time reference (preterite interpretation) and past time with current relevance (perfect interpretation). However, let us suppose that the extra past perfect tense form is used roughly like the English pluperfect. In this kind of system we would simply have to complicate the feature system in a way that destroyed the symmetry of the system in (40/41). But the system would be no more or less periphrastic.

Let us now complicate this example a little further (revealing the point of our choice of hypothetical features). Suppose that the language lacks a straightforward expression of [Tense:past] and that its present perfect forms are ambiguous as translation equivalents between the English present perfect and the English simple past. Suppose that there is still an opposition between present perfect and past perfect, but that this is expressed by taking the perfect auxiliary and applying it to the perfect auxiliary. In other words we have the set of forms shown in (43)

- (43) (i) [Perfect:no:[Tense:present]] V  
 (ii) [Perfect:yes:[Tense:present]] ha V  
 (iii) [Perfect:yes:[Tense:past]] ha ha V

Moreover, let's assume that there are morphosyntactic and morphosemantic contexts which distinguish the past tense interpretations from the non-past interpretation, such that the past tense interpretations always pattern together and the present tense interpretations always pattern together. This would naturally lead us to set up a rather more complex paradigm, of the form (44):

- (44) (i) [Perfect:no:[Tense:present]] V  
 (ii) [Perfect:no:[Tense:past]] ha V  
 (iii) [Perfect:yes:[Tense:present]] ha V  
 (iv) [Perfect:yes:[Tense:past]] ha ha V

The simple past and present perfect are always syncretic in this language but the [Tense] feature is warranted by the role it plays in accounting for the present~past

opposition in the perfect series. (We assume that there is no way of semantically interpreting the notion 'perfect of a perfect', and certainly no way of interpreting such a beast which would make it the translation equivalent of simple past in any language.)

Is the system in (44) a periphrasis? Note that there is no morphological feature intersection (we are still assuming that the hypothetical language lacks morphology). Nonetheless, there is a clear intuition that (44) is periphrastic in a sense that (40/41) or (42) aren't. What is the origin of this intuition? In (44) we have a tense opposition which is lacking in (43), but that opposition receives no independent expression. It comes about by virtue of manipulating the combinatorics of the existing auxiliary system. In this respect the system is less canonically syntactic than any of the previous examples. Does this mean that it is closer to being a canonical periphrasis? We would argue that it isn't, precisely because it is less canonical syntactically. The reason why the system in (44) 'feels' (to some linguists) as though it is a periphrase is because periphrasis represents a tension between 'pure' syntax and 'pure' morphology. If we encounter a system that can be handled by independently motivated principles of syntax we have no motivation to call it anything special, for instance, a periphrase. Hence, by being less canonically syntactic a construction type such as (44) looks as though it is different and from this the conclusion might be drawn that it represents a periphrasis. In practice, for individual languages this may well be a sound conclusion, of course, depending on the aims of the grammatical description and so on, but it isn't a conclusion that follows from Canonical Typology. The best we can say about it, *ceteris paribus*, is that it's a less than canonical syntactic construction (less canonical, say, than (40)). This conclusion should come as no surprise. There are many situations in which non-canonical syntactic behaviour is nothing to do with periphrasis. For instance, consider a language in which attributive adjectives always come after the noun they modify, but come before the noun when they have certain types of metaphorical interpretation. We would not wish to use this type of syntactic non-canonicity as evidence of a periphrastic construction.<sup>11</sup>

<sup>11</sup> The situation described in (43/44) is, of course, modelled on attested constructions such as the French *passé composé*. However, there is a crucial difference between our hypothetical example and the French construction (which genuinely is a periphrasis). In French, verbs are morphologically marked for tense, present vs. imperfect: *mange* 'eats', *mangeait* 'ate, was eating', *a* 'has', *avait* 'had, used to have'. This opposition is maintained in the perfect tense series: *a mangé* 'has eaten', *avait mangé* 'had eaten'. So far there is no evidence of (canonical) periphrasis. Now, the present perfect *a mangé* can also be given a simple past interpretation, the so-called *passé composé*. This might be taken to illustrate nothing more than vagueness of time reference for the present perfect (and is so analysed by some linguists, e.g. Vet 2007) but this isn't quite right: there is a semantic opposition between the imperfect *mangeait* 'was eating' and the *passé composé*, *a mangé*, in the sense that the imperfect lacks a simple preterite interpretation, and the *passé composé* lacks the durative, habitual etc. interpretations associated with the imperfect. Interestingly, in the past perfect *avait mangé* the imperfect form of the auxiliary isn't interpreted with imperfect semantics. This form is vague between translation equivalents 'had eaten' and 'had been eating'. The crucial fact about French is the existence of a *passé surcomposé a eu mangé*. This is formed by taking the present perfect form of the perfect auxiliary. The interpretation is that of a perfect in the (simple) past, i.e. overlapping with some of the meanings of *avait mangé*. This means that *a mangé* really does realize a [Tense:past] feature value but that the [Tense] feature is expressed indirectly, as in example (43/44) above. The crucial difference with (43/44), however, is that the French auxiliaries are marked morphologically for non-past tense. Therefore, the simple past (*passé composé*) interpretation of *a mangé* is opposed to the morphological present tense form *mange*. This means there is intersection with a morphologically expressed feature (though of a somewhat complex kind).

## 8 Conclusion

Periphrasis is an important phenomenon because of the place it occupies at the morphology-syntax interface. It is often problematic for linguistic models since they have to accommodate not only straightforwardly morphological and straightforwardly syntactic phenomena but also phenomena that seem to share properties of both. Our intention here, however, has been not to show how periphrasis can be accommodated in a particular linguistic model. Our aim, rather, has been to give a clear account of this 'mixed' nature of periphrasis and to provide a basis for the description of actual facts in different languages that despite being previously labelled as instances of periphrasis might actually be quite different.

From the viewpoint of canonical typology the paper thus represents a new departure. Previous analyses using this approach have typically set up converging criteria for a phenomenon located within a single component. Periphrasis stands at the intersection of two components, syntax (especially functional syntax) and (inflectional) morphology, each with their canonical properties. We start with the fundamental insight in the literature that periphrasis is syntax where we expect to find morphology. We set up canonical criteria for syntax and canonical criteria for (inflectional) morphology and show that periphrasis is at the intersection of these two. In real life it will be difficult to find linguistic phenomena which are canonically syntactic and canonically morphological at the same time. We think that the closest periphrasis will come to this is canonical functional syntax, i.e. it is canonical syntax with non-canonical semantic interpretation which can be found in a morphological paradigm. Our confidence that the definitions are the right ones is based on the fact that they are independently required to define canonical syntax and morphology.

Justifying our criteria for periphrasis on the basis that they are required to define canonical syntax and morphology enables us to be more principled about the variation we find among periphrastic constructions in examples coming from various natural languages. This proved a key feature of the paper. For instance, the phenomena we might wish to call periphrastic, far from conforming to some ideal of functional syntax, can be very different in terms of their syntactic properties. Indeed, recent work in periphrasis (Bonami and Webelhuth forthcoming) shows just that. Furthermore, the relation between form and meaning in periphrastic constructions can show different departures from transparency. Such departures are often taken to be the hallmark of periphrasis: for Ackerman and Stump (2004) they constitute a sufficient condition for periphrasis. However, for us they are part of the space of variation but are not criterial (Sects. 5.3, 6.3). Next, periphrastic constructions might fit into morphological paradigms in different ways. They might occupy a cell in a paradigm in an ideal feature intersection, or be a sub-paradigm or a whole paradigm (see our discussion in Sect. 5.2). Or indeed all we might want to say about a periphrastic construction (a very non-canonical one in our terms) is that it is not part of the paradigmatic organization in a language but expresses a grammatical feature (which itself might be canonical or non-canonical). We discussed this aspect of periphrasis in Sects. 5.1 and 5.2.

Our canonical approach has allowed us to stake out a theoretical space, within which the interesting examples of periphrasis can be located. It also has allowed us to engage with earlier helpful work on definitions, including Haspelmath (2000) and Ackerman and Stump (2004). The topic is one of continuing interest, and we believe that definitional work is a key contribution to progress in this area.

**Acknowledgments** The research reported here was supported by ESRC grant number RES-062-23-0696 and ERC grant number ERC-2008-AdG-230268 MORPHOLOGY. This support is gratefully acknowledged. We are grateful to Nicholas Evans, Roger Evans and Gregory Stump for their useful comments. We would also like to thank Ingo Plag and the anonymous reviewers for their helpful suggestions. Any errors are, of course, our responsibility.

## References

- Ackerman, F. (2000). *Lexical constructions: Paradigms and periphrastic expressions*. Paper read at the LFG workshop on Morphosyntax.
- Ackerman, F., Stump G., & Webelhuth G. (2011). Lexicalism, periphrasis and implicative morphology. In R. Borsley & K. Börjars (Eds.), *Non-transformational syntax: Formal and explicit models of grammar*. Oxford: Blackwell.
- Ackerman, F., & Stump, G. T. (2004). Paradigms and periphrastic expression: A study in realization-based lexicalism. In L. Sadler & A. Spencer (Eds.), *Projecting morphology* (pp. 111–158). Stanford, CA: Center for the Study of Language and Information.
- Ackerman, F., & Webelhuth, G. (1998). *A theory of predicates*. Stanford, CA: Center for the Study of Language and Information.
- Aikhenvald, A., Yu., & Dixon, R. M. W. (Eds.). (2006). *Serial verb constructions: A cross-linguistic typology*. Oxford: Oxford University Press.
- Anderson, G. D. S. (2006). *Auxiliary verb constructions*. Oxford: Oxford University Press.
- Bach, E. (1988). Categorical grammars as theories of language. In R. Oehrle, E. Bach, & D. Wheeler (Eds.), *Categorical grammars and natural language structures* (pp. 17–34). Dordrecht: D. Reidel.
- Baerman, M., Brown, D., & Corbett, G. G. (2005). *The syntax–morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.
- Bonami, O., & Samvelian, P. (2009). Inflectional periphrasis in Persian. In S. Müller (Ed.), *Proceedings of the HPSG 2009 conference* (pp. 26–46). Stanford: CSLI.
- Bonami, O., & Webelhuth, G. (forthcoming). The phrase-structural diversity of periphrasis: A lexicalist account. In M. Chumakina & G. G. Corbett (Eds.), *Periphrasis: The role of syntax and morphology in paradigms*. Proceedings of the British Academy. Oxford: British Academy/Oxford University Press.
- Booij, G. (1996). Inherent versus contextual inflection and the split morphology hypothesis. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1995* (pp. 1–15). Dordrecht: Kluwer.
- Börjars, K. (1998). *Feature distribution in Swedish noun phrases*. Oxford: Blackwell.
- Börjars, K., Vincent, N., & Chapman, C. (1997). Paradigms, pronominal inflection and periphrasis. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology* (pp. 155–180). Dordrecht: Kluwer.
- Bresnan, J. (2001). *Lexical functional syntax*. Oxford: Blackwell.
- Bybee, J. L., Perkins, R. D., & Pagliuca W. (1994). *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
- Comrie, B. (2003). When agreement gets trigger-happy. In D. Brown, G. G. Corbett, & C. Tiberius (Eds.), *Agreement: A typological perspective*. Special issue of Transactions of the Philological Society 101, no. 2 (pp. 313–337). Oxford: Blackwell.
- Corbett, G. G. (2003). Agreement: the range of the phenomenon and the principles of the Surrey Database of Agreement. In D. Brown, G. G. Corbett, & C. Tiberius (Eds.), *Agreement: A typological perspective*. Special issue of Transactions of the Philological Society 101 no. 2 (pp. 155–202). Oxford: Blackwell.
- Corbett, G. G. (2006). *Agreement*. Cambridge: Cambridge University Press.
- Corbett, G. G. (2007). Canonical typology, suppletion and possible words. *Language*, 83, 8–42.

- Corbett, G. G. (2008). Determining morphosyntactic feature values: The case of case. In G. G. Corbett & M. Noonan (Eds.), *Case and grammatical relations: Studies in honor of Bernard Comrie* (pp. 1–34). Amsterdam: John Benjamins.
- Corbett, G. G. (2010a). Features: Essential notions. In A. Kibort & G. G. Corbett (Eds.), *Features: Perspectives on a key notion in linguistics* (pp. 17–36). Oxford: Oxford University Press.
- Corbett, G. G. (2010b). Canonical derivational morphology. *Word Structure*, 3, 141–155.
- Corbett, G. G. (forthcoming). Periphrasis and possible lexemes. In M. Chumakina & G. G. Corbett (Eds.), *Periphrasis: The role of syntax and morphology in paradigms*. Proceedings of the British Academy. Oxford: British Academy/Oxford University Press.
- Corbett, G. G., Fraser, N. M., & McGlashan, S. (Eds.). (1993). *Heads in grammatical theory*. Cambridge: Cambridge University Press.
- Delsing, L. (1993). *The internal structure of noun phrases in the Scandinavian languages*. PhD dissertation, University of Lund.
- Enfield, N. J. (2007). *A grammar of Lao*. Berlin: Mouton de Gruyter.
- Evans, N. (1995). *A grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Evans, N. (2003). Typologies of agreement: Some problems from Kayardild. In D. Brown, G. G. Corbett, & C. Tiberius (Eds.), *Agreement: A typological perspective*. Special issue of Transactions of the Philological Society 101 no. 2 (pp. 203–234). Oxford: Blackwell.
- Everaert, M., van der Linden, E.-J., Schenk, A., & Schreuder, R. (Eds.). (1995). *Idioms. Structural and psychological perspectives*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hankamer, J., & Mikkelsen, L. (2002). A morphological analysis of definite nouns in Danish. *Journal of Germanic Linguistics*, 14, 137–175.
- Hankamer, J., & Mikkelsen, L. (2005). When movement must be blocked: A reply to Embick and Noyer. *Linguistic Inquiry*, 36, 85–125.
- Hankamer, J., & Mikkelsen, L. (2008). Definiteness marking and the structure of Danish pseudopartitives. *Journal of Linguistics*, 44, 317–346.
- Haspelmath, M. (2000). Periphrasis. In G. Booij, C. Lehmann, & J. Mugdan (Eds.), *Morphology: A handbook on inflection and word formation* (pp. 654–664). Berlin: de Gruyter.
- Heck, F., Müller, G., & Trommer, J. (2008). A phase-based approach to Scandinavian definiteness marking. In C. B. Chang & H. J. Haynie (Eds.), *Proceedings of the 26th West Coast conference on formal linguistics* (pp. 226–233). Somerville, MA: Cascadilla.
- Heine, B., & Kuteva, T. (2002). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Honda, I. (1996). *Negation: A cross-linguistic study*. PhD thesis, University of Buffalo.
- Hudson, R. A. (1987). Zwicky on heads. *Journal of Linguistics*, 23, 109–132.
- Julien, M. (2005). *Nominal phrases from a Scandinavian perspective*. Amsterdam: John Benjamins.
- Kahrel, P. (1996). *Aspects of negation*. PhD Dissertation, University of Amsterdam.
- Matthews, P. H. (1974). *Morphology*. Cambridge: Cambridge University Press.
- Matthews, P. H. (2007). *Syntactic relations: A critical survey*. Cambridge: Cambridge University Press.
- Nerbonne, J. (1996) Computational semantics—Linguistics and processing. In S. Lappin (Ed.), *The handbook of contemporary semantic theory* (pp. 461–484). Oxford: Blackwell.
- Nikolaeva, I. (forthcoming). Periphrasis in Nenets. In M. Chumakina & G. G. Corbett (Eds.), *Periphrasis: The role of syntax and morphology in paradigms*. Proceedings of the British Academy. Oxford: British Academy and Oxford University Press.
- Nikolaeva, I., & Spencer, A. (2008). Nouns as adjectives and adjectives as nouns. ms.
- Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language*, 70, 491–538.
- Plank, F. (1994). Inflection and derivation. In R. E. Asher & J. M. Y. Simpson (Eds.), *The encyclopedia of language and linguistics* (Vol. 3, pp. 1671–1678). Oxford: Pergamon Press.
- Polinsky, M. (2003). Non-canonical agreement is canonical. In D. Brown, G. G. Corbett, & C. Tiberius (Eds.), *Agreement: A typological perspective*. Special issue of Transactions of the Philological Society 101 no. 2 (pp. 279–312). Oxford: Blackwell.
- Popova, G. (2010). Features in periphrastic constructions. In A. Kibort & G. G. Corbett (Eds.), *Features: Perspectives on a key notion in linguistics* (pp. 166–184). Oxford: Oxford University Press.
- Popova, G., & Spencer, A. (forthcoming). Relatedness in periphrasis: A paradigm-based perspective. In M. Chumakina & G. G. Corbett (Eds.), *Periphrasis: The role of syntax and morphology in paradigms*. Proceedings of the British Academy. Oxford: British Academy/Oxford University Press.
- Poser, W. J. (1992). Blocking of phrasal constructions by lexical items. In I. Sag & A. Szabolcsi (Eds.), *Lexical matters* (pp. 111–130). Stanford, CA: CSLI.
- Rosch, E. (1973). Natural categories. *Cognitive Psychology*, 4, 328–335.

- Rounds, C. (2001). *Hungarian. An essential grammar*. London: Routledge.
- Sadler, L., & Spencer, A. (2001). Syntax as an exponent of morphological features. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 2000* (pp. 71–96). Dordrecht: Kluwer.
- Salminen, T. (1997). *Tundra Nenets inflection*. Mémoires de la Société Finno-Ougrienne (Vol. 227), Helsinki.
- Scatton, E. (1984). *A reference grammar of modern Bulgarian*. Columbus, OH: Slavica Publishers.
- Seifart, F. (2005). *The structure and use of shape-based noun classes in Miraña (North West Amazon)*. PhD thesis, Radboud University, Nijmegen.
- Spencer, A. (2001). The paradigm-based model of morphosyntax. *Transactions of the Philological Society*, 99, 279–313.
- Spencer, A. (2003). Periphrastic paradigms in Bulgarian. In U. Junghanns & L. Szucsich (Eds.), *Syntactic structures and morphological information* (pp. 249–282). Berlin: Mouton de Gruyter.
- Spencer, A. (2007). Extending deponency. In M. Baerman, G. G. Corbett, D. Brown, & A. Hippisley (Eds.), *Deponency and morphological mismatches* (pp. 45–70). Proceedings of the British Academy, 145. Oxford: Oxford University Press.
- Spencer, A. (2008). Negation in Japanese: A case of morphosyntactic mismatch. *Lingua*, 118, 997–1017.
- Spencer, A. (ms.). Sentence negation and periphrasis. Unpublished manuscript, University of Essex.
- Stump, G. T. (1998). Inflection. In A. Spencer & A. M. Zwicky (Eds.), *Handbook of morphology* (pp. 13–43). Oxford: Blackwell.
- Stump, G. T. (2001). *Inflectional morphology: A theory of paradigm structure*. Cambridge studies in linguistics. Cambridge: Cambridge University Press.
- Stump, G. T. (2005). Word-formation and inflectional morphology. In P. Štekauer & R. Lieber (Eds.), *Handbook of word-formation* (pp. 49–71). Dordrecht: Springer.
- Stump, G. T. (2006). Heteroclisis and paradigm linkage. *Language*, 82, 279–322.
- Stump, G. T., & Finkel, R. (2008). *Stem alternations and principal parts in French verb inflection*. Paper presented at Décembrettes 6: Colloque International de Morphologie, “Morphologie et classes flexionnelles”, December 4–5, 2008. France: Université de Bordeaux.
- Suthar, B. K. (2006). *Agreement in Gujarati*. PhD Dissertation, University of Pennsylvania.
- Thornton, A. (2008). *A non-canonical phenomenon in Italian verb morphology: Double forms realizing the same cell*. Paper read at the first Oxford workshop on Romance verb morphology, 27–28 August 2008, Oxford.
- Thornton, A. (2011). Overabundance (multiple forms realizing the same cell): A non-canonical phenomenon in Italian verb morphology. In M. Goldbach et al. (Eds.), *Morphological autonomy: Perspectives from Romance inflectional morphology* (pp. 362–385). Oxford: Oxford University Press.
- Toivonen, I. (2003). Non-projecting words: A case study of Swedish particles. In *Studies in natural language and linguistic theory* 58. Dordrecht: Kluwer.
- Vet, C. (2007). The descriptive inadequacy of Reichenbach's tense system: A new proposal. In L. de Saussure, J. Moeschler, & G. Puskas (Eds.), *Tense, Mood and aspect: Theoretical and descriptive issues* (pp. 7–26). Amsterdam: Rodopi.
- Westney, P. (1995). *Modals and periphrastics in English*. Tübingen: Max Niemeyer Verlag.
- Zwicky, A. M. (1985). Clitics and particles. *Language*, 61(2), 283–305.