THE PROSODIC MARKING OF DISCOURSE FUNCTIONS

Louise Mycock and John J. Lowe
University of Oxford

Proceedings of the LFG13 Conference

Miriam Butt and Tracy Holloway King (Editors)

2013

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

The prosodic marking of Discourse Functions such as Focus presents a challenge to theories that seek to model grammar in a modular way because distinct components of linguistic structure must be permitted to interact but they must be neither conflated nor assumed to be isomorphic. We present an account of prosodic Focus marking within the modular grammatical architecture of Lexical-Functional Grammar, building on the model of the syntax-prosody interface developed by Dalrymple and Mycock (2011).

1 Introduction

We hold that an absolute modularity and strict separation of, for example, phonology from syntax, semantics and pragmatics is a theoretical desideratum, and that this is best captured within the framework of Lexical-Functional Grammar (LFG), in which separate levels of linguistic representation are connected by projection functions. Motivated by this commitment to modularity and domain specificity, our main objective is to capture the fundamental differences that distinguish various aspects of linguistic structure whilst permitting these distinct elements of the grammar to interact in an appropriately restricted way. The prosodic marking of Discourse Functions presents a challenge to grammatical modelling because of the interaction of a number of distinct components of the grammar: prosody, syntax, semantics and information structure. This challenge is particularly serious for a grammatical architecture that aims for strict modularity, i.e. in which distinct components of the grammar, such as syntax and phonology, are organized according to their own rules and primitives and are effectively 'blind' to the rules and primitives of other components. The prosodic marking of Focus also represents a challenge because of the types of mismatches that arise: it is not necessarily the case that the prosodic exponent of Focus marking will be coterminous with the string of syntactic elements that represent the Focus of the sentence in information structural terms. In this paper we present an account of prosodic marking of the Discourse Function (DF) Focus in English, building on the model of the syntax-prosody interface presented in Dalrymple and Mycock (2011).

We account for the prosodic indication of Focus by marking Information Structure category status on string elements. When signalled prosodically, a label indicating Focus will be associated with a p-string element bearing main stress. The principle of Interface Harmony (Dalrymple and Mycock 2011) will require a corresponding syntactic label to be associated with a corresponding s-string element;

[†]The authors' names are arbitrarily ordered: both contributed equally to the paper. The authors are grateful to the audiences at SE-LFG09, LFG2013, CIL19 and the Fachbereichskolloquium of the FB Sprachwissenschaft at Universität Konstanz, and to the editors and the reviewers of this paper for their attention and comments, and in particular to Mary Dalrymple, Ash Asudeh, and Dag Haug. John Lowe acknowledges the financial support of the Leverhulme Trust. Louise Mycock gratefully acknowledges financial support from the Faculty of Linguistics, Philology and Phonetics at the University of Oxford, which enabled her to attend LFG13.

this will ultimately correlate with the meaning of the relevant lexical item(s) being categorized as a member of the Focus set at the level of information structure, following Dalrymple and Nikolaeva (2011). We streamline the architecture proposed in Dalrymple and Mycock (2011) and show how it can account for prosodic marking of Focus. In §2 we present the data for Focus marking in English that we will analyse in this paper. In §3 we describe the model of the syntax-prosody interface that we assume. We then present our formal account of prosodic Focus marking in English in §4. Finally, in §5, we conclude and highlight directions for future work.

2 Focus in English

In analysing the marking of Focus in English (and other languages) one must make the key distinction between *Extent of Focus* (Foc-Extent) and *Exponent of Focus* (Foc-Exponent). Foc-Extent (also known as the Focus Domain) refers to how much of a sentence can be said to have DF Focus status, while Foc-Exponent is the indication at some level of representation, e.g. syntactic (c-structure) or prosodic (p-structure), of the Focus status of part or all of a sentence. In the examples in (1), the square brackets enclose the syntactic elements that constitute the Foc-Extent, while boldface marks the word that carries the Focus marking (the main stress in the sentence).

(1) a. Q: Who hit Norman?
A: [**Anna**] hit Norman.

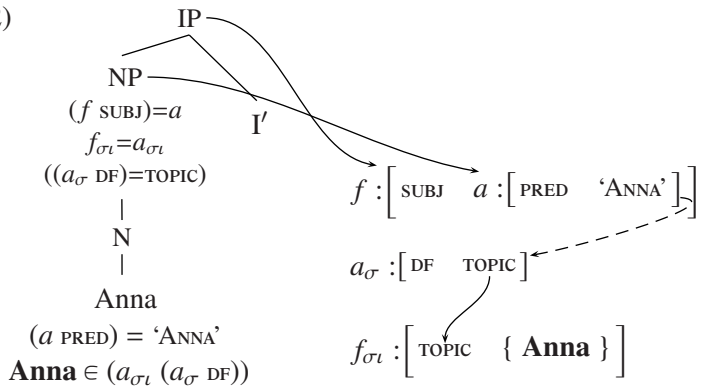
b. Q: Who hit Norman?
A: [Some old **woman**] hit Norman.

In this paper we propose an analysis which captures the relationship between Foc-Extent and Foc-Exponent when the latter is marked only by prosodic means. We confine our discussion to the prosodic encoding of Focus in English, but the analysis we propose can be extended in a straightforward manner to cover prosodic marking of Discourse Functions more generally.[1] For English, on which we concentrate in this paper, we identify the Foc-Exponent as the element in an utterance which bears the main or nuclear stress (see, for instance, Ladd 2008). We identify this element as bearing the Nuclear Tone (though we do not assume that stress need comprise a pitch cue alone). This is the final pitch accent within the relevant domain. This pitch accent is perceived as being the most prominent within the intonational contour under consideration.

[1]More specifically, and for consistency in our analysis, we treat only a single sub-type of Focus, that found in answers to wh-questions ('New Information Focus'). We do not explore the variety of sub-types of Focus which have been posited, e.g. Dik's (1997) taxonomy based on communicative purpose, but such sub-types could be distinguished in our model given a more fine-grained view of Discourse Functions, possibly building on the information feature space as defined in Butt and King (1996) and discussed in Mycock (2013).

The Foc-Extent, on the other hand, refers to how much of a sentence can be said to have Focus status; in more precise terms, the Foc-Extent is the set of syntactic elements that are associated with the DF Focus. The precise definition of Foc-Extent will be dependent on the general approach taken towards information structure and its relation to other aspects of linguistic structure within the grammar. We adopt Dalrymple and Nikolaeva's (2011) approach to information structure within LFG's parallel architecture, according to which elements of a sentence's meaning, i.e. meaning constructors (which appear in bold in (2) and throughout), are categorized according to their DF at s(emantic)-structure and consequently belong to the relevant set (e.g. Topic, Focus) at the level of i(nformation)-structure. In the following English example, based on Dalrymple and Nikolaeva (2011: 83), *Anna* is the subject of a clause, and is also the Topic at i-structure by virtue of occupying SpecIP. Key to this analysis is the attribute-value pair DF TOPIC included in the s-structure for 'Anna', a_σ. This information, combined with the annotations on the terminal and SpecIP nodes in the c-structure, serves to categorize the relevant meaning constructor as belonging to the Topic set in the clause's i-structure, $f_{\sigma\iota}$.

(2)



Given the approach to i-structure exemplified in (2), Foc-Extent is defined as those meaning constructors which are members of the set that is the value of the attribute FOCUS at i-structure. These meaning constructors are semantic units and thus correspond, imperfectly in some cases, to units at other levels of representation, e.g. syntactic elements. We take the position that the correspondence between syntax and prosody is similarly characterized by a lack of isomorphism (see, with references, Dalrymple and Mycock 2011). While it is possible to provide a relatively straightforward informal generalization concerning the relationship between Foc-Extent and Foc-Exponent in English,[2] a formal analysis presents a number of challenges. The extensive misalignment that is a feature of the correspondences between units belonging to different structural levels means that the prosodic encoding of Focus and other DFs represents a complex phenomenon whose formal analysis requires sophisticated modelling of the interfaces between these structural levels.

[2]The Foc-Exponent is associated with the Prosodic Word which corresponds to the rightmost syntactic word that is the syntactic realization of the Foc-Extent.

Central to previous analyses of prosodic Focus marking has been the difference between 'narrow' Focus, which can be characterized as those cases in which the Foc-Extent and Foc-Exponent are a close match (Exponent≈Extent), and 'broad' or 'projecting' Focus, in which the Foc-Exponent correlates with only a part of the Foc-Extent (Extent>Exponent). This distinction can be seen in (1): the answer in (1a) represents 'narrow' Focus, and the answer in (1b) 'broad' Focus.

The distinction between 'narrow' and 'broad' Focus misses an important generalization, however. Based purely on their syntax, the two types of Focus shown in (1) are fundamentally the same: in both instances the Foc-Extent is an XP constituent. Similarly, the basic facts about the Foc-Exponent are the same in both cases (and can be stated as p-structure rules of the type employed in Dalrymple and Mycock 2011): the rightmost Prosodic Word of the Focus constituent bears the Nuclear Tone. With the syntactic and prosodic aspects of Focus marking suitably captured, the challenge is to define how c-structure and p-structure are properly related, both to each other and to i-structure, in cases of prosody-only Focus marking in English. In this paper, we propose an analysis which captures the relevant interactions based on a streamlined version of the syntax-prosody interface in LFG as first introduced in Dalrymple and Mycock (2011).

3 The Architecture

A number of different approaches to prosody and how it should be integrated into the LFG architecture have been proposed, including Butt and King (1998), Mycock (2006), O'Connor (2006), Bögel et al. (2009), Lowe (2011) and Bögel (2012). In this paper, we build on the model of the syntax-prosody interface within LFG proposed in Dalrymple and Mycock (2011). Their strictly modular approach is rooted in the lexicon, with a characterization of lexical entries as comprising (among other things) a s(yntactic)-form and a p(honological)-form. These form values, linked to each other via their association in the lexicon, are the basis for two distinct but related aspects of any string, namely the syntactic string (s-string) and the prosodic or phonological string (p-string, represented using the IPA). Linear precedence relations between atomic syntactic elements are encoded in the s-string and between atomic prosodic elements in the p-string.[3] Parsing of a string therefore involves the association of p-string sequences with p-forms in lexical entries, and concomitant association of p-forms with s-forms to produce the s-string, meaning that within this model the string represents the sole point of interface between syntax and phonology/prosody. The units of the p-string group together to form hierarchical units at the level of p(rosodic)-structure, following the approach to prosodic structure advocated in Lahiri and Plank (2010).

[3]Treating syllables as the minimal prosodic unit is sufficient for the data treated here. Nevertheless there is nothing to prevent the mora being used as the minimal prosodic unit in languages where this is relevant, nor, for example, the addition of a foot level in the Prosodic Hierarchy, which is abstracted away from here as being not strictly necessary for present purposes.

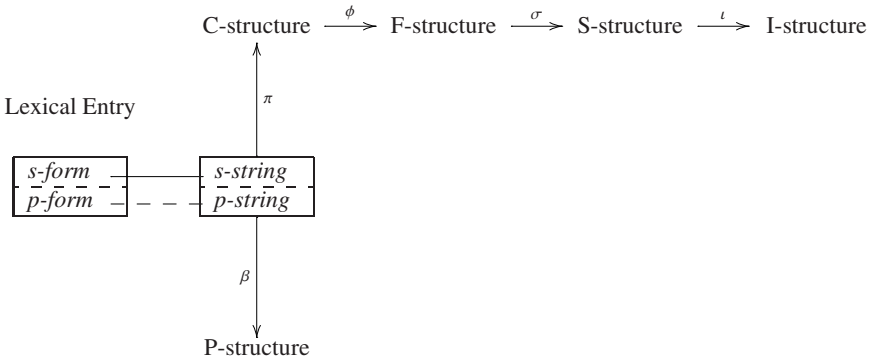
Dalrymple and Mycock propose this approach to the syntax-prosody interface in order to model the contributions that prosody can make to meaning. Their key proposals are that: (i) the 'line of communication' between prosody and semantics is mediated by syntax; (ii) while syntax is effectively blind to objects native to phonology/prosody and vice versa, the edges of syntactic and prosodic constituents represent points of contact between the two separate modules; and (iii) a principle of Interface Harmony exists which requires certain sorts of information associated with constituent edges in distinct modules to match at the string. Information potentially relevant at the interface is 'passed down' the syntactic and prosodic trees by means of separate 'interface structures' – e(psilon)-structure on the syntactic side, chi-structure on the prosodic side – and, as a result, is made available at the interface, namely in the e-structure and chi-structure projections that are associated with the units of the s-string and p-string, respectively. E-structures and chi-structures appear as attribute-value matrices (AVMs) projected from the relevant units. Each interface structure contains a L(eft) and R(ight) attribute whose value is a set whose members represent information that is potentially relevant at other levels of representation, e.g. constituent edge location. For instance, [L {IP}] denotes 'left edge of an Inflection Phrase', and [R {PhP, InP}] denotes 'right edge of a Phonological Phrase and an Intonational Phrase'. Ultimately, this information about constituent edges is associated with the relevant units in the respective aspect of the string (s-string or p-string). Here, at the point of interface, the principle of Interface Harmony applies to enforce alignment as appropriate (though the default is mass misalignment; see Dalrymple and Mycock, 2011 for discussion and references). In this way, the Dalrymple and Mycock model allows for a full investigation of an important issue: which types of information are relevant at the syntax-prosody interface and under which circumstances? We contribute to this research programme by proposing that, just like the semantic contributions discussed by Dalrymple and Mycock (2011), DFs such as Focus are a type of information that should be available at the interface between the two modules, because it is relevant at other levels of representation. However, before we come to formalize the prosodic encoding of Focus in §4, we propose some emendations to the model described in this section which both streamline the architecture and, more importantly, enable it to capture the complexities of Focus marking.

The status of e- and chi-structure in the model of Dalrymple and Mycock (2011) is in certain respects ambiguous. They are, in principle, structures or projections with status equal to that of other structures, yet their purpose is merely to mediate the interface between two other structures. Moreover, these interface structures do not correspond to separate modules of the grammar in the same way that e.g. p-structure, c-structure and s-structure correspond to the separate modules of prosody, syntax and semantics. In particular, the e- and chi-structures projected respectively from c- and p-structure nodes have no independent function: they are purely mediatory, passing information from mother nodes to daughter nodes. Only the e-/chi-structures projected from string elements have any independent (and, indeed, important) function, in mediating the interface, but even these are more

'metastructures' than independent structures per se.

The alterations to the Dalrymple and Mycock (2011) model proposed in this section result in a more streamlined grammatical architecture by eliminating e- and chi-structure, and enable this approach to account for the Focus marking phenomena treated in this paper. At the same time, the underlying principles of the model, including its commitment to strict modularity and the principle of edge-inheritance, are fully maintained. We assume the grammatical architecture given in (3), which is largely identical to that of Dalrymple and Mycock (2011) except for the absence of e- and chi-structure.[4] Note that the units of the s-string are not the terminal nodes of the c-structure. For an example, see (4).

(3)

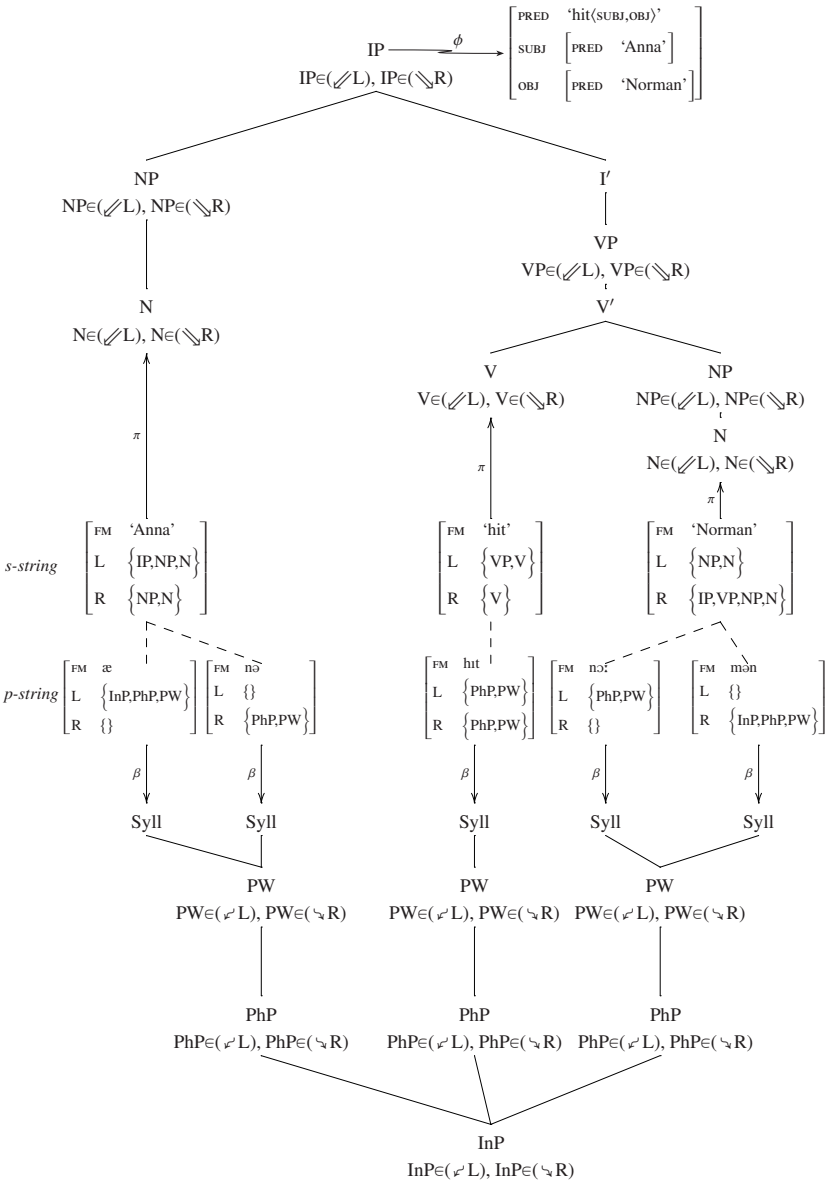


In order to dispense with e- and chi-structure we propose that s- and p-string elements be represented not as atomic elements, but as AVMs. As well as an attribute FM whose value represents the form of the relevant unit in the s- or p-string, each AVM contains among other things the L { } and R { } Attribute-Value (AV) pairs previously allocated to e- and chi-structure. This is illustrated in (4).[5] Thus the AVMs that comprise the s- and p-strings include both information about the content of the string and the interface information associated with the respective

[4]We also specify the relation between the p-string and p-structure, which Dalrymple and Mycock (2011) simply denote with a dashed line, as the mapping β (Mycock 2010: 292).

[5]Following Dalrymple and Mycock (2011), the syntax-prosody interface is represented using a 'double tree' diagram. This type of diagram includes the two aspects of the string, the s-string and the p-string, which are analysed in terms of their hierarchical structure (c-structure and p-structure, respectively). Thus the syntactic analysis is projected from the syntactically parsed 'side' of the string, while the prosodic analysis, the p(rosodic)-structure, is projected from the prosodically parsed 'side'. The syntactic analysis is given above the string in the upside-down c-structure tree, per convention, while the prosodic analysis has its root at the bottom of the diagram, below the string. Although represented together in a double tree, the projections and resulting structures are entirely separate and formed according to their own principles and primitives; i.e. there is nothing syntactic about the bottom half of the tree, nor anything prosodic about the top half. S- and p-string elements are associated via their co-occurrence in lexical entries; the string is parsed by matching prosodic and syntactic units in the string to lexical entries. This 'double tree' diagram is useful in representing the simultaneous analyses of different aspects of the linguistic structure of the utterance concerned, but this is a matter of presentation alone; the analyses belong to different modules of the grammar, they analyse different aspects of the string, and therefore they could equally be represented entirely separately.

(4) *Anna hit Norman*



edges of that unit. This means that the two aspects of the string in this revised model truly represent the sole point of possible interface between the syntax and phonology modules, mediated and constrained by information stored in the lexicon.

So that interface information can be included in the AVMs of the s- and p-strings, we need a mechanism by which the relevant edge information can be passed to string elements. Observe that the passing of edge information to string elements occurs as it were in the 'opposite' direction from the projection from c-structure to f-structure. In the latter case, projection works 'up the tree': usu-

ally, f-structures projected from lower nodes are identical to or included within f-structures projected from higher nodes, and the f-structure projected from the top node of any tree is identical to or contains every f-structure projected from all other nodes in the tree. On the other hand, if we return to the original Dalrymple and Mycock conception of passing edge information in terms of e-structures, there is no single e-structure for any one c-structure, but rather there are as many e-structures as there are c-structure nodes. In contrast to the projection out to f-structure, AV pairs at e-structure which are projected from higher c-structure nodes are incorporated into e-structures projected from lower nodes. This is highly important when it comes to formalizing the passing of edge information. Projection to f-structure, insofar as its direction is 'up the tree', requires reference only to the current node and its mother, symbolized as $*$ and $\hat{*}$. The f-structures corresponding to these nodes are obtained by applying the function ϕ, giving $\phi(*)$ and $\phi(\hat{*})$, regularly abbreviated as $\downarrow$ and $\uparrow$ respectively. But since the passing of edge information to the string works 'down the tree', as it were, these familiar annotations are insufficient; specifically, we require reference not to the current node and its mother, but to the current node and its (leftmost and rightmost) daughters. We define a relation D, which finds the set of daughter nodes of the current node; we then represent the leftmost and rightmost immediate daughters of $*$ by the functions $D_l(*)$ and $D_r(*)$ respectively, defined as in (5). The leftmost and rightmost *terminal* nodes projected from $*$, which we represent by the functions $T_l(*)$ and $T_r(*)$ respectively, are defined in (6).[6] Finally, the s-string elements corresponding to these terminal nodes are simply obtained by applying the function π^{-1}. In the spirit of $\uparrow \equiv \phi(\hat{*})$, we propose the abbreviations in (7).

(5) a. $D_l(*) \equiv$ node n, where $n \in D(*) \wedge \neg\exists x.x \in D(*) \wedge x \prec n$.
b. $D_r(*) \equiv$ node n, where $n \in D(*) \wedge \neg\exists x.x \in D(*) \wedge x \succ n$.

(6) a. $T_l(*) \equiv \begin{cases} * \text{ if } D(*) = \varnothing \\ \text{else } T_l(D_l(*)) \end{cases}$ b. $T_r(*) \equiv \begin{cases} * \text{ if } D(*) = \varnothing \\ \text{else } T_r(D_r(*)) \end{cases}$

(7) a. $\swarrow \equiv \pi^{-1}(T_l(*))$ b. $\searrow \equiv \pi^{-1}(T_r(*))$

These arrows, then, permit direct reference from any c-structure node to the s-string elements corresponding to the leftmost and rightmost terminal c-structure nodes that are descendants of the node in question.[7] We can therefore pass syntactic category information to string elements by means of simple rules such as those in

[6]These rules can be informally read as follows: the leftmost/rightmost terminal node from the current node is the current node if the current node has no daughters, else it is the leftmost/rightmost terminal node from the leftmost/rightmost (respectively) daughter of the current node. The rule will apply recursively to find the appropriate terminal descendant from any node.

[7]Note that this model permits simple reference only to constituent edges. In principle it may be possible to make reference to constituent-internal elements, but given the edge-oriented nature of prosody we believe that in general it is edges that will be crucial in the analysis of interface phenomena such as Focus marking.

(8). These rules are given under the nodes in (4) in order to be explicit, but we assume they can be stated as more general principles, and we omit them from subsequent trees.

(8) a. For any XP, XP $\in$($\swarrow$L) and XP $\in$ ($\searrow$R)

b. For any X, X $\in$($\swarrow$L) and X $\in$ ($\searrow$R)

We can do precisely the equivalent, of course, for the passing of information down the prosodic structure to the p-string. Using $\diamond$ for the current node in the p-structure, the functions $\swarrow$ and $\searrow$, equivalents for syntactic $\swarrow$ and $\searrow$, can be defined along entirely parallel lines, that is:

(9) a. $\swarrow \equiv \beta^{-1}(T_l(\diamond))$ b. $\searrow \equiv \beta^{-1}(T_r(\diamond))$

This will allow us to pass prosodic category (and other) information directly from nodes in the p-structure to the relevant p-string elements (syllables, cf. fn. 3), e.g. for any PhP, PhP $\in$($\swarrow$L) and PhP $\in$ ($\searrow$R), and so on. In order to account for Focus marking, we also need to be able to refer to the rightmost and leftmost p-string elements in a prosodic projection *that are marked for primary stress*, that is the leftmost and rightmost syllables within a projection that are specified as being the location of primary stress (represented as the p-string feature SYLLSTRESS p).[8,9]

(10) a. $\swarrow^s \equiv \beta^{-1}(T_{ls}(\diamond))$ b. $\searrow^s \equiv \beta^{-1}(T_{rs}(\diamond))$

(11) a. $T_{ls}(\diamond) \equiv \begin{cases} T_l(\diamond) \text{ if } (\beta^{-1}(T_l(\diamond))\text{SYLLSTRESS}) = \text{p} \\ \text{else } T_{ls}(\beta(N(\beta^{-1}(T_l(\diamond))))) \end{cases}$

b. $T_{rs}(\diamond) \equiv \begin{cases} T_r(\diamond) \text{ if } (\beta^{-1}(T_r(\diamond))\text{SYLLSTRESS}) = \text{p} \\ \text{else } T_{rs}(\beta(N^{-1}(\beta^{-1}(T_r(\diamond))))) \end{cases}$

In this way, potentially relevant information can be passed to s-string and p-string elements directly, without the need for additional structures, such as Dalrymple and Mycock's e-structure and chi-structure, to mediate the passing.[10]

[8]In fact, in the examples discussed below, there is only ever one stressed syllable in the relevant prosodic projections, such that it might have been possible to define only a single arrow under (10); nevertheless we define both on the grounds that it is possible for more than one stressed syllable to appear in some prosodic projections and Focus marking seems to be edge-oriented cross-linguistically. We utilize only $\searrow^s$ in the rules below, since English Focus marking seems to be consistently oriented to the right edge, but using $\swarrow^s$ would have made no difference here, at least.

[9]The function N, which appears in (11), finds the next element in linear order when applied to string elements, as defined in Asudeh (2009: 111); N^{-1} finds the preceding element.

[10]It is possible to retain e- and chi-structure in the current model, but as structures projected only from string elements. The conceptual justification for this would be that Left and Right edge information is not information about string elements, but information associated with string elements,

4 Formalizing Focus

In formalizing the prosodic marking of Focus in English, the parallel architecture of LFG, comprising distinct but related levels of linguistic representation, requires us to simultaneously provide an analysis purely in terms of prosody and an analysis purely in terms of syntax. This means that there will be a rule or rules treating the relationship between a unit or units at p-structure and the DF Focus at i-structure; this covers the Foc-Exponent (§4.1). Similarly, there will be a rule or rules treating the relationship between a unit or units at c-structure and DF Focus at i-structure; this covers the Foc-Extent (§4.2). The principle of Interface Harmony will ensure that the appropriate relationship exists between the Foc-Exponent, corresponding to some prosodic cue(s), and the Foc-Extent, corresponding to some syntactic unit(s). In this way, our analysis truly models prosodic marking of Discourse Functions as an interface phenomenon, sited at the one point in the grammar – the string – where the phonology and syntax macromodules are in contact.

4.1 Foc-Exponent

We assume the rule in (12) for declarative intonation in English. The rightmost Phonological Phrase (PhP) in an Intonational Phrase (InP) is specified as having (in default cases) N_TONE = H (a Nuclear Tone whose value is High) associated with its rightmost stressed syllable, and RB_TONE = L (a Right Boundary Tone whose value is Low) associated with its rightmost syllable. These features are related to the label *DeclSem* that appears in the relevant p-string AVM and is required to interface with an equivalent property in the corresponding s-string AVM, which will mark the sentence as semantically declarative. The rule for Focus presented in (13) will apply alongside the rule in (12), since the type of Focus we are discussing involves declaratives. (We include the *DeclSem* label for the sake of completeness, but do not otherwise utilize it here.)

(12) InP → PhP* PhP
$((\searrow^{s}\text{N_TONE}) = \text{H})$
$(\searrow\text{RB_TONE}) = \text{L}$
$\textit{DeclSem} \in (\searrow\text{R})$

It is possible to formulate a single p(rosodic)-structure rule (13) to account for nearly all types of intonational Focus marking. This rule states that any Prosodic Word (PW) in a Phonological Phrase may be marked for N_TONE = H, and if so, the label *DF_Focus* appears as a value of the attribute R of the corresponding p-string

such that it should not really be represented in the string AVMs (which we would retain in any case), but in AVMs associated with string elements. If such a model were pursued, all of the above equations and symbols could remain, but with the addition of projection out to e- and chi- structure from the s- and p-string elements respectively in the definitions of $\swarrow$, $\searrow$, $\nwarrow$ and $\searrow$; so e.g. $\swarrow \equiv \epsilon(\pi^{-1}(T_r(*)))$, etc. For simplicity we do not make use of e- and chi-structure, but we leave it an open question which representation is conceptually and architecturally most desirable.

element that carries the Nuclear Tone.[11] The p-string element in which this feature appears is, specifically, the rightmost syllable that bears primary stress ($\searrow^{s}$) in the relevant PW.[12]

(13) PhP → PW* PW PW*
$((\searrow^{s}{}_{\text{N_TONE}}) = \text{H}) \Rightarrow$
$DF_Focus \in (\searrow^{s}{}_{\text{R}})$

The Foc-Exponent serves to delimit the Foc-Extent. Foc-Extent is captured by separate c-structure rules which, together with the p-structure rule in (13) and the principle of Interface Harmony, play an equally important role in the analysis of prosodic marking of Focus as an interface phenomenon. We now turn to these c-structure rules.

4.2 Foc-Extent

As discussed in §2, most examples of Focus in English, including not only Extent>Exponent Focus (broad or projecting Focus), but also many examples of Exponent≈Extent Focus (narrow Focus), correlate to an XP at c-structure. It is also possible for the Foc-Extent to correlate with an X′, although examples are harder to come by (14–15).

(14) V′ Focus:
Q: What did Anna do for Norman?
A: Anna [bought **fudge**] for Norman.

(15) N′ Focus:
Q: What/Which blue thing did you find?
A: I found a blue [piece of **pa**per].

We therefore assume that in principle any phrasal category at c-structure may delimit the Foc-Extent in a clause. The PS rule in (16) licenses this: any non-terminal node in the syntactic structure may be classified as in Focus, i.e. its semantic structure may include the feature DF FOCUS. If this is the case, the label DF_Focus is specified as a value of the attribute R of the s-string element corresponding to the rightmost terminal daughter of the phrasal category.[13]

[11] It is important to note that the value *DF_Focus* is simply a label associated with a prosodic feature, and does not represent semantic information within the prosodic component. The fact that we call the label *DF_Focus* and not, say, *xyz*, is intended to make the requirement of Interface Harmony – that *DF_Focus* in the p-string match with (the equally meaningless) DF_Focus in the s-string – as clear and obvious as possible, and nothing more. This does not, therefore, conflict with our stated aim of absolute modularity. The same applies, of course, to other labels such as *DeclSem* in (12) and the labels for constituent edges (PW, NP, etc.) used in Dalrymple and Mycock (2011). Interface Harmony requires label matching, which will have consequences; it does not involve storage or 'passing' of semantic or any other type of information.

[12] It is sufficient for the present purposes that we specify and require harmony for only the R features in the s-string and p-string, which reflects the fact that Focus marking in English is fundamentally right-edge based. See also Dalrymple and Mycock (2011) on declarative questions, in which the crucial meaning constructor is similarly proposed to be right-edge based.

[13] We distinguish labels on the prosodic side from labels on the syntactic side by giving the former

(16) C-structure rule for phrasal Focus:

$$\Sigma \rightarrow \Sigma^* \qquad \Sigma \qquad \begin{pmatrix} (\uparrow_\sigma \text{DF}) = \text{FOCUS} \\ \text{DF_Focus} \in (\diagdown_R) \end{pmatrix}$$

Due to the principles governing the passing of semantic information through syntax in LFG, it is not in fact enough to state that the XP concerned is marked for Focus at s-structure. The f-structure projected from an XP is identical to or contains the f-structures related to its subordinate nodes at c-structure, but the corresponding semantic structure projected from the XP's f-structure does not include s-structures projected from subordinate f-structures.[14] The rule in (17) applies alongside that in (16), extending the feature DF FOCUS to the semantic structures projected from any node that is a daughter of a node which is itself marked for Focus.

(17) C-structure rule for cascading Focus:

$$\Sigma \rightarrow \qquad \Sigma^* \qquad \left\{ (\uparrow_\sigma \text{DF}) \neq \text{FOCUS} \;\middle|\; \begin{matrix} (\uparrow_\sigma \text{DF}) =_c \text{FOCUS} \\ (\downarrow_\sigma \text{DF}) = \text{FOCUS} \end{matrix} \right\}$$

The implementation of phrasal Foc-Extent by means of these rules, and the prosodic structure rule given above, is exemplified in (25) and (26) in the Appendix.[15]

It is also possible, though rare, that the Foc-Extent may correlate with an X^0, i.e. non-phrasal, category at c-structure. The clearest example of X^0 Foc-Extent is Exponent≈Extent verb Focus, as in example (18a). Here, the complement to the verb, *Norman*, is backgrounded and therefore not included in the Foc-Extent, meaning that in syntactic terms only the V^0 constituent is focused.[16] In (18b) the

in italics, hence *DF_Focus* in (13) but DF_Focus in (16). This is purely for the purpose of clarity in representing the fact that these labels relate to distinct aspects of the grammar, one to prosody, the other to syntax.

[14]This is necessary to overcome the granularity problem; see King (1997).

[15]Note that the example analysed in (26) would be problematic under an approach that defined Foc-Exponent in terms of PWs because in this case the relevant PW also includes an enclitic (related via the lexicon to the s-string unit *for*) which does not constitute part of the Foc-Extent.

[16]Verb Focus is somewhat more complicated than simple XP Focus. In question-answer pairs such as we have here, the event (EV) element of a verb's meaning is presupposed in the question, and hence is not focused in the answer; rather, what is questioned, and hence focused in the answer, is the *type* of event (REL) that occurred, see Mycock (2006). We abstract away from the formal details of this here (details which will in fact apply in all cases where a verb is focused, including, for example, IP focus). In formal terms, distinct meaning constructors will correspond to the EV and REL elements of a verb's meaning, and only the latter will appear as a value of FOCUS at i-structure. We believe this is pragmatically determined, i.e. the EV element is excluded from Focus specifically because it is presupposed in the preceding discourse. While some work has been done on cross-sentential discourse structure in LFG (King and Zaenen 2004, Gazdik 2011), a simple integration with Dalrymple and Nikolaeva's i-structure, which would permit us to account for the issue under discussion, is beyond the scope of the present paper. We therefore set this problem aside for present purposes and assume that, when a verb is focused, all its associated meaning constructors are focused, just like all other categories of words.

Foc-Extent includes only the D^0 head/specifier of the verb's complement.

(18) a. Q: What did Anna do to Norman?
A: Anna [**hit**] Norman.

b. Q: Which car did you damage?
A: I damaged [**your**] car.

It is difficult or impossible to focus other X^0 categories when only New Information Focus is considered. However, it is clear that for other types of Focus any X^0 can function as the Foc-Extent, e.g. when contrast is involved as in *This letter is **[to]** my sister (not from her)*.[17] We therefore frame a phrase-structure rule that permits X^0 Foc-Extent of any sort (19); X^0 Foc-Extent is exemplified in (27) in the Appendix.

(19) C-structure rule for X^0 Focus:

$$\Sigma \rightarrow \Sigma^* \quad X \quad \Sigma^* \qquad \begin{pmatrix} (\downarrow_\sigma \text{DF}) = \text{FOCUS} \\ \text{DF_Focus} \in (\diagdown_R) \end{pmatrix}$$

4.3 Single-syllable Exponent>Extent

Examples such as the ones in (20) appear to represent Exponent>Extent Focus, that is cases in which the Foc-Extent corresponds to only a part of the Foc-Exponent. This appears problematic for our model, since we define the Foc-Exponent as the stressed syllable, which here appears to correspond to two elements in the s-string, only one of which is semantically focused.

(20) a. Q: Who's gone?
A: [**Kay**]**'s** gone.

b. Q: Who'll help?
A: [**I**]**'ll** help.

Instances of this phenomenon in English are in fact rare and lexically restricted (which may account for the lack of attention paid to them in the literature): they prototypically involve reduced auxiliaries. The restricted inventory of forms involved means this can be treated most efficiently via a lexical restriction. For example, we assume the lexical entry in (21) for *'s* (= *has*), seen in (20a). In (21), $\bullet$ refers to any s-string element with which the lexical entry is associated; $\bullet_\pi$ therefore refers to the c-structure exponent of the corresponding s-string element, and $\bullet_{\pi\phi}$ to the corresponding f-structure element. $\blacklozenge$ refers to any p-string element with which the lexical entry is associated. As stated in §3, the feature FM represents the form of the unit in the s-/p-string.

[17] We do not consider here cases such as *No, don't retie it, **un**tie it!*; as argued by Vallduví and Engdahl (1996: pp. 504–505) these are best treated as metalinguistic corrections, rather than instances of the type of Focus phenomena analysed in this paper.

(21)

(●FM) = ''S'
$\bullet_{\pi}$ = I
($\bullet_{\pi\phi}$ TENSE-ASPECT) = PRES-PERF
DF_Focus ∉(●R)
(◆FM) = /Z/

This lexical entry prevents DF_Focus from being a value of the attributes L and R in an s-string exponent of the auxiliary *'s* (= *has*). In terms of Interface Harmony, in cases where *'s* is part of a stressed syllable which is the Foc-Exponent, this will enforce the appearance of DF_Focus as a value of L and R in the s-string element directly preceding *'s*, meaning that only the word preceding the auxiliary, and not the auxiliary itself, will have an s-structure that contains the AV pair DF FOCUS.[18] This is illustrated in (28) in the Appendix. In this example, the Foc-Exponent is /keɪz/ in the p-string, so this syllable's R attribute-value set includes *DF_Focus*. According to Interface Harmony as applied thus far, this R value should match with the R value of either (or perhaps both) of the corresponding s-string units, which are *Kay* and *'s*. A match with the s-string unit for *Kay* is unproblematic, and is in fact what we want, but a match with the s-string unit for *'s* must be excluded: *'s* (= *has*) is not part of the Foc-Extent. The lexical entry in (21) deals with this issue by stating (in line 4) that in no case can DF_Focus appear as a member of the attribute-value set R in an s-string unit associated with the lexical entry for *'s*. Therefore there is only one possibility that will satisfy Interface Harmony in (28), namely to have the R *DF_Focus* AV pair in the p-string unit for /keɪz/ match with an R DF_Focus AV pair in the s-string unit for *Kay*. Therefore in (28), DF_Focus is necessarily a member of the R attribute-value set in the AVM for *Kay* in the s-string.

5 Further Issues and Conclusion

In this paper we have concentrated on the formal mechanisms of the syntax-prosody interface and the specifics of modelling prosodic Focus marking in general terms. In doing this we have abstracted away from some more complex data. We noted above (fn. 16) that the technical details of verb Focus are in fact more complex than we have assumed for the present purposes. Similarly, we have intentionally avoided treating examples of non-edge-marked Focus (22–23), and examples of discontinuous Foc-Extent (24).

(22) Q: What did Anna do?
A: Anna [put the **cat** out].[19]

(23) Q:What happened?
A: [Anna **hit** me].

[18]Note that it is still possible for *'s* to appear within a Foc-Extent, e.g. within a focused IP, since the restriction is on s-string labels, not i-structure categorization.

[19]It is not the case that a particle like *out* cannot be the Foc-Exponent. Indeed, if the answer were *Anna put it **out***, as in a reply to *What did Anna do with the cat?*, this would be the default production pattern. Any future analysis will have to be able to account for these two possibilities.

(24) Q: What did Charlie do with the cake?
A: He [put] it [in the **cup**board].

A more comprehensive analysis of Focus marking in English must, of course, account for the entire range of data. As noted above this requires some reference to inter-sentential discourse relations, which is beyond the scope of the present paper. We have shown that it is relatively simple to model many examples of Focus marking in English in the architecture proposed here, and it is only a more sophisticated articulation of the rules governing English Focus marking that would be required to enable it to model the entire range of data. We contend that this also holds for extending this approach to the prosodic marking of DFs in other languages as well. We have not attempted this here as we seek to emphasize and explore the architectural issues. Our aim has been the demonstration that it is possible to model prosodic contributions to information structure in the streamlined version of Dalrymple and Mycock's (2011) architecture presented here, representing this challenging phenomenon while maintaining strong modularity of the grammatical architecture.

References

Asudeh, A. (2009). 'Adjacency and Locality: a Constraint-Based Analysis of Complementizer-Adjacent Extraction'. In M. Butt and T. H. King (eds.), *Proceedings of the LFG09 Conference*, CSLI Publications, pp. 106–126.

Bögel, T. (2012). 'The P-Diagram – A Syllable-Based Approach to P-Structure'. In M. Butt and T. H. King (eds.), *Proceedings of the LFG12 Conference*, CSLI Publications, pp. 99–117.

Bögel, T., M. Butt, R. M. Kaplan, T. H. King, and J. T. Maxwell III (2009). 'Prosodic Phonology in LFG: A New Proposal'. In M. Butt and T. H. King (eds.), *Proceedings of the LFG09 Conference*, CSLI Publications, pp. 146–166.

Butt, M. and T. H. King (1996). 'Structural Topic and Focus without Movement'. In M. Butt and T. H. King (eds.), *Proceedings of the LFG96 Conference*, CSLI Publications.

Butt, M. and T. H. King (1998). 'Interfacing Phonology with LFG'. In M. Butt and T. H. King (eds.), *Proceedings of the LFG98 Conference*, CSLI Publications.

Dalrymple, M. and L. Mycock (2011). 'The Prosody-Semantics Interface'. In M. Butt and T. H. King (eds.), *Proceedings of the LFG11 Conference*, CSLI Publications, pp. 173–193.

Dalrymple, M. and I. Nikolaeva (2011). *Objects and Information Structure*. Cambridge University Press.

Dik, S. C. (1997). *The Theory of Functional Grammar* (2 vol.). Mouton de Gruyter, second edition. Edited by K. Hengeveld.

Gazdik, A. (2011). *Multiple Questions in French and Hungarian: A Lexical-Functional Analysis with Special Emphasis on the Syntax-Discourse Interface*. Ph.D. thesis, Université Paris Diderot - Paris 7.

King, T. H. (1997). 'Focus Domains and Information-Structure'. In M. Butt and T. H. King (eds.), *Proceedings of the LFG97 Conference*, CSLI Publications.

King, T. H. and A. Zaenen (2004). 'F-structures, Information Structure, and Discourse Structure'. In M. Butt and T. H. King (eds.), *Proceedings of LFG04*, CSLI Publications.

Ladd, D. R. (2008). *Intonational Phonology*. Cambridge University Press, second edition.

Lahiri, A. and F. Plank (2010). 'Phonological phrasing in Germanic: the judgement of history, confirmed through experiment'. *Transactions of the Philological Society 108*(3), pp. 370–398.

Lowe, J. J. (2011). 'Ṛgvedic Clitics and 'Prosodic Movement''. In M. Butt and T. H. King (eds.), *Proceedings of the LFG11 Conference*, CSLI Publications, pp. 360–380.

Mycock, L. (2006). *The Typology of Constituent Questions : A Lexical-Functional Grammar analysis of 'wh'-questions*. Ph.D. thesis, University of Manchester.

Mycock, L. (2010). 'Prominence in Hungarian: The Prosody-Syntax Connection'. *Transactions of the Philological Society 108*(3), pp. 265–297.

Mycock, L. (2013). 'Discourse functions of question words'. In M. Butt and T. H. King (eds.), *Proceedings of the LFG13 Conference*, CSLI Publications.

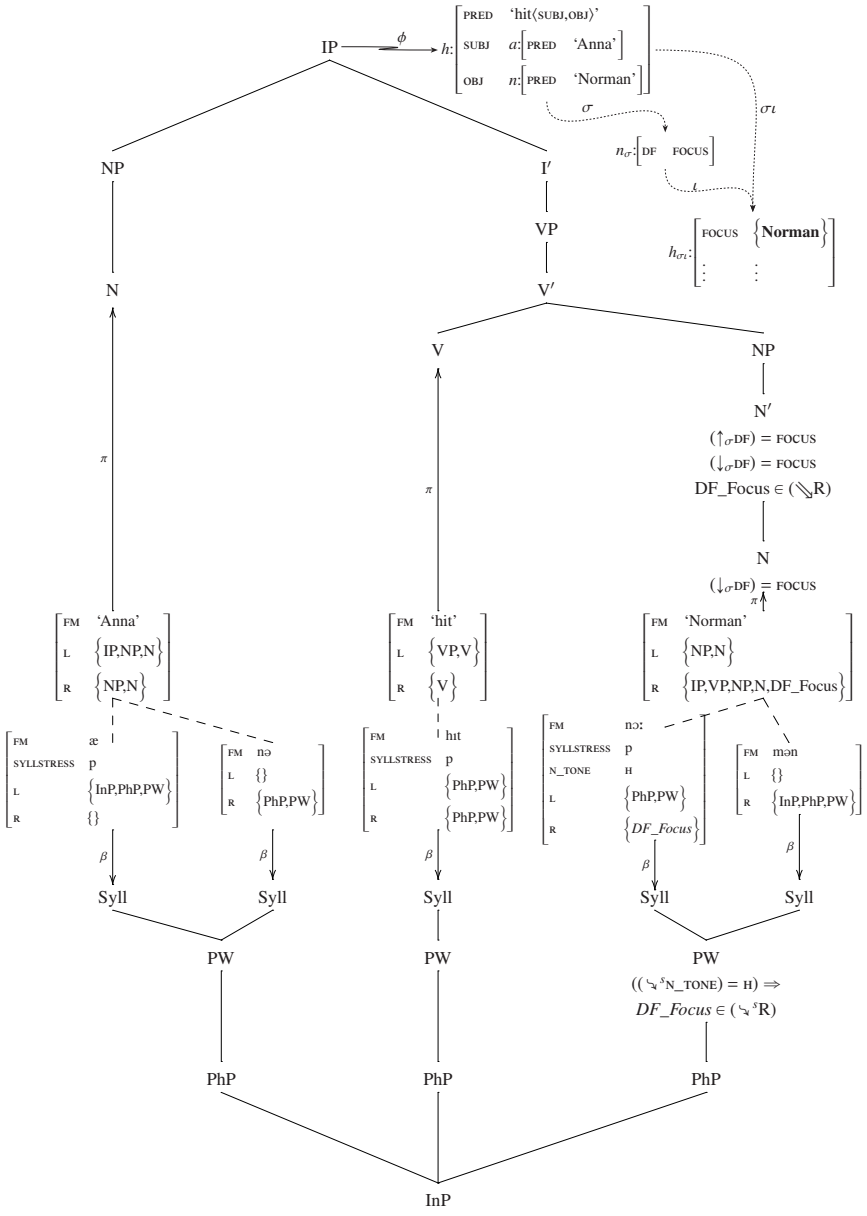
O'Connor, R. (2006). *Information Structure in Lexical-Functional Grammar: The Discourse-Prosody Correspondence*. Ph.D. thesis, University of Manchester.

Vallduví, E. and E. Engdahl (1996). 'The linguistic realization of information packaging'. *Linguistics 34*(3), pp. 459–519.

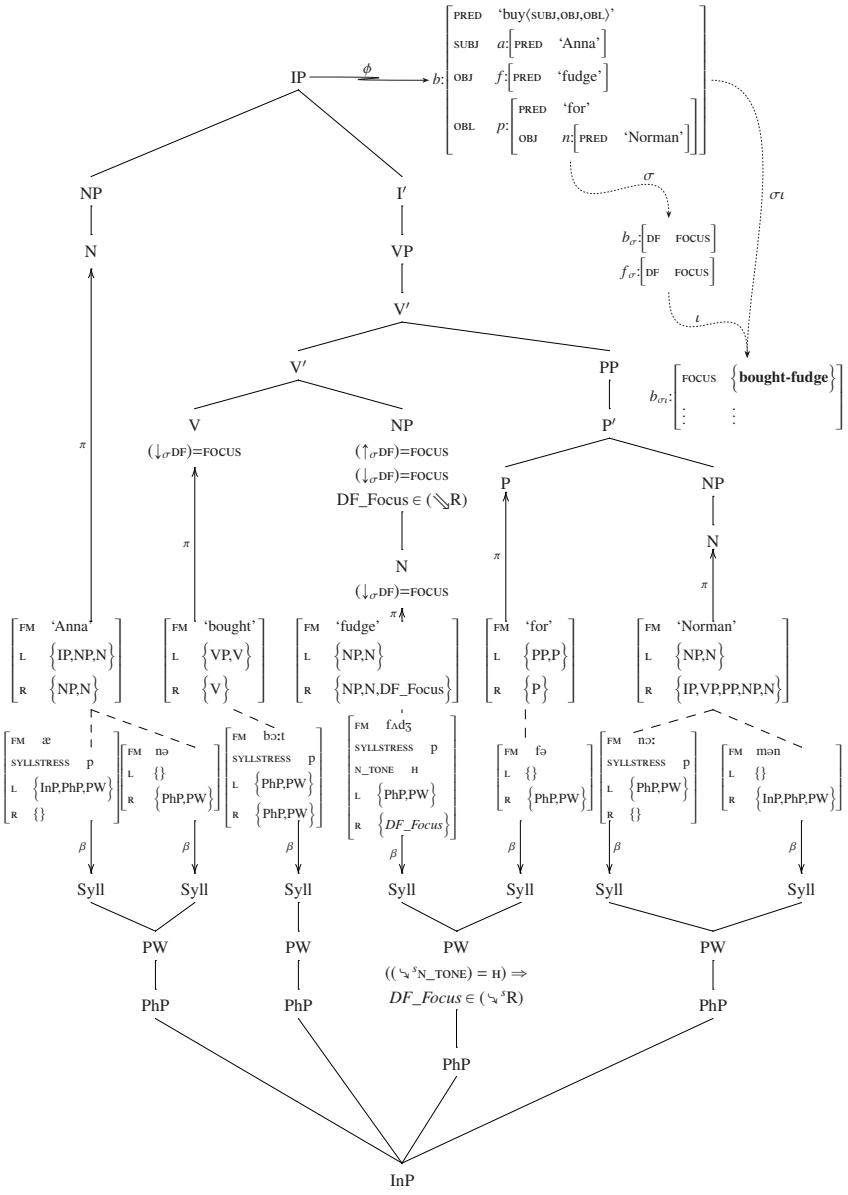
Appendix: 'double-tree' analyses of prosodic Focus marking in English

The examples in this Appendix show utterances spoken at a regular tempo. As Dalrymple and Mycock (2011) point out, p-structure (and the extent of its alignment with c-structure) can vary depending on a number of factors including speech tempo, even in the case of a single sentence. While the diagrams included in this Appendix depict cases in which there is a one-to-one relationship between PWs and PhPs, this need not be the case; our analysis can be extended in a straightforward manner based on the proposals in Dalrymple and Mycock (2011), which allow for such variation to be captured. At syllable level, the stress indicated in a p-string AVM indicates the location of primary prominence in the relevant prosodic word.

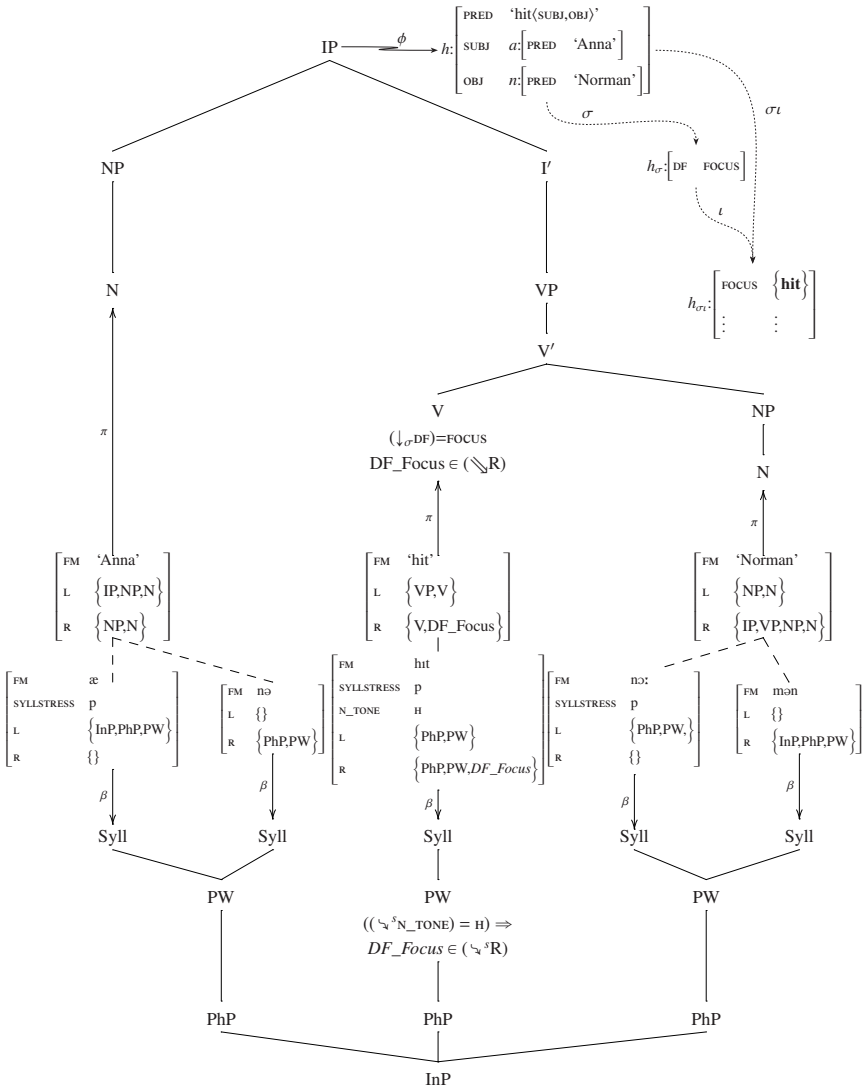
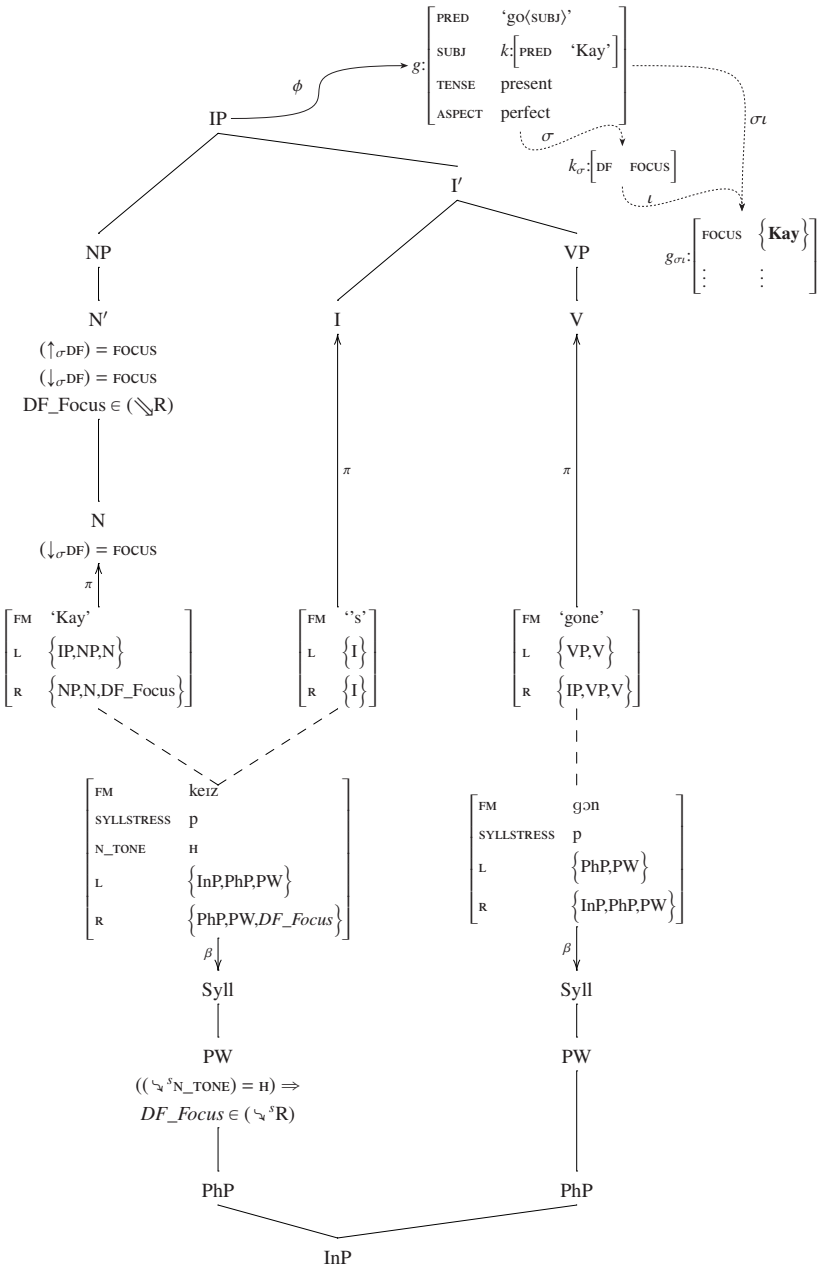
(25) XP Focus: *Anna hit [**Nor**man]*



(26) X′ Focus: *Anna [bought **fudge**] for Norman*



(27) X^0 Focus: *Anna **[hit]** Norman*



(28) Single-Syllable Exponent>Extent: ***[Kay]'s*** *gone*