

## CONSTRAINING XP-SEQUENCES\*

Katherine Demuth  
Brown University

Jeffrey Gruber  
Université du Québec à Montréal

### 1. Introduction

Current approaches to syntactic theory offer several proposals for the structure of the Inflectional Phrase (henceforth IP) (e.g. Chomsky 1989, Pollock 1989, Belletti 1990). At the heart of this controversy have been questions about how many functional projections are included in the IP, what the order of those projections is, and whether these characteristics are universal or language particular. These questions are part of the general problem of determining what the functional XP sequences are and how they cooccur and interact with the various lexical or thematic projections (e.g. VP, NP) (see Abney 1987, Fukui and Speas 1985) to form an 'extended projection' (Bittner and Hale 1993, Grimshaw 1991, 1993).

In this paper we find solutions to aspects of each of these problems in a strong claim about the form of the basic X' structure that feeds the syntactic system. Much of our discussion draws on evidence from morphologically rich Bantu languages, where 'compound tense' forms like those in (1) pose problems for traditional analyses of IP structure.

- (1) Ke-n-e ke-tla-reka dijo. (Sesotho/Setswana)<sup>1</sup>  
AGR-COP-PAST AGR-IRR-buy food  
'I would have bought food.'

Our analysis of these constructions provides strong evidence for two types of syntactic organization. First, the BASIC PROJECTION SEQUENCE (BPS) constitutes a basic unit which in the verbal domain is realized as AgrP-TP-VP. The BPS thus constrains the order and composition of functional and lexical projections, both within the IP and in the composition of extended projections generally (e.g. CP's, DP's). Second, our analysis of 'compound tense' forms shows that these are sequences of IP's, each a BPS unit. Thus, BPS units combine to form a LEXICAL PROJECTION SEQUENCE (LPS). The LPS constrains the order and composition of thematic and non-thematic projections among the iterating BPS structures. In general, syntactic structures built by repetition of these basic units make it possible to avoid postulating a large variety of functional projection types.<sup>2</sup>

While our discussion centers on Sesotho and Setswana, two closely related southern Bantu languages, our findings have direct implications for the structure of other languages. This is evidenced in the provision of a fruitful account of language variation involving Bantu, Italian, French and English. We argue that the structures we identify are more widely found than a cursory examination of surface structures might show.

\*The work of the second author was supported in part by SSHRCC grant no. 411-92-0012 'La Modularité de la Grammaire: Arguments, Projections et Variations' and by the Fonds Institutionnel de Recherche of the Université du Québec à Montréal to Anne Marie Di Sciullo. We thank Chris Collins, Anne Marie Di Sciullo, Kenneth Hale, Morris Halle, Kisangati Kinyalolo, Elisabeth Löbel, 'Malillo Machobane, Alec Marantz, Juvénal Ndayiragije and Ed Stabler for helpful comments and criticism.

<sup>1</sup>Sesotho (or Southern Sotho) is spoken by approximately 4 million speakers residing in the countries of Lesotho and South Africa. Setswana is spoken by approximately 5 million speakers residing in the countries of Botswana and South Africa. The two languages are part of the Sotho group and are generally considered mutually intelligible. Our data is drawn from Doke and Mofokeng (1957) and Cole (1955). See Machobane (1985) for a semantic treatment of some of the issues discussed here.

Unless otherwise specified, the examples provided are equivalent for both Sesotho and Setswana, with a modified Sesotho orthography. Bantu languages are generally characterized by a noun class and 'concordial' agreement system composed of several singular/plural pairs (e.g. 9/10), identified by standardized numbering (see Welmers 1973: 165). Glosses are as follows: 1SG = 1st person singular, 1 = noun class #1, AGR<sub>O</sub> = object agreement formative, AGR<sub>S</sub> = subject agreement formative, AFOC = anti-focus, ASP = aspect, COP = copula, COND = conditional tense, CONT = continuous aspect, EXCL = exclusive, FPT = future participial, FUT = future, HAB = habitual, INF = infinitive, IRR = irrealis, MOD = Modal, NEG = negative formative, OCC = occasional, PAST = past, PERF = perfect, POSS = possessive, POT = potential, PRES = present, T = Tense, V = Verb. We do not address the possible modal status of the verb-final vowel -a, which alternates with -e in the negative imperfect present.

<sup>2</sup>We are concerned here with the phrasal projections AgrP and TP or DP immediately extending each lexical phrasal projection, VP or NP, respectively, making up a BPS, and not primarily with its further extension to CP or KP, considered in the work of Bittner and Hale (1993) and Grimshaw (1991, 1993). In our terms the latter would constitute the higher extension of an LPS. But see §4 for discussion of the role of CP.

Indeed, their fundamental motivation, on both empirical and conceptual grounds, suggests that they are universal.

Our analysis focuses on syntactic structure, and assumes recent work on head movement (Travis 1984, Chomsky 1989, Pollock 1989) and its implications for word formation (Li 1990, Poletto 1991). Furthermore, it treats functional items (grammatical morphemes) as syntactically independent abstract feature bundles, thereby allowing for the possibility of their null or coalesced morphological realization as a matter of language variation (Demuth 1994, Gruber 1976, Halle and Marantz 1993; see in contrast Anderson 1992, Lieber 1992). The analysis thus elucidates the relation among functional elements and items of lexical content as they occur in the BPS, as well as some of the mechanisms involved in determining their representation in phonetic form (PF). A unified explanation is thereby provided for several seemingly disparate morpho-syntactic problems such as *pro*-drop, auxiliary ordering, affix-hopping and *do*-support.

The paper proceeds as follows. In §1 we examine simple and complex Bantu IP structures, showing how both adhere to an AgrP-TP-VP structure. In §2 we present the BASIC PROJECTION SEQUENCE as the fundamental architectural unit which constrains the number and type of functional projections within the IP. In §3 we show how the LEXICAL PROJECTION SEQUENCE (LPS) incorporates the BPS in a principled fashion, providing constraints on the ordering among thematic and non-thematic IP's. In §4 we employ restrictions inherent in the LPS to explicate the syntactic distinctions between auxiliary, raising and non-raising thematic verbs, and to explore the mechanisms of Case assignment in constructions with multiple AgrP's. In §5 we discuss interactions at the morpho-syntactic interface, showing how the BPS provides a principled set of constraints on word formation. Finally, we conclude in §6 with a sketch of how the BPS might be generalized to DP structures.

### 1.1 IP structure in Bantu languages

Bantu languages are generally characterized as SVO *pro*-drop languages, where an obligatory subject marker prefixes to the verb, and the lexical subject can be freely reordered. The lexical object can also be freely reordered when an object marker is prefixed to the verb (Bresnan and Mchombo 1987, Demuth and Johnson 1989). In this paper we assume that subject markers and object markers are AGR<sub>S</sub> and AGR<sub>O</sub> functional heads respectively (see Demuth 1992, 1994). The underlying structure of Bantu IP's is then relatively transparent, generally consisting of an AGR<sub>S</sub>, a TENSE/MODAL/ASPECT formative (henceforth T(ense)), an optional AGR<sub>O</sub>, and the Verb to which the other formatives are prefixed, in that order. This is illustrated in the Sesotho/Setswana examples in (2).

- (2)a. Ke-tla-reka dikhomo. [AGR<sub>S</sub>-[T-V]] N  
 1SG-FUT-buy 10cattle  
 'I will buy cattle.'
- b. Ke-tla-di-reka. [AGR<sub>S</sub>-[T-[AGR<sub>O</sub>-V]]]  
 1SG-FUT-10them-buy  
 'I will buy them.'

This is the most salient and common order amongst IP formatives. Apparent 'exceptions' are of two types. The perfect-tense/aspect formative *-ile* suffixes to the verb in (3a), and the present tense formative *-a-* in (3b) is realized by a null morpheme when a lexical object follows the verb in (3c).<sup>3</sup>

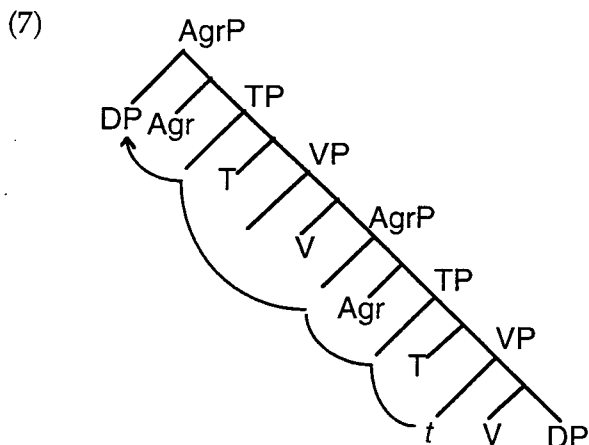
- (3)a. Ke-di-rek-ile. [AGR<sub>S</sub>-[[AGR<sub>O</sub>-V]-T]]  
 1SG-10them-buy-PAST  
 'I (have) bought them.'
- b. Ke-a-di-reka. [AGR<sub>S</sub>-[T-[AGR<sub>O</sub>-V]]]  
 1SG-PRES-10them-buy  
 'I am buying them.'

<sup>3</sup>We treat the perfect marker *-ile* as a simple suffix. It is sometimes analyzed as *-il-*, an allomorph of *ya* 'go' with final affix *-e*. All other T's are prefixes. The perfect form can have simple past significance. The temporal status of *a-* is debatable—it may rather be thought of as a prosodic place holder or 'Antifocus' marker. We do not address these issues here as they do not bear on our analysis.



order. We identify these constructions as iterative IP's, where each IP conforms to the basic underlying structure ascribed to the simple tense forms discussed above. That is, 'compound tense' constructions appear to be successively nested IP's, each IP composed of AgrP, TP and VP, in that order, with succeeding IP's serving as complements to VP. This is shown in (7), where the subject raises from the lowest (main) verb to the highest Spec AGR<sub>S</sub>.

We note here that the basic underlying IP pattern shown in (4) is given here twice, with a minor difference between the initial and final IP's in (5); the initial IP never takes a lexical object (or AGR<sub>O</sub>P), while the final IP will if the verb is transitive. In these constructions the 'thematic' verb, i.e. the main verb or the verb with predicative significance that determines argument structure, is the last in the sequence, while a non-thematic verb, or auxiliary verb (or copula), resides in the first IP. Generalizing to syntactic structure, the thematic verb is the lowest in a series of iterative IP's, while auxiliary verbs are higher.



Iterative IP structure is a robust phenomena found not only in Sesotho and Setswana, but in other Bantu languages as well. Consider the following examples from Kiswahili (our glosses) (Carstens and Kinyalolo 1989:3):

- (8)a. Juma a-ta-kuwa a-me-pika chakula.                    AGR<sub>S</sub>-T-V    AGR<sub>S</sub>-T-V  
       I-FUT-COP I-PERF-cook food  
       'Juma will have cooked food.'
- b. Juma a-li-kuwa a-ki-pika chakula.                    AGR<sub>S</sub>-T-V    AGR<sub>S</sub>-T-V  
       I-PAST-COP I-CONT-cook food  
       'Juma was cooking food.'

Note the similarity between the Sesotho/Setswana examples in (5) and the Kiswahili examples in (8), where the verb in the higher IP is once again a copula, and the verb in the lower IP is a thematic verb. In other words, the basic sequence AGR<sub>S</sub>-T-V is strictly followed in Kiswahili. It would therefore appear that IP structure in Bantu languages is strongly constrained. Specifically, every TP has a VP complement, and every VP must be headed by a T. Furthermore, every TP-VP sequence must have its own AGR<sub>S</sub>P, as shown by the presence of multiple AGR<sub>S</sub> formatives in Iterative IP structures.

We will argue below that this AgrP-TP-VP structure, which we call the Basic Projection Sequence (BPS), is universal, even when obscured by surface realization. For example, Tense formatives and auxiliary verbs may be realized as a single morpheme:

- (9)a. Ke-ne                    ke-Ø-di-reka.  
       ISG-COP+PAST ISG-CONT-1othem-buy  
       'I was buying them.'
- b. Ke-se                    ke-Ø-di-reka.  
       ISG-EXCL+PAST ISG-CONT-1othem-buy  
       'I am already buying them.'

In (9a) the formative *-ne* is derived from the copula *-na*, which has combined with the perfect formative *-e*, the phonological variant of *-ile* when preceded by a nasal ( $n + ile > ne$ ). Thus, *-ne* consists of T+V, and adheres to the sequence of formatives already seen in IP structures. Likewise, the auxiliary *-se* is the reduced perfect form of the verb *sala* ‘to remain’ ( $sal + ile > setse$ ;  $s + e > se$ ), and therefore also adheres to the basic sequence of IP formatives (Doke and Mofokeng 1957). In both cases, continuous aspect is encoded by a null morpheme on the main (thematic) verb. Thus, while the constructions in (8) do not appear to conform to the Basic Projection Sequence on the surface, we argue that they do at the structural level.<sup>6</sup>

Another possible ‘exception’ to the BPS might be the copula *-na*. Unlike the copula *-be/bo*, the copula *-na* is used in constructions meaning to ‘have’ or to ‘be (with)’. Forms such as those in (10) indicate that the copula *-na* cooccurs with a null present or continuous tense/aspect formative.

- (10)a. Ke-Ø-na        le buka.  
 1SG-PRES-COP with 9book  
 ‘I have (be with) a book.’
- b. \*Ke-tla-na        le buka.  
 1SG-FUT-COP with 9book  
 [‘I will have (be with) a book.’]
- c. Ke-tla-be/bo ke-Ø-na        le buka.  
 1SG-FUT-COP 1SG-CONT-COP with 9book  
 ‘I will have (be with) a book.’

Note that *-na* cannot occur with another Tense formative, hence the ungrammaticality of (10b). Furthermore, an Iterative IP construction employing the copula *-be/bo* must be used if the future tense is employed (10c).

We therefore argue that apparent exceptions to the BPS are only surface morphological variants, not structural exceptions. This implies that ‘exceptions’ in other languages may be analyzed in a similar fashion. Consider the following examples from Zulu (Doke 1927: 169, 204; our glosses):

- (11)a. Ngi-be            ngi-Ø-thanda.  
 1SG-COP+PAST 1SG-CONT-love  
 ‘I was loving.’
- b. Ngi-zo-ke        ngi-Ø-thande.  
 1SG-FUT-OCC 1SG-CONT-love  
 ‘I will love sometime soon.’

The examples in (11) are very similar to those seen for Sesotho and Setswana in (9) above. We propose that they also have the same syntactic structure – i.e. in every case a Tense formative is complemented by a Verb, and every Agr formative is complemented by a Tense formative. In other words, ‘compound tense’ constructions in other Bantu languages also conform to the BPS.

In sum, we have shown that ‘compound tense’ constructions in Bantu languages appear to be iterative IP structures where each IP is composed of an AgrP, a TP and a VP, in that order. Furthermore, we have seen that the heads of TP and VP in some IP’s may morpho-phonologically coalesce, obscuring their status as separate formatives. In the following section we will show that such forms are quite widely attested amongst the world’s languages. We turn now to a discussion of the BPS and the implications it holds for the structure of language.

## 2. The Basic Projection Sequence (BPS)

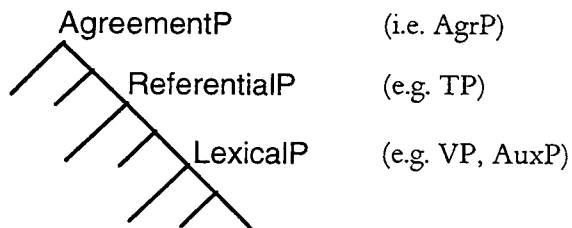
Drawing on the foregoing evidence from Bantu languages, we argue that each IP must contain the sequence of syntactic projections described above as a basic architectural requirement of language. Furthermore, we note that the sequence requirement cannot be one of ‘strict subcategorization’ or ‘selection’. While VP may be subcategorized by T, this is not sufficient to require every VP to co-occur with a TP, since such a condition should normally be in the direction of c-command. In addition,

<sup>6</sup>Note that, despite historical processes of categorical reanalysis (Heine 1991), the BPS is still met.

subcategorization or selection would not be sufficient to require, or make available, AgrP's in exactly the positions where they may or must occur.

Rather, the Basic Projection Sequence, or the BPS, is of conceptual significance to syntactic theory. In Grimshaw's (1993) terms, the well-formedness of an extended projection is a matter of principles of projection. Firstly, the structure reflects the functional necessity to associate every lexical or 'thematic' head (e.g. Verb) with an immediately c-commanding head providing reference-related properties (e.g. Tense, i.e. tense, modality or aspect). Thus, as observed in Di Sciullo and Williams (1987), lexical phrasal projections in syntax are associated with reference, either to objects (NP's), time (IP's), or truth values (CP's). The projection of each referential and thematic head is then associated with an Agreement Phrase, this also arising out of functional necessity, viz., the licensing of arguments and words, as discussed below. The resulting structure, as instantiated for an IP, is shown in (12).

(12) Basic Projection Sequence (BPS)



Thus the BPS constitutes an elemental linguistic unit consisting of three types of projection, nested in a particular sequence and distinguished by function, content and their mutually supportive roles.

The structure of the BPS given in (12) is similar to that proposed for IP's elsewhere. First, it incorporates the split-Inflection hypothesis of Pollock (1989), where the IP is analysed as an AgrP and a TP. Secondly, it incorporates the ordering insights of Chomsky (1989) and Belletti (1990), where the AgrP dominates the TP. Thirdly, it allows for the possibility of cyclic head-to-head movement, where V can move to T, and where both V+T can then raise to Agr, as is commonly assumed in the case of languages like Italian and French. Finally, wherever the structural presence of AgrP is obligatory, there is the necessity for a subject to move into it for the discharge of its  $\phi$ -features, suggestively encompassing thereby the Extended Projection Principle. Thus, the BPS captures many of the formal characteristics that have been posited as critical for the structure of language, while at the same time providing a principled explanation for their presence. We pursue these explanations below.

### 2.1 Restrictions on number and type of functional projections

The specific contribution made here to the structure of IP is to provide a principled basis for *constraining* the number, type and order of functional projections that human language allows. In general, this obviates the need for postulating a variety of types of functional projection and constructions in which they occur. Instead, the BPS constitutes an elemental linguistic unit from which more elaborate forms are constructed and various surface forms are derived. Accordingly, we expect the BPS to generalize across languages and constructions. For the moment we restrict our discussion to the BPS as it applies to the IP. The instantiation of the BPS in other domains (e.g. CP's and DP/NP's) will be considered in §4 and §6 respectively.

Consider the status of the ReferentialP (we will generally refer to this as TP within the verbal domain). Some of the controversy surrounding the number and type of functional projections permitted in human language has centered around the possibility of having both a Tense Phrase and an Aspect Phrase as functional projections within the same extended projection (e.g. Carstens and Kinyalolo 1989, Belletti 1990). From the perspective of the BPS, such constructions are disallowed. That is, the presence of both a functional Tense Phrase and a functional Aspect Phrase implies the extended projections of two distinct VP's, or two BPS's. But, perhaps more importantly, the BPS also provides a logical explanation for why such proposals have been made. We have already seen that the surface realization of tense/aspect formatives is subject to variation. We have also seen that such formatives frequently coalesce with semantically bleached auxiliary verbs. This has led to multiple interpretations regarding both the semantic content of TP's, as well as the number and type of functional categories permitted within a given IP.

The BPS is agnostic regarding the particular semantic content of a given TP, or ReferentialP. It allows for the possibility that this structural position is filled by a formative that specifies either temporal, modal, or aspectual content. For example, we saw in (3a) above that Sesotho/Setswana verbs can be marked for the perfect suffix *-ile*. Given that *-ile* has a possible aspectual semantic interpretation, and follows the verb, we might expect, given a slot-filler type of morphology, that another Tense formative could precede the verb. However, this results in the ungrammatical form in (13a). Rather, the Tense formative *-tla-* must occur within its own BPS, as shown in (13b).

- (13)a. \*Ke-tla-rek-ile buka. \*AGR<sub>S</sub>-T-V-ASP  
 1SG-FUT-buy-PERF 9book  
 ['I will have bought the book.']
- b. Ke-tla-be/bo ke-rek-ile buka. AGR<sub>S</sub>-T-V [AGR<sub>S</sub>-[V-ASP]]  
 1SG-FUT-COP 1SG-buy-PERF 9book  
 'I will have bought the book.'

Similarly, a tense and modal formative cannot occur within the same BPS. Thus, (14a) is grammatical with two BPS's, the first with a tense formative and the second with a modal formative. In contrast, (14b), with both a tense and a modal formative within one BPS is ungrammatical.

- (14)a. Ke-tla-be/bo n/ke-ka-di-reka. AGR<sub>S</sub>-T-V AGR<sub>S</sub>-MOD-AGR<sub>O</sub>-V  
 1SG-FUT-AUX 1SG-POT-10them-buy  
 'I will possibly buy them.'
- b. \*Ke-tla-ka-di-reka. \*AGR<sub>S</sub>-T-MOD-AGR<sub>O</sub>-V  
 1SG-FUT-POT-10them-buy  
 'I will possibly buy them.'

Examples such as (13) and (14) show that only one Tense formative, either temporal, modal, or aspectual, is permitted with a given BPS, each dominated by an AgrP, and each complemented by a VP. In other words, only one ReferentialP is permitted within the BPS. This means that tense, modality and aspect, each with its referential-like properties, have the same structural requirements, each occurring within their own BPS. The number and type of functional categories permitted across languages is thereby constrained.

Interestingly, these constraints also appear to apply to negation. While we leave open the possibility of distinct alternative base generated positions for negation formatives (e.g. Di Sciullo and Tremblay 1993, Zanuttini 1991), we suggest that the embedded type of negation can also occupy the head of the ReferentialP. Consider the Sesotho/Setswana examples in (15).

- (15)a. Ke-tla-be/bo ke-sa-di-reka. AGR<sub>S</sub>-T-V AGR<sub>S</sub>-NEG-AGR<sub>O</sub>-V  
 1SG-FUT-AUX 1SG-NEG-10them-buy  
 'I will not buy them.'
- b. Ke-tla-\*(be/bo) \*(ke)-sa-di-reka. AGR<sub>S</sub>-T-\*(V)\* (AGR<sub>S</sub>)-NEG-AGR<sub>O</sub>-V  
 1SG-FUT-AUX 1SG-NEG-10them-buy  
 'I will not buy them.'

Example (15) shows that negation must occur within its own BPS; it can not cooccur with a TP within the same IP. We therefore propose that TP, or more generally ReferentialP, is the structural projection under which tense, modality, aspect, or negation can be realized. We pursue these issues further in §3.

In sum, each TP must be complemented by a VP. The universality of this requirement is suggested by the use of 'dummy' auxiliaries to support Tense in languages generally. The requirement that each TP has a VP complement is demonstrated in Sesotho/Setswana by the fact that the dummy copula *-be/bo* is inserted to support TP where necessary (e.g. 14a). Dummy auxiliaries in other languages play the same supporting role. In English, the function of *have* and *be* as well as *do*-support may be viewed in this way. The tense formative must be supported by the dummy verb *do*, not because tense is an affix as in traditional analyses, but because each TP must be associated with a VP. The negative formative that apparently separates the TP from its VP complement can then be analyzed as being structurally within its own BPS, as indeed it is overtly in Sesotho/Setswana (15a). Thus in English as in Sesotho and Setswana, negation must appear