Language Sciences xxx (2015) 1–13

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona

Raúl Aranovich

University of California Davis, United States

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Realizational morphology
Periphrasis
Auxiliary
Bantu
Lexical Functional Grammar
Lexical Integrity Principle

ABSTRACT

In this paper, I argue that the choice of a periphrastic over a synthetic verbal form in Shona is a strategy to avoid the realization of specific combinations of morphosyntactic properties by the same word form. I argue that periphrastic forms with *-nga* are formally equivalent to synthetic forms in the verbal paradigm of Shona, based on the assumption that inflected words are not formed in the lexicon, but are the realization of properties specified on the terminal nodes of syntactic structures. The competition between synthetic and analytic forms in the inflectional paradigm of Shona is accounted for in a model of grammar that associates a clause with a representation that is functional and non-configurational, similar to f-structure in Lexical-Functional Grammar. But instead of letting lexical items project this information onto terminal nodes, I follow Anderson's (1992) proposals closely by incorporating phrasal realizational rules to the model, the function of which is to distribute inflectional features among the constituents of the clause.

1. Introduction

In his summary of the characteristic properties of the Bantu languages, Doke points out that they "show an extraordinarily high development in the conjugation of the verb" (Doke, 1967:47). The Bantu verb typically consists of a stem preceded by a series of inflectional prefixes expressing tense, aspect, etc., the first of which is a subject agreement marker (often referred to as a concordial affix). It is also common for a stem to have suffixes for voice and other operations on argument structure. The languages in the South-Eastern and South-Central zones of Doke's (1945, 1967) surveys are also characterized by the occurrence of **deficient** (i.e. auxiliary) verbs (cf. also Mberi, 2006). In this paper, I will examine complex verbs formed with the help of the auxiliary *-nga* in Shona, a Bantu language of Zimbabwe. I will show that this auxiliary verb denotes no morphosyntactic property of its own. Rather, it is used to construct a periphrastic verbal form when specific combinations of inflectional properties prevent the formation of a synthetic form in the paradigm of the Shona verb.

To anticipate the discussion, I present a few illustrative examples here, from the 1.SG indefinite incidental (the terms are from Fortune, 1955, and are explained in detail in Section 2). The simple past form *nda-tora* 'I took' is formed with the prefix *nda-* '1.SG.PAST' attached directly to the root *tora* 'take'. Shona has a progressive prefix *-cha-*, which goes in a slot between the subject agreement marker and the root, as in *ndi-cha-tora* 'I am taking'. But the form **nda-cha-tora* does not exist. In its place, to express the past progressive, Shona employs a periphrastic form with the auxiliary *-nga* in the past, followed by a progressive participle, as in (1).

E-mail address: raranovich@ucdavis.edu.

http://dx.doi.org/10.1016/j.langsci.2015.05.004
0388-0001/

(1) *nda-nga* *ndí-cha-tora*
1.SG.PAST-AUX 1.SG.PART-PROG-take
'I was taking.'

Intuitively, the auxiliary and participle together express a complex combination of morphosyntactic properties through a division of labor: the auxiliary realizes the tense, and the participle the progressive aspect. The periphrastic form, then, is used to avoid two properties (past and progressive) from being expressed jointly in the same word form. But given the fact that the periphrastic past progressive is in a paradigmatic opposition to two synthetic forms (simple past and present progressive), the morphology of the Shona verb must have mechanisms to realize different combinations of properties either synthetically (in a single word form) or analytically (as a sequence of words). To account for the competition between synthetic and analytic forms in the inflectional paradigm of Shona, I will propose a rule of phrasal morphology that spreads inflectional features around the constituents of the clause, depending on the specific inflectional features of the predicate. I will adopt a realizational view of inflection (Anderson, 1992; Stump, 2001) in which rules of exponence specify the morpho-phonological realization of a verb's grammatical properties. But I will argue that the linguistic behavior of periphrastic verbs supports a model in which inflectional features are not projected from lexical items onto syntactic structure. Rather, they are specified as properties of terminal nodes. The paradigmatic opposition between synthetic and periphrastic verbal forms, I will suggest, is captured in a level where morphosyntactic properties are represented as functional features, in a nonconfigurational fashion. Within the framework of Lexical-Functional Grammar (Bresnan, 1982, 2001), this is the level of functional structure (f-structure).

In Shona, a periphrasis is employed to avoid having two incompatible morphosyntactic properties on the same terminal node. In this sense, the Shona periphrastic verb seems to differ from the more familiar verbal periphrasis that emerges in Latin perfect passives. A Latin form like *amatus sum* 'I have been loved' contains a participle expressing the features passive and perfect, and a BE auxiliary realizing other features of the paradigm (person and tense, for instance). Unlike the Shona case, then, the two offending features (perfect aspect and passive voice) are realized on the same terminal node (the participle). It is their joint occurrence that blocks the main verb from also realizing the other inflectional properties, requiring the addition of the auxiliary verb. Embick (2000) develops an analysis of the Latin periphrastic verb in terms of movement, within the Distributed Morphology framework (Halle and Marantz, 1993). While a comprehensive comparison between DM and the approach I am proposing is beyond the scope of this paper, a discussion of Embick's analysis of the Latin periphrastic forms in the light of the Shona data will serve to illustrate some of the differences between the two approaches, suggesting topics for further research.

The paper is organized as follows. Section 2 describes the inflectional system of the Shona verb (based on Fortune, 1955), limiting the discussion to those regions of the verbal paradigm that are relevant to understanding the distribution of the auxiliary *-nga*. Forms with *-nga* are characterized as verbal periphrases because they satisfy the Featurally Intersective Distributional Criterion (Ackerman and Stump, 2004). Section 3 introduces basic concepts of Lexical Functional Grammar (Bresnan, 1982, 2001), showing the role of f-structure in capturing paradigmatic relations. Section 4 introduces fundamental concepts of realizational morphology, and it sketches a realizational analysis of the Shona verb, introducing rules of exponence for the basic affixes along the lines of Stump (2001) and Ackerman and Stump (2004). Section 5 is where a realizational analysis of the periphrastic forms is developed. The modifications to the LFG framework that are necessary to accommodate phrasal realizational rules are explained, showing how the analysis departs from the strict adherence to lexicalism found in many monostratal models. Some differences and similarities with other approaches to the grammar of periphrastic forms (the Distributed Morphology theory of Halle and Marantz, 1993, and Embick, 2000, in particular) are summarized in Section 6. Section 7 concludes the paper, with a discussion of the relationship between syntax and inflectional morphology.

2. Elements of the Shona conjugation

Like other Bantu languages, Shona has a rich set of morphological noun classes, which is also reflected in the system of verbal agreement (Déchaine et al. 2014). Class I and II prefixes indicate person, in addition to number (class II is the plural of class I). To simplify the presentation of the material, I will limit the examples to forms with the 1.SG (class I) prefix *ndi-*, and its variants (as explained below). There are 20 different class prefixes in Shona, in addition to four prefixes for the local persons (I and II) in classes 1 and 2. After the subject marker, a number of tense/aspect prefixes can be found. There are also some inflectional affixes, and a final vowel that sometimes indicates mood.

Fortune (1955) classifies the Shona verb into two general **conjugations**: *affirmative* and *negative*. Each conjugation is further classified into **moods**: *imperative*, *indicative*, *potential*, *participial*, *relative*, *subjunctive*, and *hortative*. There are also four **tenses**: *present*, *future*, *recent past*, and *remote past*. Tenses are divided according to **aspect**, **mode**, and **implication**. The aspects are *indefinite*, *continuous*, and *perfect*. The modes are *incidental* (i.e. an aorist) and *habitual*. The implications have no direct correlation with inflectional categories commonly found in Indo-European languages, and are more easily translated with the help of adverbs. There are three implications: *simple*, *exclusive*, and *progressive*. The exclusive implication is translated with the adverb *now*, and the progressive implication with the adverb *still*. For the sake of clarity, I will focus the analysis on the indicative mood of the affirmative conjugation, using the verb *-tora* 'to take' as an example.

An examination of the indefinite simple, summarized in Table 1, reveals the basic prefixes of the Shona verb.

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

Table 1
Tenses and modes of the Indefinite Simple.

	Indefinite simple (1st person singular)	
	Incidental	Habitual
PRES	*ndi-no-tora* 'I take'	*ndi-no-tora* 'I usually take'
PAST	*nda-tora* 'I took'	*nda-i-tora* 'I used to take'
RM. P.	*nda-ka-tora* 'I took'	———
FUT	*ndi-cha-tora* 'I shall take'	———

The 1.SG prefix *ndi-* combines with *-no-* to render the incidental present form. The future tense of the incidental is expressed by the prefix *-cha-*, which follows the 1.SG prefix *ndi-*. The past tenses are characterized by a second series of subject markers with *a* as a characteristic vowel. Thus, the prefix *nda-* is immediately added to the root to form the incidental recent past. The remote past is formed with the prefix *-ka-*, immediately after the past person marker. The habitual mode has only two tenses, present and past (the future and remote past occur in the incidental only). The habitual present is syncretic with the incidental present, and the habitual past is formed with insertion of the prefix *-i-* right after the past tense subject concordial prefix.

Like the simple implication, the exclusive implication of the affirmative indicative has six tenses: four in the incidental mode, and two in the habitual. These are summarized in Table 2.

The characteristic affix of the exclusive implication is *-chi-*, added to the corresponding tenses of the incidental simple, following all other prefixes. It occurs in combination with the future prefix, and in the recent and remote past tenses. For the present exclusive, however, there is a third set of subject markers, characterized by the vowel *o*. So, in the present incidental exclusive, the first person singular present exclusive prefix *ndo-* attaches directly to the verbal stem. Unlike the tenses reviewed so far, the two tenses of the habitual mode of the exclusive implication are periphrastic, formed with the auxiliary verb *-nga* 'be' and a participle. While the auxiliary realizes tense and mood, the participle expresses the implication. The participle is identical in its segmental composition to the present incidental, but with a characteristic high tone on the personal prefix. Notice that the auxiliary is conjugated like the corresponding forms in the simple implication, even though the interpretation of the periphrasis as a whole is in the exclusive implication.

The periphrastic forms with *-nga* and participle occur in all tenses of the progressive implication, except for the incidental present. This is summarized in Table 3.

The only synthetic form of the progressive is characterized by the prefix *-cha-*, observable in the first person singular incidental present progressive. A high tone on the subject marker of this form provides the progressive participle. This participle combines with the auxiliary verb *-nga*, inflected for mode and tense, to complete the basic paradigm of the imperfect affirmative indicative. Notice again that the auxiliary is inflected like the forms for the simple implication, now also marking the distinction between the incidental and habitual modes.

It is possible now to see that the periphrastic form composed of the auxiliary *-nga* (expressing tense and mode) and a participle (expressing implication and aspect) are not limited to one particular implication, mode, or tense. The incidental is periphrastic in the past progressive, but not in the past exclusive. The exclusive is periphrastic in the habitual, but not in the incidental. The habitual is periphrastic in the exclusive and progressive, but not in the simple implication. Table 4 illustrates the distribution of synthetic and periphrastic forms, using the present and past tenses for comparison between the modes and implications.

The periphrastic forms *-nga* + PART, then, are always in paradigmatic opposition with a synthetic form in some morphosyntactic dimension. It is not possible, then, to associate a single, individual feature as the content of the periphrasis. It is this sort of "feature intersectivity" that characterizes true periphrases, according to Ackerman and Stump (2004). This observation is made explicit in (2)

(2) Featurally intersective distributional criterion: there is no single feature that is always expressed as a periphrase (Ackerman and Stump, 2004).

The periphrastic forms of the Shona verb, then, are integral to its inflectional paradigm because they are in a minimal opposition (i.e. by changing only one of its distinctive morphosyntactic features) to a synthetic form for every one of their morphosyntactic features. Moreover, there isn't a feature that the auxiliary *-nga* expresses, since it is not in exclusive use for

Table 2
Tenses and modes of the Indefinite Exclusive.

	Indefinite exclusive (1st person singular)	
	Incidental	Habitual
PRES	*ndo-tora* 'I now take'	*ndi-ne-nge ndó-tora* 'I usually then take'
PAST	*nda-chi-tora* 'I then took'	*nda-i-nge ndó-tora* 'I used to then take'
RM. P.	*nda-ka-chi-tora* 'I then took'	———
FUT	*ndi-cha-chi-tora* 'I will then take'	———

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

Table 3
Tenses and modes of the Indefinite Progressive.

	Indefinite Progressive (1st person singular)	
	Incidental	Habitual
PRES	*ndi-cha-tora* 'I still take'	*ndi-ne-nge ndí-cha-tora* 'I habitually still take'
PAST	*nda-nga ndí-cha-tora* 'I still took'	*nda-i-nge ndí-cha-tora* 'I used to still take'
RM. P.	*nda-ka-nga ndí-cha-tora* 'I still took'	———
FUT	*ndi-cha-nge ndí-cha-tora* 'I will still take'	———

Table 4
Schematic distribution of periphrastic forms in the indefinite.

	Indefinite (1st person singular)			
	Incidental		Habitual	
	PRES	PAST	PRES	PAST
SIMP	*ndinotora*	*ndatora*	*ndinotora*	*ndaitora*
EXC	*ndotora*	*ndachitora*	***ndinenge ndótora***	***ndainge ndótora***
PROG	*ndichatora*	***ndanga ndíchatora***	***ndinenge ndíchatora***	***ndainge ndíchatora***

one particular feature.[1] The only purpose of the *-nga* auxiliary is to act as a support for the **tense** and **mode** formatives in the periphrastic forms (leaving the **aspect** and **implication** to be expressed in the participle).

A sound analysis of the inflectional system of the Shona verb, then, should capture the paradigmatic nature of the oppositions between synthetic and periphrastic forms. For that, there must be a level of analysis in which the morphosyntactic similarities between synthetic and periphrastic forms can be established, regardless of the differences in their syntactic realizations. I suggest that such a level of analysis can be found in LFG's f-structure. Ultimately, I will argue that the architecture of LFG needs to be slightly changed to account for the properties of periphrastic verbs within a realizational theory of morphology. But first I will introduce the classical treatment of inflection in LFG.

3. Synthetic and analytic verbs in Lexical Functional Grammar

A conception of morphology as the level of linguistic analysis that deals with word structure is often based on the assumption that the linguistic units defined by morphological rules are also lexical units. These lexical units are the terminal nodes found in the structural representations of phrases and sentences. Information carried by lexical items (i.e. semantic, categorial, and morpho-syntactic features) is projected onto the syntax, but syntactic rules do not have access to the internal structure of words, nor can they change the information associated with a particular lexical item. Such strict segregation of morphology and the lexicon from syntax is made explicit in Lapointe's (1980) Generalized Lexical Hypothesis, adopted by LFG as the Lexical Integrity Principle (Bresnan and Mchombo, 1995):

(3) Lexical Integrity Principle (LIP): Syntactic rules do not have access to the internal parts of words

Periphrastic forms are problematic for the LIP because they are composed of parts that are treated as independent units by the syntax, in spite of being paradigmatically equivalent to single inflected words.[2] To account for the paradigmatic opposition of synthetic and periphrastic forms, then, a theory of inflection must incorporate a level of representation in which all verbal forms are given a uniform description. These goals can be achieved within a theory of grammar that has a modular architecture in which constituent structure and functional structure are distinct levels of representation. Work in the LFG formalism, for instance, represents morphosyntactic properties as matrixes of functional features (Börjars et al., 1997, Butt et al. 1996; Frank and Zaenen, 2002). In my account, these non-configurational representations will provide the way to uniformly represent functionally equivalent information expressed in different constituent structures (e.g. synthetic vs. analytic).

Lexical-Functional Grammar is a mono-stratal, multi-level theory of syntax. Sentences are associated with a structure showing the hierarchical organization of words and phrases into constituents (c-structure) and a structure containing the functional information associated with the clause and its constituents (f-structure). In f-structure, linguistic information is organized as a matrix of features (represented as attribute-value pairs). The information that makes up f-structure is contributed by individual lexical items. These are the terminal nodes of a sentence, grouped into syntactic constituents. C-structure (i.e.

[1] Doke (1967:225) seems to suggest that *-nga* is used to form past continuous tenses, but it should be clear from Fortune's data, and from the analysis developed here, that the meaning of *-nga* cannot be restricted to a particular set of features.

[2] The components of a periphrastic form can often be separated by other lexical items, or they can be deleted or displaced independently of the other components. Some prosodic processes may also treat the components of a periphrastic form as independent words.

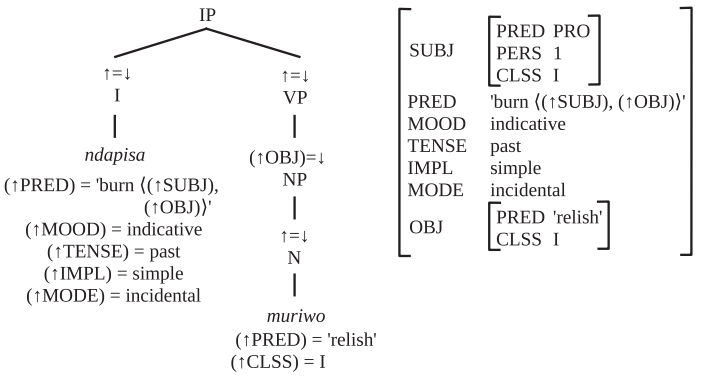
Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

constituent structure) provides a way to pool together (through the logic of unification) the diverse linguistic features projected from those lexical items, resulting in an f-structure that is consistent and complete. The constituents, or c-structure nodes, are annotated with functional equations specifying how linguistic information flows through the sentence. Thus, an equation like $\uparrow = \downarrow$ is read as 'all features of the daughter node ($\downarrow$) are features of the mother node ($\uparrow$)', and it typically identifies the head of a constituent. The equation $(\uparrow \text{OBJ}) = \downarrow$, on the other hand, specifies that the features associated with the daughter node (i.e. the NP) are those in the embedded matrix that constitutes the value of the OBJECT feature in the main f-structure.

Morphosyntactic properties are represented as features of f-structure as well. As an example, consider the Shona sentence in (4). This sentence is associated with the c-structure and its corresponding f-structure in (5). The f-structure includes features for mood, tense, implication, and mode. It also includes features for the subject and the object, the values of which are nested f-structures with person and number features. I assume that the properties that define the verbal inflectional paradigms in Shona are properties of IP.[3]

(4) a. *nda-pisa* *muriwo*
1.SG.PAST-burn relish
'I burned the relish'

(5)



a) c-structure b) f-structure

In classical LFG, fully inflected lexical items project the features that end up being pooled together in f-structure as a result of the functional equations that annotate the tree.[4] The lexical item *ndapisa* 'I burned' contributes information about the predicate, the subject, and also tense, mood, implication, and mode. Notice that these are not features of the lexical item. Rather, an annotation like $(\uparrow \text{TENSE}) = \text{past}$ on a lexical item specifies that its mother node (i.e. the head of IP) has the TENSE = past feature to be passed on to the f-structure associated with IP.

In a sentence with a periphrastic verb, on the other hand, the morphosyntactic features that make up f-structure are contributed separately by different lexical items (the components of the periphrasis), which occupy different c-structure positions (the auxiliary as the head of IP, the main verb as the head of VP). In classical LFG, the auxiliary verb and the main verb are f-structure co-heads.[5] Features from the co-heads unify in the mother f-structure. The information associated with I, then, is merged with the information associated with the VP in a single f-structure. In this analysis, the auxiliary has no PRED features, but it contributes some morphosyntactic features of its own (Mohanan, 1982; Falk, 1984; Kroeger, 1993). These features unify with those of the main predicate.[6] A sentence like (6), then, with the periphrastic verb *ndanga ndichapisa* 'burn.1SG.PAST.INCD.PROG', will have a c-structure/f-structure pair like the one in (7).

[3] Following Kroeger (1993) and Bresnan (2001), I take the finite verbal element to be the categorial head of the sentence. It occupies a privileged position as the head of IP, a functional extension of the VP. Tense, mode, aspect, implication, mood, and the subject features are realized in I, and the object features are realized in the VP. Notice the headless VP. When all of the information associated with the matrix IP is realized by I, then the head of V is optional. This is how V-to-I movement is modeled in a monostratal theory of syntax.

[4] Zaenen and Frank (2004) state that the contribution of independent morphemes to the composition of lexical meaning is accomplished by rules that are independent of syntax, since LFG strictly abides by the Lexical Integrity Principle. At the sublexical level, rules of morphology can be encoded as a finite-state transducer (Zaenen and Frank, 2004: 44).

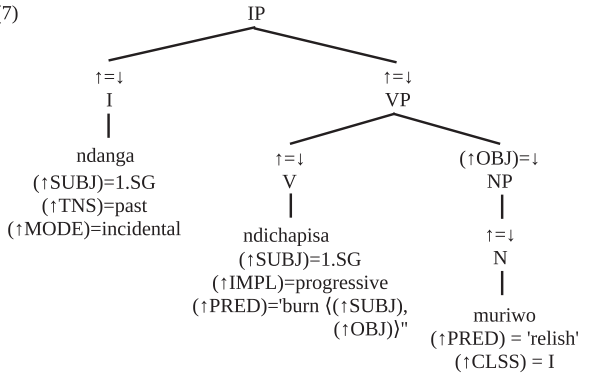
[5] Gazdar et al. (1982) elaborate an alternative analysis of auxiliaries as complement-taking verbs, similar to raising predicates. This approach is adopted in earlier analyses of auxiliaries in LFG (Kaplan and Bresnan, 1982), and in much of the HPSG literature on periphrasis. Even though the general tendency in LFG is to treat auxiliaries as mere feature carriers, Falk (1984, 2008) develops a more nuanced analysis, in which only some auxiliaries are treated as raising predicates. Yokota (2005) makes similar observations regarding light verbs in Japanese.

[6] The flexible modular architecture of LFG makes it unnecessary for f-structure heads to always correspond to c-structure heads (Bresnan, 2001:102). C-heads are f-heads, but also complements of functional categories (i.e. the lexical categories that the functional categories extend) are f-structure co-heads. This captures the intuition that "the relation of the functional F° categories to their complements is not that of predicator to argument", but that of an inflectional extension of a lexical category.

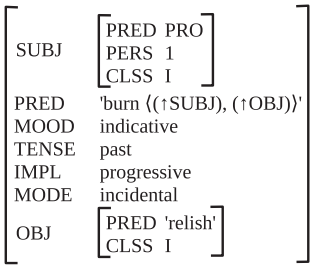
Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

(6) b. *nda-nga* *ndi-cha-pisa* *muriwo*
1.SG.PAST.INCD-be 1.SG-PROG-burn relish
'I was still burning the relish'

(7)



a) c-structure



b) f-structure

Comparing the f-structures in (5) and (7), it becomes apparent that the only difference between the two f-structures is the value of the IMPL(ICATION) attribute. At the level of f-structure, then, (4) is in a paradigmatic opposition with (6), even though the former has a synthetic verbal form, but the latter a verbal periphrasis.

There are known problems with this treatment of auxiliaries in classical LFG. The problem faced by this approach in Shona is that not all the information from the co-heads can be passed on to the f-structure in correspondence with the IP. The participle's tense and mood cannot unify with those of the auxiliary, and the implication of the auxiliary cannot unify with that of the participle. Some solutions have been proposed in the literature (cf. Zaenen and Frank, 2002 for discussion), but the root of the problem is the assumption made in LFG that all morphosyntactic properties must be lexically specified, and then passed on upwards the syntactic tree to unify in the f-structure associated with the mother node. I will suggest a different approach, extending some ideas from Realizational Morphology. Before I do so, though, it is necessary to sketch a realizational analysis of the morphology of the Shona verb, following ideas in Anderson (1992) and Stump (2001).

4. A realizational analysis of inflection in Shona

Over the last four decades, morphological theory has become dominated by realizational models of word formation (Matthews, 1972, 1974; Anderson, 1992; Halle and Marantz, 1993; Stump, 2001). The structuralist approach to the analysis of word formation was to segment words into morphs, each of them a unit of sound and meaning. As the meaning of a word is built by the addition of morphemes, so is its structure, in a compositional fashion. In realizational morphology, however, morphological processes are dissociated from their material realization. A process is a change (or specification) of a morphosyntactic property (i.e. the past tense form of a verb, or the plural of a noun). Morphs are not units of matter and meaning, but pieces of structure in correspondence with features of content (i.e. morphosyntactic features). The recognizable pieces that make up the form of a word are the **exponents** of its morphosyntactic properties. In the simplest case, there is a one-to-one correspondence between an exponent and a morphosyntactic property. But realizational morphology allows for mismatches in the correspondences between the planes of content and expression (i.e. cumulative or extended exponence).

In realizational treatments of inflection exponents and morphosyntactic properties are related by realizational rules. Following this approach, I will formulate a set of realizational rules for the morphology of the Shona verb. In the simplified

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

notation in Ackerman and Stump (2004) and Stewart and Stump (2007), I will represent realizational rules by the notation X_C, σ:τ -> Y, where X is a stem of class C, σ is a complete set of morphosyntactic properties, and τ a sub-set of the properties in σ. Thus, whereas σ defines a particular cell in a paradigm, τ identifies the features of that cell that are realized by the form Y. The hierarchical organization of affixes in a word form is achieved by assigning the realizational rules to different rule blocks (Anderson, 1992; Stump, 2001). Affixes introduced by earlier rules attach to the stem first. Moreover, only one affix from each block can be selected. In this way, each block does the job of a positional class in descriptive morphology. Mberi (2006) presents a full positional class analysis of the Shona verb, requiring at least one slot for the stem (and final vowel) and one slot for the subject concord, with seven slots in between (the complete template has 13 slots). The fragment of the Shona verb analyzed here requires only two slots between subject concord and stem. This template is necessary to accommodate forms like *ndi-cha-chi-tora* 'I will now take' and *nda-ka-chi-tora* 'I then previously took', which have three distinct prefixes before the root.[7] The positional classes I postulate for the Shona fragment analyzed here are summarized in (8).

(8)

4	3	2	1	0
H[	ndi	cha	chi	tora
	nda	ka	cha	
	ndo	i		
		no		

Going outwards from the bare stem, then, the first rule block will include the exclusive prefix *chi*, while the second block will include the future *cha* and the remote *ka*. The third block includes the concord prefixes *ndi* and *nda*, which realize the features 1.SG. *Nda* also realizes the feature {past}. As in Mberi (2006) and other discussions of Shona morphology, I consider that the complexity of the concord system does not justify assigning the person/number and tense features of the subject concord affixes to two different slots or rule blocks. But I do assign the high tone that characterized the participial mood to a separate rule block after the one introducing the concordial prefixes.

Forms like *nda-tora* 'I took', *nda-ka-tora* 'I previously took', *ndi-cha-tora* 'I will take', and *nda-chi-tora* 'I then took' are missing a prefix from block 1 or 2 (or both). Each block contains a rule that, by default, returns the same stem form as the input from the previous block. Thus, for a set like {indicative, 1SG, past, remote, incidental, simple}, Block 1 will return the stem *tora-*, since the set does not contain the property {habitual}, or {exclusive}, or the subset {simple, present}. Block 2 will add *ka*, and block 3 *nda*. Thus, *nda-ka-tora* can be inferred to be the realization of a 1SG remote past in the simple implication, because of the absence of an affix in Block 1. Likewise, *nda-chi-tora* is the result of a rule that returns an identical stem as the output of Block 2, in this case when the morphosyntactic set is {indicative, 1SG, past, incidental, exclusive}.

I will also assign to Block 1 the progressive *cha.* The past habitual *i*, and the prefix *no*, which realizes the simple present, are assigned to Block 2. Thus, a form like *ndi-no-tora* realizes the feature set {1SG, simple, present}, being ambiguous about the mode (incidental or habitual), while a form like *nda-i-tora* realizes the set {1SG, past, habitual). The prefix *ndo* requires special treatment, since it is in competition with prefixes from more than one block. *ndo* realizes the features {1SG, present, exclusive}. Being more specific than {exclusive} *chi*, it blocks its choice in Block 1. It also blocks the choice of {1SG} *ndi* in Block 3. To account for this, I introduce a portmanteau block (following Stump, 2001), covering blocks 1 to 3 (hence Block 1:3), and I assign the rule introducing the exponent *ndo* to this block. Finally, Block 4 includes a rule to add a floating high (H) tone to the left of the input form in the participial mood. The full rule set realizing the morphosyntactic features for the simple verbal forms is shown in (9):

(9) Shona rules of exponence for verbal inflection

1:3		X_V, σ:{1SG, pres, excl}	→ *ndo*X
1.	a.	X_V, σ:{excl}	→ *chi*X
	b.	X_V, σ:{prog}	→ *cha*X
2.	a.	X_V, σ:{spl, pres}	→ *no*X/*ne*X
	b.	X_V, σ:{hab, past}	→ *i*X
	c.	X_V, σ:{remote}	→ *ka*X
	d.	X_V, σ:{future}	→ *cha*X
3.	a.	X_V, σ:{past, 1SG}	→ *nda*X
	b.	X_V, σ:{1SG}	→ *ndi*X
4.		X_V, σ:{participial}	→ H[X

Realizational rules like the ones in (9) relate a set of morphosyntactic properties to a phonological string. Assuming their domain of application to be a lexical item, these realizational rules can be integrated with the LFG architecture by translating individual morphosyntactic properties into their corresponding lexical feature specifications, as Orgun (1996) suggests. LFG is committed to a sign-based perspective of grammar, so at the sub-lexical (i.e. morphological) level each morpheme will contribute a piece of form and a piece of meaning. Nevertheless, Orgun shows that it is equally possible to develop a realizational model of

[7] For the most part, the rule blocks I suggest and Mberi's (2006) slots are congruent. The additional slots in Mberi's template are needed to accommodate affixes for negation, object concord, and other modal affixes I do not discuss here. An important difference is that whereas Mberi has three different TMA slots, I have only two blocks. I group *ka* with the other tense markers (*i*/*no*/*cha*), closer to the subject concord prefix; and I group exclusive *chi* and progressive *cha* (which is indicated as *chi* in Mberi, 2006) together and closer to the root.

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

morphology in unification-based formalisms. In the approach defended by Orgun, however, word structure is built up using similar resources as phrase structure. Even this can be avoided. It isn't difficult to conceive of a realizational approach to morphology in LFG, such that f-structure feature specifications in a lexical item constitute the property sets that characterize the cells in the paradigms, and subsets of those features are mapped onto morpho-phonological material by rules of exponence.

While this approach may work as a realizational implementation of synthetic inflectional morphology in LFG, a successful extension to a treatment of verbal periphrases is not quite immediate. First, the rule set in (9) does nothing, yet, to account for the shape and distribution of the periphrastic forms. That is, if nothing else is said, it would be possible to have a form like *nda-cha-tora* as the realization of the 1SG past incidental progressive, or have the set {1.SG, past, habitual, exclusive} realized by *nda-i-tora* or *nda-chi-tora*, for instance. In recent years, several realizational analyses of periphrastic forms have been proposed in the literature. Stump (2001) develops an analysis within his theory of Paradigm Functional Morphology (PFM) that seeks to incorporate verbal periphrases into series of paradigmatic oppositions inside the lexicon. In PFM, realization rules are of two basic types: rules of exponence and rules of referral. Rules of exponence relate morphosyntactic properties to morpho-phonological formatives, including stems and affixes. Rules of referral, on the other hand, specify that the form of a particular cell is identical to the form of some other cell in the paradigm (Zwicky, 1985), and they play a prominent role in the analysis of periphrastic forms in PFM.

The realization rules for periphrastic forms map a single complete set of morphosyntactic properties onto a sequence of word forms. These word forms may be identical to other word forms in the same paradigm (for the main verb) or in the paradigms of other verbs (for the auxiliaries). Stump (2001:230–235) applies this approach to the analysis of the periphrastic future of Sanskrit. The periphrastic future is formed with the auxiliary verb AS 'be' and an agentive noun. This is shown in (10).

(10) *pratihrahītā tām asmi.*
receive:AGENT.N her:ACC.SG I.am
'I will receive her.'

The realization rule for the periphrastic future, stated in (11), specifies that two different stems, S_A and S_B, are needed to realize the morphosyntactic properties of the cells of the future sub-paradigm.

(11) A set of features M including [FUT] for a stem S is realized as a two stem sequence S_A-S_B, where
a) S_B is the realization of S for the feature [AGENT.NOUN], and
b) S_A has the same realization as the set of features M with [PRESENT] instead of [FUTURE] for the stem *AS*.

First, the person, number, and tense features are realized by the auxiliary S_A. This is done by means of a Rule of Referral, specifying that the form of the future auxiliary, for each cell, is identical to that of the present of the verb AS. Second, the main verb S_B is also specified in the realization rule, as the agentive noun derived from the main verb.

Ackerman and Stump (2004) implement realizational analyses of periphrasis in LFG, while still operating under the assumption that morphology is contained within the lexical component of the grammar. They do so by allowing the lexicon to include multi-word forms. For them, "the contentive information associated with a periphrase is specified morpholexically" (Ackerman and Stump, 2004:115). In this model, then, the periphrasis is associated with a single f-structure. But instead of building out this f-structure by unification of partial morphosyntactic information that is lexically associated with f-heads (as in Borjars et al., 1997, or Frank and Zaenen, 2002), in Ackerman and Stump's model "the clausal predicate information in f-structures is projected from a single lexical representation which receives either synthetic or periphrastic expression in c-structure."

This model faces the challenge of providing a mechanism to relate an f-structure to a c-structure in which the components of a periphrastic form are inserted as terminal nodes, since the usual means to relate the two levels of representation in LFG will not yield the right result. As they state in a footnote, "within LFG the possibility of multi-word lexical items requires modifying the conventions used for annotating c-structure expressions associated with single-word lexical items so that appropriate lexical information will produce well-formed f-structures." (Ackerman and Stump, 2004:116). The implementation details of this proposal are left to be developed, but a workaround is already proposed in Sadler and Spencer (2001). They propose a new kind of rule of referral, which realizes the form of a cell as a syntactic construction, not as just another word-form (or set thereof). The syntactic construction's f-structure is subsumed by the set of morphosyntactic features that define the periphrastic cell.

One shortcoming of this approach, noticed by the authors themselves, is that the notion of a syntactic construction has no clear ontological status within the LFG formalism.[8] A related problem with the construction referral solution is that rules of referral in the PFM theory are set to refer to another cell in the same paradigm (syncretism) or to a cell of another paradigm (suppletion). This is a desirable constraint on the power of rules of referral, from which most of their explanatory force stems, along with support for the conceptual value of the paradigm. Syntactic constructions, however, do not constitute autonomous cells in any paradigm.

I will argue that the distribution of periphrastic forms is accounted for by a different kind of morphological rule, a rule that spreads the features of a full complement of morphosyntactic properties over more than one verbal constituent. In the next sections I will show how this insight can be made explicit within a modified version of Lexical Functional Grammar, in which the theory's strict commitment to lexicalism is relaxed to allow for a more radically realizational approach to morphology.

[8] Sadler and Spencer propose to adapt the notion of a constructional predicate from Ackerman and Webelhuth (1997) to their LFG analysis.

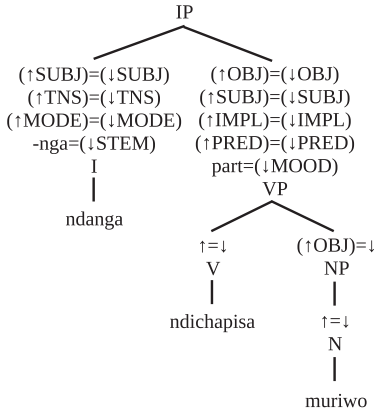
Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

5. Inflection as phrasal morphology in a functional theory of syntax

The classical approach to the treatment of periphrastic forms within the LFG framework, then, faces serious problems, even when enriched with a realizational treatment of verbal periphrases by rules of referral. The problem disappear, I suggest, if the LIP is relaxed to let the syntax build up inflected words. Following Anderson (1992), I take inflection "to be just the morphology that is accessible to and/or manipulated by rules of the syntax" (Anderson, 1992:83). In this approach, the set of morphosyntactic features that occurs in the input to realizational rules is a property of terminal nodes in a syntactic structure, not a property of lexical items. The lexicon contains only uninflected stems. A realizational rule, then, interprets the phonological content of a terminal node given its morphosyntactic features and the form of the stem. The central difference between a synthetic form and a periphrastic one is that in the latter the morphosyntactic features that make up a complete set of inflectional properties are spread over more than one terminal node.

In any sentence, the features that make up f-structure need to be realized phonologically. The role of c-structure is to assign the different features from f-structure to constituents small enough to serve as the domain for realizational rules to apply. In a sentence with a synthetic verb, all the morphosyntactic features that define a cell in the verb's inflectional paradigm show up on the same terminal node. In a periphrastic form, on the other hand, it takes two separate lexical items (the auxiliary and the participle) to realize the full set of morphosyntactic features in f-structure. Thus, I will assign a sentence with a periphrastic verb like the one in (6) a c-structure like the one in (12). The main verb and auxiliary verb are no longer treated as co-heads. Rather than serving as the vehicle to unify the morphosyntactic features contributed by the individual lexical items, the role of the c-structure in (12) is to make sure that the exponents for all the features in the sentence's f-structure show up in the right places (i.e. in the right terminal nodes). It achieves this by specifying subsets of the f-structure features of the sentence as f-structure features of each terminal node.

(12)



The functional equations that annotate the tree indicate that the auxiliary realizes the TENSE and MODE features of the clause, in addition to subject concord, while the main verb (i.e. the participle) realizes the implication. The participle also expresses the predicate feature and, like the auxiliary, it inflects for subject concord. This kind of multiple exponence of subject agreement features across the members of the periphrasis is easily accounted for in a realizational model of morphology that allows syntax to have access to inflectional properties.

Realizational morphology also offers a way to deal with another known problem in the analysis of periphrastic forms. Periphrastic forms are not wholly compositional, since some of the features of the individual constituents of the periphrastic forms do not correspond to any features of the periphrasis as a whole. The participle, for instance, has its own mood (participial) and tense (present), which do not correspond to the mood and tense of the clause (indicative past). The participle's MOOD and TENSE features, then, are only specified for the VP (and its head), and are not equal to the MOOD and TENSE features of IP. These features are realized by the auxiliary. But the auxiliary contributes no PRED feature to IP. The choice of a particular verb (or rather a verb stem) to function as the auxiliary, then, is not driven by its content. For this reason, the STEM feature of the auxiliary needs to be stipulated as *-nga* in the c-structure of the periphrastic construction.

The role of c-structure annotated with functional equations, then, is to specify where in the tree a specific morphosyntactic property will be realized by the rules of exponence of the language. This is a small but significant weakening of the LIP, necessary to implement a truly realizational approach to inflection in a monostratal theory of syntax. In a lexicalist framework abiding by the LIP, like standard LFG, most of the features in an f-structure are projected from the lexical entries of the terminal nodes. Unification ensures that the only lexical items that combine with each other in a sentence are those whose functional features are compatible. In my approach, on the other hand, morphosyntactic features are not specified in the entries of lexical items, but only in the syntax. As Anderson (1992) suggests, lexical items contribute to specify the stem of a terminal node, and its semantic content. Inflectional affixes, then, are not part of lexical entries. Rather, they are pieces of the phonological string associated with terminal nodes, licensed by the realizational rules of inflectional morphology. The lexical

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

entry for *pisa* 'burn' will then look like the one in (13). The value of the STEM feature is a phonological string, which is the input to the realizational rules. The stem and the rules together specify the phonological form of the terminal node.

(13) -pisa: (↓PRED) = 'burn ⟨(↑SUBJ), (↑OBJ)⟩'
(↓STEM) = /-pisa/

The idea that some functional features are introduced by syntactic rules, instead of being projected from the lexicon, can already be found in previous literature on LFG, most noticeable on treatments of structural case. Early studies (Neidle, 1982:398) suggested that the accusative case that marks direct objects in many languages is a feature of the phrase structure rule expanding the VP with a complement, as in (14).

(14) VP -> V NP
(↑OBJ) = ↓
(↓CASE) = ACC

The lexical entry of a noun inflected for accusative would include a constraining equation of the form (↑CASE) $=_c$ ACC, specifying that it must be inserted under a node marked ACCUSATIVE in the syntax. More recently, Spencer (2003) argues that a constraint equation for case becomes unnecessary when a distinction between syntactic features and morphological features is introduced in the grammar. Syntactic case features are properties of the phrase, and are not projected from lexical items. Inflected lexical items, however, include morphological case features. These morphological features are the exponents (in morphosyntactic terms) for the syntactic case features.

As in Neidle (1982) and Spencer (2003), I let phrase structure rules specify morphosyntactic features in the syntactic representation of a sentence, instead of being projected from lexical items. But I adopt a more radically realizational approach to inflection, since I allow inflectional affixes to be specified in the syntax. My approach, then, dispenses with the constraining equations or the morphological features specified in inflected lexical entries in previous approaches. Affixes are direct exponents of the morphosyntactic properties specified in the syntax. Moreover, following a suggestion made in Anderson (1992), I treat morphosyntactic features as phrasal in nature. Crucially, morphosyntactic phrasal features need not be realized by all the constituents of the phrase. In German, for instance, complex patterns of case marking may exclude the head, if there is a case-marked modifier.[9] The grammar must therefore include rules specifying how the morphosyntactic features of a phrase are distributed among its constituents. This can be achieved by phrase structure rules annotated with functional equations like the one in (14).

A general rule for IP, shown in (15), distributes those morphosyntactic features of IP to I, except for the information about the object, which is realized by a constituent of VP. This rule applies to the synthetic structures.

(15) IP → I VP
↑ = ↓ (↑OBJ) = (↓OBJ)

If the verb is periphrastic, on the other hand, a competing rule of phrasal morphology, shown in (16) splits the morphosyntactic features of IP among more than one head.

(16) IP → I, VP
(↑SUBJ) = (↓SUBJ) (↑SUBJ) = (↓SUBJ)
(↑MOOD) = (↓MOOD) (↑OBJ) = (↓OBJ)
(↑TENSE) = (↓TENSE) (↑PRED) = (↓PRED)
(↑MODE) = (↓MODE) (↑IMPLIC) = (↓IMPLIC)
-nga = (↓STEM) participial = (↓MOOD)

The properties are distributed in such a way that the auxiliary realizes the clause's tense, mode, mood, and subject concord features, while a participial form inside VP realizes the implication. The VP also gets the predicate feature (ensuring that the stem for the main verb is realized as the head of VP), and a copy of the subject features. The phrasal morphology rule also defines the stem that the auxiliary is associated with: the stem *-nga*. Each terminal node is then interpreted according to the realizational rules introducing the exponents for verbal morphology.

Rules (15) and (16) are in competition. To make sure that the right one is selected, one of them should be specified to apply in a given context, letting the other one apply elsewhere. I suggest that the rule introducing the periphrastic forms is the most restricted one. The context of application specifies the features that are in conflict, meaning that they cannot be realized together on the same word. For instance, rule (16) will be constrained to apply when the f-structure associated with IP contains the features [IMPLICATION:exclusive, MODE:habitual]. Other feature combinations that trigger the periphrastic rule are [IMPLICATION:progressive, MODE:habitual], [IMPLICATION:progressive, TENSE:past], and [IMPLICATION:progressive, TENSE:future]. Rule (15), then, is blocked by rule (16). Blocking is also used in other approaches to account for the distribution of synthetic and periphrastic forms. The advantage of this approach over others is that the rules that are in competition are both syntactic rules, so the blocking relation can be defined within the same component of the grammar.

[9] Complex patterns in the distribution of the German genitive case are also at the core of Spencer's (2003) phrasal analysis of CASE in LFG.

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

My analysis of periphrastic forms avoids these shortcomings by letting the syntax take care of inflection. The rule I proposed in (16) does in the syntax what Stump's (2001) rule (11) does in the lexicon. As in Stump's analysis of verbal periphrases, features that normally would appear on a single verb are distributed between the auxiliary and its nonfinite complement. Also, the auxiliary stem is specified in the rule, doing the job of the rule of referral used by Stump to introduce the auxiliary verb. But because rule (16) is a rule of syntax, the problem of how to insert a two-word lexical unit into c-structure does not even arise. As the structure is built, morphosyntactic features are spread over appropriate nodes to be realized later by rules of exponence.

The rule in (16), then, spreads the morphosyntactic properties of IP over its c-structure constituents, so that they can be realized by the appropriate terminal nodes. The rule also specifies that some features are only associated with the constituents themselves, without being projected from the mother node. This may look like a violation of the principle of monotonicity, since features of the heads are not features of the mother node. But because morphosyntactic features are realized by the terminal nodes, and not projected from them, there is no actual change or deletion of features in my approach.

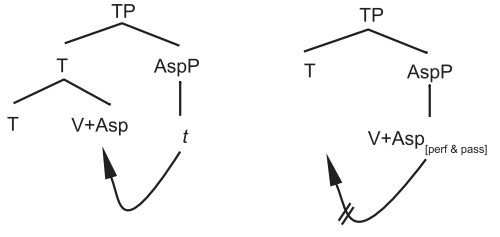
6. An alternative account of periphrastic phenomena

The analysis I have just developed combines insights from realizational morphology with the formal syntactic architecture of LFG to account for the paradigmatic opposition between synthetic and periphrastic verbal forms in Shona, but relegating the lexicon to a more modest role. Because not all functional features are projected from lexical items, the specific structural aspects of my analysis of auxiliaries departs from previous proposals. The realizational approach to the morphology of periphrastic forms I have just presented marries a syntactic view of inflection with a monostratal theory of syntax. Morphosyntactic properties are not specified in the lexicon but in a functional, non-configurational level of syntax, where they can be given a uniform representation across synthetic and periphrastic realizations. My proposal, then, innovates over linguistic models that adopt a mono-stratal, non configurational view of syntax (i.e. LFG), which place inflectional morphology strictly within the lexicon.

Taking inflectional morphology out of the lexicon may be a novelty in frameworks that traditionally adhere to the LIP, but it is accepted as a matter of course in theories that favor a configurational, multi-stratal approach to grammar. The proposal that inflectional morphology (including periphrastic forms) is carried on by word formation rules that are no different from other syntactic rules can be traced back to Chomsky's (1957) analysis of the English auxiliary system. He suggests that discrete morphosyntactic properties are heads of functional projections in sentence structure. This hypothesis is further developed in Distributed Morphology (Halle and Marantz, 1993). Functional heads in the syntax are bundles of morphosyntactic features only, lacking phonological features. The phonological representation of those categories is accomplished by insertion of **vocabulary** items after transformational rules (head movement, merger, fusion, etc.) rearrange the terminal nodes of a syntactic structure into bound sequences. Vocabulary items are like the traditional morphemes of structural linguistics, consisting of sets of phonological features and morphosyntactic features, but they do not furnish the syntax with their morphosyntactic features. The only purpose of a morphosyntactic specification on the vocabulary items is to make sure they are inserted under compatible terminal nodes. A fundamental insight of DM is that most facts about the linear arrangement of morphemes are explained as a consequence of the hierarchical organization of the syntax.

In DM, the equivalence of periphrastic and synthetic forms is resolved by making them configurationally identical to one another, giving them the same abstract structural representation. A synthetic form is the result of merging two functional heads under a single node in the syntax. It is along these lines that Embick (2000) develops a DM analysis of the alternation between the synthetic and analytic perfects of Latin. In Latin, the perfect passive is periphrastic (formed with the auxiliary verb ESSE 'be' and a participle), even though the perfect active and the present passive are synthetic. Embick's analysis (simplified somewhat) would have the V + Asp complex (i.e. Verb and aspectual morphemes) move upwards to adjoin to T(ense), resulting in a synthetic form. But if Asp contains the features [perfect] and [passive], then the movement of Asp to T is forbidden. In this case, two distinct categorial nodes are provided by the morphosyntax for vocabulary insertion: one under T, another one under Asp. Because Asp contains V but no Tense features, a verbal participle is inserted under it. T, on the other hand, prompts the insertion of the verb ESSE by default. This is how the periphrastic perfect is derived.

(17)



a. synthetic form b. periphrastic form

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

A comparison between Embick's analysis of the Latin periphrastic perfect and my analysis of the Shona periphrastic forms with *-nge* reveals an interesting difference. Embick's insight is that a combination of morphosyntactic features fixes a head in place, preventing the verb and the aspectual morpheme to move out of AspP. In my analysis of Shona, on the other hand, a periphrasis is chosen to avoid a combination of features from occurring on the same head. For instance, the Shona past progressive is periphrastic, but the features {past} and {progressive} *never occur on the same word form*. The claim is that the feature combination {past, progressive} cannot be features of the same terminal node, not that their combination is circumscribed to the main verb (as the Latin {passive} and {perfect} in Embick's analysis). In other words, the intersection of features that defines the periphrastic cell in the paradigm of Latin is realized on the same word form (i.e. the participle), but this is not the case in Shona.[10]

It would be interesting to see if an account of periphrastic verbs in Shona along the lines of Embick's analysis of Latin is feasible or not, but space limitations force me to leave the discussion for another time. My purpose in this article is not to evaluate the merits and shortcomings of DM with respect to the realizational analysis I am proposing here, but to highlight the similarities and differences between the two. In both approaches the lexicon ceases to exist as the domain where inflected words (i.e. grammatical words) are put together. The task of building up a grammatical word is shared between a syntactic component that furnishes sets of inflectional features, and the realizational component that relates those feature sets to their exponents. In DM this is done through the insertion of vocabulary items, while paradigm functions achieve that in the model I am proposing. As in DM analyses of periphrastic forms, I propose that the level at which a paradigmatic opposition between synthetic and analytic verbs is established is not the lexicon, but the syntax.[11] However, the syntactic theory I adopt is very different from the one DM assumes. In DM, a synthetic and a periphrastic form share a configurational representation, an abstract stratum in the derivational history of a clause related to a surface form by way of transformational operations. In LFG, on the other hand, the shared structure between clauses with synthetic and analytic forms is not configurational, but functional – an attribute-value matrix of linguistic features, related to a surface configuration by correspondence rules. Future work will investigate if these fundamental contrasts in the architecture of the theories make different predictions with respect to the grammar of periphrastic forms.

7. Conclusions

In this paper, I showed that complex verbs formed with the auxiliary *-nga* in Shona satisfy Ackerman and Stump's (2004) Featurally Intersective Distributional Criterion, and therefore should be analyzed as verbal periphrases in the most restrictive sense of the term. Synthetic forms like *nda-tora* 'I took' and *ndi-cha-tora* 'I am taking', for instance, are in a to-way contrast with analytic *nda-nga ndí-cha-tora* 'I was taking'. The competition between synthetic and analytic forms in the inflectional paradigm of Shona, I have suggested, can be accounted for in a model of grammar that, like LFG, associates a clause with a representation that is functional and nonconfigurational (i.e. f-structure). This level of syntactic representation contains the morphosyntactic properties that define inflectional paradigms. But instead of letting lexical items project this information onto terminal nodes, I enriched the model with phrasal realizational rules, the function of which is to distribute inflectional features among the constituents of the clause. My analysis captured the observation that the auxiliary verb *-nga* has no morphosyntactic properties of its own, and that it is used to express tense and mood features of the verb, in combination with a participle that realizes the implication features. The choice of a periphrastic over a synthetic form, then, is a strategy to avoid the realization of specific combinations of tense/mood and implication properties by the same word form.

Verbal periphrases like the ones I have discussed here put into question the received wisdom about the separation between words and phrases, between syntax and the lexicon, since these forms seem to behave as a single word from a paradigmatic perspective, but as two separate words from a syntagmatic perspective. The literature offers two major strategies to account for the paradigmatic opposition between synthetic and periphrastic forms. One solution to this problem is to allow the lexicon to contain multi-word entities (Stump, 2001; Sadler and Spencer, 2001; Ackerman and Stump, 2004). This provides a unified view of inflection as a lexical phenomenon, but within a theory that has no workable mechanism for lexical insertion. The other solution to the problem is to allow the syntax to build up inflection for synthetic and analytic forms alike, as in Distributed Morphology (Halle and Marantz, 1993; Embick, 2000), but this requires a rigid hierarchical organization of morphosyntactic properties, with additional levels or rules to account for mismatches between morphosyntax and morphophonology. Analyses that attempt to define the formal equivalence between synthetic and periphrastic forms in the syntax are committed to a configurational view of syntax. The limitation of nonconfigurational theories like LFG, on the other hand, is their lexicalist conception of inflection.

[10] Arregi and Nevins (2012) discuss the morphology of auxiliaries in Basque within the DM framework. Most verbs in Basque are analytic, with the main verb expressing aspect and the auxiliary expressing tense and agreement. Arregi and Nevins focus mostly on the morphology of the auxiliaries, however, without going into much detail about the analysis of the morphosyntax of complex predicates in Basque.

[11] In the preceding discussion I have assumed that synthetic verbal forms, whether in Latin or in Shona, are the result of a syntactic rule that constructs a tree with a single terminal node for the verb and its associated morphosyntactic properties. An alternative is to let the opposition between synthetic and analytic (i.e. periphrastic) forms appear as the result of postsyntactic rules, as in Embick and Noyer's (2001) treatment of comparative and superlative adjectives. Embick and Noyer argue that the contrast between synthetic *bigger/biggest* and analytic *more/most interesting* is postsyntactic on the grounds that *-er/-est* suffixation shows lexical sensitivity and that it is blocked by intervening adverbs, as in **amazingly smartest*. I am not aware of analogous effects among the verbal paradigms I am investigating, so I will not pursue this alternative any further.

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004

In this paper I adopted elements from these seemingly incompatible approaches, but with different assumptions about the nature of morphosyntactic representations. Like other LFG analyses, I defined the equivalence between synthetic and periphrastic forms at a level of syntactic representation where morphosyntactic properties are represented in a nonconfigurational manner. In so doing, I was able to develop a realizational analysis of the morphology of the Shona verb, along the lines of Anderson (1992). I have show that the grammar of verbal periphrases in Shona offers evidence that in order to bring periphrastic forms into a morphological paradigm, inflectional morphology must be done outside the lexicon, and that a realizational approach to the problem is best formulated in a syntactic framework that includes a functional level of representation.

Acknowledgments

This paper was first presented at the Interdisciplinary Workshop on the Grammatical Word at UC Davis, in May of 2013. I benefited greatly from the comments offered by the participants, whom I thank here. Comments from Brian Joseph were particularly useful. Special thanks go to Dunstan Brown, who offered valuable comments on an earlier version of the paper.

References

Ackerman, Farrell, Stump, Gregory, 2004. Paradigms and periphrastic expression: a study in realization-based lexicalism. In: Sadler, Louisa, Spencer, Andrew (Eds.), Projecting Morphology. CSLI Publications, Stanford, pp. 111–158.
Ackerman, Farrell, Webelhuth, Gert, 1997. A Theory of Predicates. CSLI Publications, Stanford.
Anderson, Stephen, 1992. A-Morphous Morphology. Cambridge University Press, Cambridge.
Arregi, Karlos, Nevins, Andrew, 2012. Morphotactics: Basque Auxiliaries and The Structure of Spellout. Springer Verlag, Dordrecht.
Bresnan, Joan, 1982. The passive in lexical theory. In: Bresnan, Joan (Ed.), The Mental Representation of Grammatical Relations. The MIT Press, Cambridge, pp. 3–86.
Bresnan, Joan, 2001. Lexical-functional Syntax. Blackwell, Oxford.
Bresnan, Joan, Mchombo, Sam, 1995. The lexical integrity principle: evidence from Bantu. Natural Lang. Linguist. Theory 13, 181–254.
Butt, Miriam, Niño, María-Eugenia, Segond, Frédérique, 1996. Multilingual processing of auxiliaries within LFG. In: Gibbon, Dafydd (Ed.), Natural Language Processing and Speech Technology. Mouton de Gruyter, Berlin, pp. 111–122.
Börjars, Kersti, Vincent, Nigel, Chapman, Carol, 1997. Paradigms, periphrases and pronominal inflection: a feature-based account. In: Booij, Geert, van Marle, Jaap (Eds.), Yearbook of Morphology 1996. Kluwer, Amsterdam, pp. 155–180.
Chomsky, Noam, 1957. Syntactic Structures. Mouton, The Hague.
Déchaine, Rose-Marie, Girard, Raphaël, Mudzingwa, Calisto, Wiltschko, Martina, 2014. The internal syntax of Shona class prefixes. Lang. Sci. 43, 18–46.
Doke, Clement M., 1945. Bantu: Modern Grammatical, Phonetical, and Lexicographical Studies Since 1860. International African Institute, London.
Doke, Clement M., 1967. The Southern Bantu Languages. International African Institute, London.
Embick, David, 2000. Features, syntax, and categories in the latin perfect. Linguist. Inq. 31, 185–230.
Embick, David, Noyer, Rolf, 2001. Movement operations after syntax. Linguist. Inq. 32, 555–595.
Falk, Yehuda, 1984. The English auxiliary system: a lexical-functional analysis. Language 60, 483–509.
Falk, Yehuda, 2008. Functional relations in the English auxiliary system. Linguistics 46, 861–889.
Fortune, George, 1955. An Analytical Grammar of Shona. Longmans, Green, and Co, London.
Frank, Anette, Zaenen, Annie, 2002. How we say WHEN it happens: contributions to the theory of temporal reference in natural language. In: Kamp, Hans, Reyle, Uwe (Eds.), Tense in LFG: Syntax and Morphology. Niemeyer, Tübingen, pp. 17–52.
Gazdar, Gerald, Pullum, Geoffrey K., Sag, Ivan A., 1982. Auxiliaries and related phenomena in a restrictive theory of grammar. Language 58, 591–638.
Halle, Morris, Marantz, Alec, 1993. Distributed morphology and the pieces of inflection. In: Hale, Kenneth, Keyser, Samuel Jay (Eds.), The View from Building 20. MIT Press, Cambridge, pp. 111–176.
Kaplan, Ronald, M., Bresnan, Joan, 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In: Bresnan, Joan (Ed.), The Mental Representation of Grammatical Relations. The MIT Press, Cambridge, pp. 173–281.
Kroeger, Paul, 1993. Phrase Structure and Grammatical Relations in Tagalog. CSLI Publications, Stanford.
Lapointe, Steven, 1980. A Theory of Grammatical Agreement. PhD Dissertation. University of Massachusetts Amherst.
Matthews, Peter, 1972. Inflectional Morphology. Cambridge University Press, Cambridge.
Matthews, Peter, 1974. Morphology. Cambridge University Press, Cambridge.
Mberi, Nhira E., 2006. The Cateogrical Status and Functions of Auxiliaries in Shona. ALLEX Project, Oslo.
Mohanan, K.P., 1982. Grammatical relations and clause structure in Malayalam. In: Bresnan, Joan (Ed.), The Mental Representation of Grammatical Relations. The MIT Press, Cambridge, pp. 427–503.
Neidle, Carol, 1982. Case agreement in Russian. In: Bresnan, Joan (Ed.), The Mental Representation of Grammatical Relations. MIT Press, Cambridge, pp. 391–426.
Orgun, Cemil Orhan, 1996. Sign-based Morphology and Phonology. PhD Dissertation. University of California Berkeley.
Sadler, Louisa, Spencer, Andrew, 2001. Syntax as an exponent of morphosyntactic features. In: Booij, Geert, van Marle, Jaap (Eds.), Yearbook of Morphology 2000. Kluwer, Amsterdam, pp. 71–96.
Spencer, Andrew, 2003. A realizational approach to case. In: Butt, Miriam, King, Tracy Holloway (Eds.), Proceedings of the LFG03 Conference. CSLI Publications, Stanford, pp. 387–401.
Stewart, Thomas, Stump, Gregory, 2007. Paradigm-function morphology and the morphology-syntax interface. In: Ramchand, Gillian, Reiss, Charles (Eds.), The Oxford Handbook of Linguistic Interfaces. Oxford University Press, Oxford, pp. 383–421.
Stump, Gregory, 2001. Inflectional Morphology: a Theory of Paradigm Structure. Cambridge University Press, Cambridge.
Yokota, Kenji, 2005. The structure and meaning of Japanese light verbs. Lang. Sci. 27, 247–280.
Zwicky, Arnold M., 1985. Rules of allomorphy and phonology-syntax interactions. J. Linguist. 21, 43I–436.

Please cite this article in press as: Aranovich, R., Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona, *Language Sciences* (2015), http://dx.doi.org/10.1016/j.langsci.2015.05.004