Information Structuring begins with the Numeration[1]

Enoch O. Aboh, ACLC, Universiteit van Amsterdam, e.o.aboh@uva.nl

1. Introduction

In this paper, I argue for what seems to me the most natural approach to information structure, and, which in the end, appears an inescapable conclusion if one adopts the strict minimalist approach to the study of language (Chomsky 1995, 1999, 2005). In current minimalist theory, a linguistic expression is a pair (π, λ) consisting of a PF representation (i.e., sound) associated to an LF representation (i.e., a meaning). Under this view, the role of the computational system C_{HL} is to map some array of lexical choices to the pair (π, λ). The array of lexical choices is referred to as the numeration (N), which Chomsky (1995: 225) defines as:

(1) "a set of pairs (LI, *i*), where LI is an item of the lexicon and *i* is its index, understood to be the number of times that LI is selected".

Each time a lexical item is selected from the numeration (N), its index is reduced by one so that the converging derivation (i.e., the one that forms a linguistic expression) is the one for which N is reduced to zero. Given this description, it seems obvious that there should be nothing in the pair (π, λ), for instance with regard to its basic (compositional) semantics, that is not already part of the numeration (N). As often assumed in the literature without further demonstration, this would imply that additional discourse information (e.g., topic and focus)

[1] This paper has been presented in various forms at the MPI conference on Information Structure in Adult and Child Language, Nijmegen, at the Syntax Circle, Meertens Institute, and at the Syntax seminar, University of Groningen. I thank all the participants of these meetings for their reactions. I'm also grateful to Boban Arsenijević, Anna Cardinaletti, Michel DeGraff, Marina Dyakonova, Irene Heim, Luigi Rizzi, Jenneke van der Wal, and Jan Wouter Zwart for their comments.

that discourse participants assign to a linguistic expression, and which is not part of the numeration, is invisible to C_{HL}. Therefore, adding these features as the derivation proceeds would violate the inclusiveness condition (Chomsky 1995).

While one can live a happy life with this conclusion, it is still an empirical fact that, in many languages of the world, notions of information structure such as Topic, Focus, and Interrogative force are determined by lexical choices that are manipulated by C_{HL} in the course of various syntactic operations (e.g., feature matching, displacement). In these languages therefore, the numeration N of a sentence containing a Topic, a Focus or an Interrogative expression (π, λ) must include Topic, Focus, or Interrogative lexical choices. This leads me to the formulation in (2).

(2) A numeration N pre-determines the Information Structure of a linguistic expression.[2]

If we take languages where lexical items encode information structure non-ambiguously not to be too exotic or dramatically different from English, generative comparative syntax teaches us that (2) should generalize to all languages in a uniformitarian way (see Chomsky 1999: 2 for discussion of the uniformity principle). Languages therefore only vary with regard to the various choices they make in implementing (2).[3] Taking this

[2] In this paper, I only consider Interrogative, Topic, and Focus, the cornerstone of Information Structure, which involves non-ambiguous morphological expressions correlating with distinct semantic properties.

[3] The alert reader may wonder why we should extend the formulation in (2) to all languages given that it is apparently empirically motivated for certain languages only. Therefore in languages where there are no distinct markers for topic, focus, and question, one may assume that no corresponding functional structure is needed for these languages. Anticipating the discussion in section 5 where I propose that topic and focus features are optional formal features comparable to Case, I would like to stress that decades of comparative syntax should guide our steps here. In his (1977) personal letter to Lasnik and Chomsky, Vergnaud made the crucial point that even languages with virtually no case morphology pattern with case morphology languages with regard to the distributive properties of cased marked NPs. This observation later on inspired Chomsky's theory of abstract case that is now commonly accepted within generative grammar (but see Marantz 2000 and much related work for an alternative). Yet, if we consider languages like Gungbe, to which this theory extends, they show no case morphology at all on full DPs (i).

(i) a. Àjàkà dù gbàdó
mouse eat corn
'A/the mouse ate corn'

b. Àsé dù àjàkà
cat eat mouse
'A/the cat ate the/a mouse'

stance further implies that core syntax embeds properties of Information Structure. In terms of Rizzi (1997) and much related work, this means that core syntax must involve information sensitive functional projections (e.g., TopP, FocP, InterP). While not revolutionary, this controversial position must be motivated both empirically and theoretically. This paper tries to address both issues on minimalist grounds. Section 2 briefly introduces the traditional debate on the interaction between Information Structure and syntax and concludes that there is a certain bias in the literature in considering Topic and Focus as pragmatic notions, while other comparable expressions (e.g., Interrogative) are seen as proper formal features encoded in syntax (i.e., in C, Chomsky 1995: 240). This paves the way for sections 3 and 4 where I present strong empirical evidence for (2) at the CP and DP level, respectively. Section 5 discusses the implications of (2) with regard to the Topic feature and the Focus feature as proper formal features comparable to tense, Case, and ϕ-features. Section 6 concludes the paper

2. Information Structure and Syntax

It is a well-known fact that Information Structure affects syntax in some particular ways, though the interaction between these two modules is still not well understood.

2.1. The Debate

Aside from cleft constructions, which I'm not discussing here, examples that one often finds in the literature concern Left or Right Dislocations (3a-b) or Focus movement constructions

---

The only traces of case that we find are confined to a corner of the pronominal system (i.e., 1sg and 2sg as in *ùn* 1sg-Nom versus *mì* 1sg-Acc, or *à* 2sg-Nom versus *wè* 2sg-Acc, see Aboh 2004, chapter 4.). Jamaican creole and several creoles exhibit no case distinction at all, and the same forms and set of pronouns occur both in nominative and accusative positions. Despite their complete lack of case morphology, the theory of abstract case naturally extends to these languages as well since they manifest similar set of restrictions on cased marked NPs. Following this tradition therefore, I assume that it is just as natural to extend the theory of information structure proposed here to even languages with little (or no) morphology for expressing information structure.

(3c). In addition to word order, these sentences have specific intonation patterns that distinguish them from the 'neutral' statement in (4).

(3) a. John, I like him very much [LD: Topic reading]

b. I like him very much, John [RD: Topic reading]

c. John I like very much [LD: Focus reading]

(4) I like John very much [Basic order: neutral or default new information focus]

If we ignore, for the moment, the various movement operations (i.e., the so-called surface effect) that may have happened in (3a) to (3c), it appears that the numeration leading to the derivations in (3a-c) and (4) is virtually the same and we predict it should lead to roughly the same meaning in all contexts, namely: [X likes Y very much]. Under this view, the fact that the DP *John* may front to check some EPP feature under C follows from general UG principles and is not intrinsic to the lexical entry *John* (Chomsky 1995: 236). Given this description, the fact that, in actual discourse, the sentences in (3a-c) acquire additional meaning in relation to the displaced categories (e.g., topic, focus) must be explained, but the process producing this discourse-related meaning cannot be directly ascribed to syntax (i.e., the computation that produces the phrase marker).

In this regard, Chomsky (1995: 220) suggests that the observed surface effects (e.g., topic-focus articulation) "seem to involve some additional level or levels internal to the phonology component, postmorphology but prephonetic, accessed at the interface along with PF and LF". Put differently, such effects are extraneous to aspects of the computation that are imposed by the numeration (i.e., the N→λ computation). Similar approaches are found in recent work by several researchers (e.g., Zubizaretta 1998, Szendroi 2001, Fanselow 2006, 2007) who assume that constituent rearrangements as in (3) and (4) are prosodically-driven

and therefore differ from displacement rule triggered by checking operations. Under such a view, information structure and syntax interact, but indirectly (via PF and presumably LF).

2.2. On the inclusiveness condition

Indeed, when we consider the examples under (3), there appears to be no clear (context-invariant) property of the alleged Topic and Focus features driving the fronting operations which in turn correlate with specific meanings (e.g., topic vs. focus) acquired by the displaced phrase. As often suggested in the literature, such displacements rather correlate with stress assignment in Germanic and Romance rather than the assignment of topic and focus features *per se* (Zubizarreta 1998, Szendroi 2001, Samek-Lodovici 2006). The inconclusive status of examples such as those in (3) in Germanic and Romance languages reinforces the general view that crucial notions of Information Structure, such as, Topic and Focus are not part of syntax but must be added to the linguistic expression once computed by C_{HL}. The wisdom is to claim that introducing these notions in syntax violates Chomsky's (1995: 228) inclusiveness condition:

(5) Given the numeration N, C_{HL} computes until it forms a derivation that converges at PF and LF […] A "perfect language" should meet the condition of inclusiveness: any structure formed by the computation […] is constituted of elements already present in the lexical items selected for N; no new objects are added in the course of computation apart from rearrangements of lexical properties.

2.2.1. On question-answer pairs

As a way of evaluating the inclusiveness condition, let us consider question-answer pairs, which typically involve focus constructions. Adopting (5) would suggest that question (7) derives from the numeration in (6).

(6) $\{$C, T, who, did, love, John$\}$

(7) [CP who [C did [TP John [T t_{did} [$_{VP}$ t_{John} love t_{who}]]]]]

The idea here is that interrogative force, the so-called wh-feature, is a property of C, which in English, forces wh-movement of the interrogative wh-phrase to [spec CP] and I-to-C movement, as a requirement of clause-typing (Cheng 1991). It is important to observe at this stage already that the lexicon of English contains the functional category C endowed with the feature interrogative (i.e., C_{wh}). Given this state of affairs, it is admissible (and even widely accepted in the literature) that the lexicon contains certain discourse-related (or information-structure-sensitive) functional items that project in syntax. This is the first crack in the common argumentation rejecting the idea of any direct interaction between information structure and syntax.[4] As the above description shows, the correct characterization is that certain information sensitive features (e.g., interrogative force) are directly encoded in syntax. Given this state of affairs, the next question to ask is what is so unique to interrogative feature that necessitates a different treatment as opposed to topic and focus features?

[4] Another feature that is commonly assumed to project in syntax though it primarily relates to discourse (and therefore information structure) is definiteness/specificity as encoded by (in)definite articles (Abney 1987, Szabolcsi 1987, 1994, Longobardi 1994). Given the close relation between definiteness/specificity and topicality (see following discussion and section 4.1.) it is not clear to me why one is seen as a legible formal feature expressed by the nominal periphery D, while the other is banned from the clausal periphery.

Suppose that the question in (7) 'naturally' requires the answer in (8b), with new information focus on *Mary*, though the numeration (8a) contains no functional focus category that may combine with *Mary*.

(8) a. {C, T, –ed, love, John, Mary} b. John loved Mary[New Information Focus]

Some questions that immediately come to one's mind are: (i) why is (7) a product of syntax to the exclusion of the focus structure (8b) that it requires? (ii) Why must a natural answer to a (wh)question always contain a focused expression? (iii) What triggers the matching between question and focus as observed in (6-7)? (iv) Where does the (extra) focus feature in (8b) come from if it is not present in the numeration?

This paper tries to answer these questions by considering questions and answers in parallel. The interaction between (7) and (8) suggests that even though speakers have the freedom of choice with regard to which linguistic expression to use in a particular context or discourse, the form of this expression is a product of syntax that directly relates to the Numeration (see also Zwart 1998). Assuming that discourse participants are collaborative, we can make the reasonable hypothesis that the trigger of the focus feature in a felicitous reply represents an attempt of the speaker to provide a matching expression to the asked question. This recalls Zubizaretta's (1998: 4) ordered Assertion Structure (AS) where,

> "in the case of question-answer pairs, the presupposition provided by the context question is part of the AS of the answer statement […] The focus-presupposition structure of such statements is represented in terms of two ordered assertions. The first assertion (A1) is the existential presupposition provided by the context question. The second assertion (A2) is an equative relation between a definite variable (the restriction of which is the presupposition provided by the context question) and a value."

In looking at question-answer pairs this way, we realize that the structure of the answer (i.e., with regard to clause-typing and focus) is not completely orthogonal to that of the context question. That the matching answer contains a focused expression seems to be a requirement of the question operator in the question.[5] If indeed clause-typing and wh-movement are context-invariant because they satisfy formal syntactic requirements in question formation, we can formulate the hypothesis that focus assignment (which in some languages, e.g., Gungbe, require constituent displacement) also satisfies a syntactic requirement.

Treating questions and answers this way, appears compatible with the format of question-answer pairs in Mandarin Chinese, where we observe a more transparent correlation between the form of a question and that of the answer. In their discussion of word order in Mandarin Chinese, Li & Thompson (1975) gave the contrast in (9) to show that the sentence in (9b), though a perfect syntactic choice, "is extremely unnatural as an answer to (9a) [MC, Li & Thompson 1975: 173]."

(9) a. Yàoshi **ne** ?
key Inter
'What about the key(s)?'

b. ??Wǒ wàng le yàoshi le
I forget Asp key Asp
'I forgot the key(s)'

According to Li & Thompson, (9b) is strange because in the particular context of question (9a), the noun phrase *yàoshi* 'the key(s)' is forced to be interpreted as discourse anaphoric, and therefore as a topic. The authors further give a number of alternatives that satisfy this requirement and represent felicitous answers for (9a). These are given under (10).

[5] This observation actually goes back to Chomsky (1971) who argues that in question-answer pairs, the focused expression in the answer provides a value to the variable in the question.

(10) a. Yàoshi, wǒ wàng le [Topicalization]
Key(s) I forget Asp
'As for the key(s) I forgot it/them'

b. Wǒ bǎ yàoshi wàng le [The *bǎ*-marked DP is discourse anaphoric]
I ba key forget Asp
'I forgot the keys'

c. Wǒ wàng le [Topic drop]
I forget Asp
'I forgot them'

The different strategies in (10) indicate that the felicitous answer to (9a) requires the noun phrase *yàoshi* 'the key(s)' to be syntactically marked as a topic, hence its sentence-initial position in (10a). In (10b) we have a *bǎ*-construction about which Li & Thompson (1981) report that a noun so marked denotes "something about which the speaker believes the hearer knows". As I suggest later, this recalls Prince's (1981) notion of familiarity or aboutness which relates to the notion of topicality. In this regard, it is interesting to note that (10b) is an SOV construction, a fact that would suggest that 'object shift' here interacts with definiteness and topicality. Finally, example (10c) is an instance of topic drop, which means that the DP *yàoshi* is discourse-anaphoric.

The correct description therefore seems to be that question (9a) does not open the floor for a flexible situation where any choice between the grammatical sentences under (9b) and (10) could fit in. Instead, (9a) calls for the choice of a linguistic expression in (10) excluding (9b). It is clear from this description that the choice of a linguistic expression is sensitive to discourse considerations external to syntax. But what matters for the present discussion is that whatever linguistic expression is chosen by the speaker has very specific

syntactic properties. In the case of (9a), the felicitous expressions have a syntactic topic that formally distinguishes then from the infelicitous answer in (9b), which itself has a specific syntactic format (i.e., simple SVO declarative). There seems to be no principled way of approaching these facts, if we interpret (as often implied in the literature) the inclusiveness condition as an exclusion of topic and focus features from core syntax. Indeed, it cannot be said that the sentences under (10) acquired the topic reading only in context. If this were the case, they would be ambiguous and therefore undistinguishable from (9b).

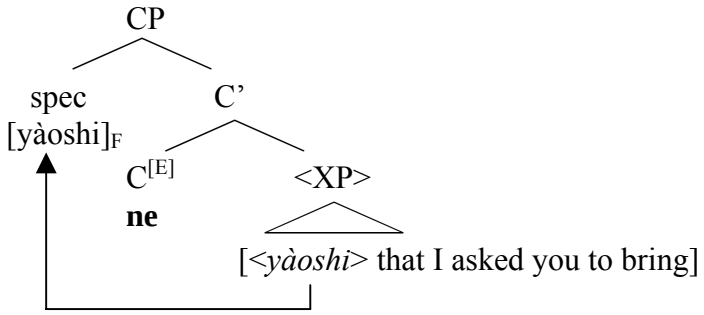
Based on (2), I argue in what follows that strict application of the inclusiveness condition requires that core notions of information structure (interrogative force, topic, focus) be part of the numeration, and project in syntax.

2.2.2. wh-words and interrogative force

Let us first step back and consider the wh-question in (7). In discussing this example, we suggested that the numeration contained an interrogative C_{wh} that requires wh-movement in English. In this view, the wh-phrase moves to [spec CP] to clause-type the question (Cheng 1991, Cheng & Rooryck 2000). This analysis cannot extend to example (9) though: there is no wh-phrase in this example, all we have is a noun phrase followed by a question marker: *ne*. In terms of Li and Thompson (1981: 305–306), this declarative typing particle can also be used to type various questions, including truncated questions that include a noun phrase only (9a). The authors report that though *ne* retains its declarative meaning (e.g., signifying: 'this is what I say in connection with your previous claim') its meaning in questions can be interpreted as: 'In connection with your claim or expectation, let me find out…' or 'with respect to what you have just said, let me ask you…' This description suggests that (9a) should not be regarded as an isolated DP including the particle, but rather as a fragment question. Adopting Merchant's (2004) view of fragment answers, this would mean that (9a)

involves ellipsis of the complement of *ne*. More precisely, (9a) presumably represents a structure where *ne*, the ellipsis site, heads a functional projection within C, whose specifier contains the attracted DP *yàoshi*, and whose complement (i.e., the presupposition described by Li & Thompson) is silenced as described in (9'a), see section 3.2.2. on Gungbe. This question therefore recalls English what-about questions where the relative clause is elided (e.g., as in *what about the keys <that I asked you to bring>*?). The only difference between English and Mandarin Chinese therefore seems to be that in the latter, the DP constituent that moves to [spec FP] does not contain a wh-word.

(9')



What appears from this description is that the particle *ne* has typical properties of functional elements responsible for ellipsis (as described in Merchant 2004): it instructs non-pronunciation of it complement, which must be given or presupposed. In addition, *ne* appears to interact with interrogative force in Mandarin Chinese. As such it is endowed with an EPP-feature that forces attraction of the DP-topic in its specifier. These properties of *ne*-constructions, clearly distinguish them from Mandarin Chinese wh-questions, which involve genuine wh-phrases (e.g., *shéi*: who; *shénme*: what; *duō*: how).

Given that the fragment question (9a), as described in (9'), is properly interpreted as wh-question though it lacks a wh-word, we can reasonably assume that wh-words do not express the feature interrogative. In the context of the English example (7b), this leads us to hypothesize that: (i) the expression of interrogative force does not depend on the wh-phrase *who*, but on some silent question particle within C (e.g., as in Mandarin Chinese), and (ii)

wh-movement satisfies other requirements (e.g., focus, EPP) than clause typing. Contrary to what is often discussed in the literature therefore (e.g., Cheng 1991, and much related work) this description suggests that we must tease wh-movement and clause-typing apart. The distribution of wh-phrases in English provides us with a first set of empirical facts that support this view. Indeed, English wh-phrases are not restricted to interrogatives only but can freely occur in declarative clauses that also involve wh-movement. These examples indicate that English wh-movement is not contingent on interrogative clause typing.

(11) a. I showed Bill the man who wrote the poem

b. I know the person to whom Bill talked

c. I saw the student who Bill invited to the party

As extensively discussed in Aboh & Pfau (2006), English is not unique in this respect and data from typologically different languages confirm the view that wh-phrases do not generally clause-type. For instance, certain well behaved wh-movement languages display wh-questions without wh-phrases, indicating that the overt occurrence of a wh-phrase is not necessary for clause typing. A case in point is the sign language of The Netherlands (*Nederlandse Gebarentaal*, NGT). NGT is arguably SOV and has a full paradigm of wh-signs. As we can see from the examples in (12a-b), these wh-phrases often occur to the right edge, therefore indicating that NGT is a wh-movement language (see Aboh & Pfau 2006 and references cited there).

		wh		wh
(12)	a.	BOOK STEAL WHO	b.	INDEX$_2$ DRINK WHAT
		'Who steals the book?'		'What do you drink?'

Quite often, these right edge wh-phrases co-occur with a sentence-final question marker glossed as PU in (11c-d). This particle also occurs in yes-no questions (11e).

(12) c. wh
INDEX$_3$ SAY WHAT PU
'What did s/he say?'

d. wh
INDEX$_2$ BIKE STEAL WHO PU
'Who stole your bike?'

e. yn
INDEX$_1$ OFTEN USE PU
'Do I use it often?'

Interestingly, the wh-signs can be dropped in appropriate contexts, but not the question marker, which is always present in the form of PU as in (13) (or else in the form of some prosodic cue).

(13) a. wh
YESTERDAY INDEX$_2$ BUY PU
'What did you buy yesterday?'

b. wh
TRAIN FRANKFURT LEAVE PU
'When/where does the train to Frankfurt leave?'

I conclude from these facts that questions always involve a question particle, regardless of whether the language involves wh-phrases and/or wh-movement. This question particle is silent in English, but not in Mandarin Chinese, and NGT.[6] In addition, recent studies on the syntax of wh-phrases indicate that when they move, they primarily do so for reasons of focus rather than interrogative clause typing (Bošković 2000, 2001, Aboh & Pfau 2006). Put together, these facts lead me to suggest that in questions, clause-typing and wh-movement are related to two different probes: interrogation (Force) and focus (Foc).

[6] In languages like English and French the presence of non-overt particles in the structure is betrayed by other cues (e.g., prosody, see Cheng & Rooryck 2000, Aboh & Pfau 2006).

If we grant this hypothesis in the context of the question in (7), we may, in a next step, ask the question of whether these probes are not present in the matching answer in (8). Put another way, is it possible that the covert question particle under Force, and the licensing of focus operators in the question, are somehow connected to clause-typing and focus assignment in the answer (in a way similar to Mandarin Chinese)?

3. The solution

With this question in mind, I argue that:

(14) Core syntactic properties that trigger question formation also correlate with the information structure required in the answer (e.g., New information focus vs. identificational or contrastive focus, see Kiss (1998) and much related work).

3.1. Information structure and the numeration

Under (14), the numeration in (6), now repeated as (15a), should be refined as in (15b). In the latter case, the numeration contains both a silent question marker and a focus marker, which I assume are included in the lexicon.

(15) a. {C, T, who, did, love, John} b. { Inter, Foc, T, who, did, love, John }

On the basis of (15b), we reach the partial representation in (16) in which the focus head attracts the wh-phrase to its specifier, while Inter hosts a null morpheme that clause-types the sentence.

(16) ....[$_{InterP}$ ∅….[$_{FocP}$ who [$_{Foc}$ did [$_{TP}$ John [$_{T}$ t_{did} [$_{VP}$ t_{John} love t_{who}]]]]]]]

The immediate conclusion here is that the interaction between Inter and Foc is a result of the numeration, not pragmatics (see Chomsky 1971). Accordingly, the required answer has a

type (e.g., declarative) and contains a focused expression that matches the question (Zubizaretta 1998). Given this discussion, we reach the description that the functional heads (Force/Inter, Foc) are present in the numeration of both the question and the answer. These findings therefore show that just as questions have a type (expressing speech act modality) and a focus (the request of information), matching answers have a type (also indicating speech act modality) and a focus (providing the requested information).

A language that readily supports this viewpoint is Maale, an SOV North Omotic language spoken in Southern Ethiopia. In this language, all sentence types must be morphologically marked on the verb, which interestingly enough happens to be final in this language (Amha 2001, chapters 7, 8).[7] In the question answer pairs in (17) therefore, interrogative force is necessarily marked on the verb by the suffix *–y* in the question (17a), while simple declarative force is realized by the suffix *–ne* in the answer (17b) (Amha 2001: 286).[8] The examples in (17c-e) involve a permissive question inflected by the permissive question element *–ondó*, which conditions the occurrence of the imperative modal elements *–é* or *–étera* in the answers.[9]

(17) a. kan-ó naʔʔ-ó-na ʔas-ommá-na wóí-t-é-**y** ?
do-Abs child-Def-Abs-Inst person-DIM-Instr be-PF-Inter
'What did the dog and the little boy do?'

b. kóí-tsi ʔízo ʔark'-é**ne**
search-Inf 3sg.F:Abs hold-PF-A:DCL
'They started looking for it'

[7] That the verb is final could be understood as a consequence of V-to-C movement followed by pied-piping of IP into [spec CP], as proposed in Aboh & Pfau (2006) for Indian sign language.
[8] Abs = Absolutive; DCL = declarative; Def = definite; DIM = diminutive; IMP = Imperative; Inf = Infinitive; Instr = Instrument; PERM = Permissive; PF = perfective aspect; POL = Polite form; 3sg.F = 3sg feminine;
[9] I thank Azeb Amha for helping with these Maale data.

c. Zíró mukk-**ondó**?
Tomorrow come-PERM:Q
'May I come tomorrow?'

d. Híyyo mukk-**é**
yes come-2sg:IMP
'Yes, come!'

e. Híyyo mukk-**étera**
yes come-2sg:POL:IMP
'Yes, please come!'

That discourse modality manifests itself in the form of inflection in Maale is a remarkable fact pointing to the relevance of core syntax in manipulating these notions. Another noteworthy point for this discussion is that the type of modality chosen for the question strongly constrains that of the answer. These examples therefore indicate that both the question and the answer have an active Force (interrogative vs. declarative, permissive interrogative vs. imperative) that must have been part of their respective numerations.

Lele, a Chadic language, provides us with additional supporting evidence. Lele is SVO, but clitics (i.e., 1sg, 2sg) precede the verb, while 3sg and plural forms immediately follow. In this language, both yes-no questions (18a) and wh-questions (18b-c) require the presence of a sentence-final question particle. Further observe that Lele displays both in situ and ex situ wh-questions. Interestingly, the ex situ construction (18c) occurs in the context of the focus marker *ba*. Here, the wh-phrase moves to the left periphery because it is focused, and the required answer presumably involves a contrastive focus (Frajzyngier 2001).

(18) a. Kiya hàb kùlbá ke-y **gà**? [p. 278]
Kiya find cow Gen-3sg[M] Inter
'Did Kiya find his cow?

b. Mè ày *wéy* **gà**?
2sg[F] marry who Inter
'Who did you marry?' [p. 284]

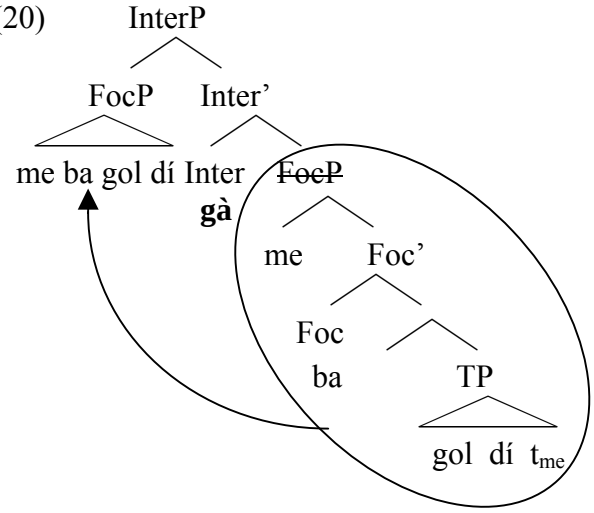
c. *Me* **ba** gol dí **gà**?
What Foc see 3sg[M] Inter
'What did he see?' [p. 286]

If C_{HL} is mainly concerned with mapping a numeration N to a linguistic expression (π, λ), it must be the case that N→λ computation in these cases involves both the question and focus markers. This in turn suggests that the ex situ versus in situ contrast in these Lele examples translates into the numerations in (19a-b). Numeration (19a), which generates question (18c), contains the question marker and the contrastive focus marker, while numeration (19b), which represents (18b), contains the question marker and a new information focus head, which I assume to be null in this context.

(19) a. { Inter[gà], Foc[ba], T, Wh[me], dí, gol } b. {Inter[gà], Foc[∅],T, Wh[wéy], mè, ày }

Under this description, the Lele example in (18c) can be represented as in (20). Here we see that the wh-phrase moves to [spec FocP], while the whole FocP pied-pipes to [spec InterP] presumably to check an interrogative (or EPP) feature under Inter (see Aboh & Pfau 2006, Aboh in press).

(20)



As suggested previously, representation (20) is compatible with the view that the syntax of wh-questions involves two probes: Foc and Inter, each of which may trigger displacement operations leading to various chain formations. On a more general note, the observed facts also imply that speakers of languages like Maale, Lele (and Gungbe which I discuss in section 3.2.) must acquire the described information-structure-sensitive particles as part of the lexicon. Though never discussed in the literature, this point would support the view defended here that the information structure has direct access to syntax via the numeration. Based on this, we can further claim (21) for which I now provide both empirical and conceptual motivations.

(21) The only interface between Information Structure and Syntax is the Lexicon.

3.2. A Lesson from Gungbe

This strong claim, which perfectly squares with minimalist considerations, finds a robust empirical support from Gungbe, an SVO language of the Gbe group.[10] Data from these languages indicate that they literally spell out the left peripheral backbone.

3.2.1. Topic, focus, and interrogative markers in Gungbe

As a way of illustration, consider the following Gungbe examples involving a topic phrase and its marker in (22a), a focused phrase and its marker in (22b) and a wh-phrase and the focus marker in (22c). In all these examples, the topic or focus marked element occurs to the right of the complementizer *dɔ́*, the equivalent of English 'that'.

[10] These are Niger-Congo languages of the Kwa family. Though there are differences, the properties described here are generally found in the Gbe group as well as the Kwa family.

(22) a. Ùn sè ḍɔ̀ dàn lɔ́ **yà** Kòfí hù ì
1sg hear that snake Det Top Kofi kill 3sg
'I heard that, as for the snake, Kofi killed it'

b. Ùn sè ḍɔ̀ dàn lɔ́ **wɛ̀** Kòfí hù
1sg hear that snake Det Foc Kofi kill
'I heard that Kofi killed THE SNAKE'

c. Ùn kànbíɔ́ ḍɔ̀ étɛ̀ **wɛ̀** Kòfí hù ?
1sg ask that what Foc Kofi kill
'I asked what did Kofi kill?'

As extensively discussed in Aboh (2004a, b, 2006a), these languages support Rizzi's (1997) split-C hypothesis in a surprisingly transparent way. Indeed, the examples in (22) show distinct C-type markers that occupy different positions within a space delimited to the left by the complementizer and to the right by the subject. In addition, example (23) indicates that these markers are not in complementary distribution and can freely cooccur.

(23) Ùn sè **ḍɔ̀** xwé lɔ́ **yà** Kòfí **wɛ̀** Àsíbá gbá-ɛ̀ ná
1sg hear that house Det Top Kofi Foc Asiba build-3sg for
'I heard that, as for the house, Asiba built it for KOFI'

Before getting onto the discussion, it is crucial to stress that the described markers do nothing else in Gungbe but mark discourse properties (e.g., topic, focus). The focus and topic markers are not found in any other context. They never occur as copulas or main predicates, and they don't seem to have any lexical meaning apart from indicating that what is to their left is a

focus or a topic constituent. Therefore these markers encode topic and focus non-ambiguously.

In accounting for these markers, I have proposed in previous work that they head different functional projections, as represented in (24a). Under this view, the markers probe over the relevant elements/constituents within the clause that are attracted to [spec TopP] and [spec FocP] as represented in (24b) (Aboh 2004a, chapter 7, 8, 2006a).

(24) a …[$_{\text{ForceP}}$ [$_{\text{Force}}$ **ɖɔ̀** [$_{\text{TopP}}$ [$_{\text{Top}}$ **yà** [$_{\text{FocP}}$ [$_{\text{Foc}}$ **wɛ̀** [$_{\text{FinP}}$ ….[VP]]]]]]]]]

b. …[$_{\text{ForceP}}$ [$_{\text{Force}}$ **ɖɔ̀** [$_{\text{TopP}}$ xwé lɔ́ [$_{\text{Top}}$ **yà** [$_{\text{FocP}}$ Kòfí [$_{\text{Foc}}$ **wɛ̀** [$_{\text{FinP}}$ Àsíbá [gbá-ɛ$_{\text{[lxwé lɔ́]}}$ ná $t_{\text{Kòfí}}$]]]]]]]]]

In addition, to these markers which typically occur to the left edge, Gungbe has certain sentence-final discourse particles, such as the yes-no question marker, which in this example is an additional low tone on *mótò* ‘car’. Contrast the minimal pair (25a) and (25b) (see Aboh 2004a, chapter 8, 2004b).

(25) a. Kòfí xɔ̀ mótò
Kofi buy car
‘Kofi bought a car’

b. Kòfí xɔ̀ mótɔ̀?
Kofi buy car-Inter
‘Did Kofi buy a car?’

A priori, one could think of the Gungbe low tone as a simple prosodic requirement on yes-no questions similar to rising or falling intonation in some languages. This would follow because yes-no questions require a falling intonation in Gungbe and other Gbe languages, contrary to the rising intonation often found in Indo-European languages. A look at the closely related language Fongbe, however, indicates that yes-no questions include a full question morpheme:

à (26). This suggests to me that the Gungbe low tone is presumably a leftover of a full morpheme that the language has had at earlier stages.[11]

(26) É ɖù nú **à**? [Fongbe]
3sg eat thing Inter
'Did he eat?

Granting this, it appears that, in Gungbe as well as in Fongbe, topic and focus markers, which mainly occur in the left periphery and the right peripheral marker (i.e., the yes-no question marker) may co-occur. When this happens, we may obtain the sequences in (27) for Gungbe, and (28) for Fongbe (see Aboh 2004a, chapter 8 for discussion).

(27) a. Ùn kànbíɔ́ ɖɔ̀ Kòfí ɖù lɛ́sì **wɛ̀**? [Gungbe]
1sg ask that Kofi eat rice Foc-Inter
'I asked whether KOFI ATE RICE [e.g. he shouldn't do so because he is taking medicine]?'

b. Ùn kànbíɔ́ ɖɔ̀ Kòfí ɖù lɛ́sì **yà**?
1sg ask that Kofi eat rice Top-Inter
'I ask whether Kofi ate rice [as planned/mentioned]?'

c. Ùn kànbíɔ́ ɖɔ̀ Kòfí ní xɔ́ mótò **wɛ̀** **yà**?
1sg ask that Kofi MOOD buy car Foc Top-Inter
'I asked whether KOFI SHOULD BUY A CAR [as planned/mentioned]'?

d. *Ùn kànbíɔ́ ɖɔ̀ Kòfí ní xɔ́ mótò **yà** **wɛ̀**?
1sg ask that Kofi MOOD buy car Top Foc-Inter

[11] See Aboh 2004a and references cited there for discussion on the Gbe languages.

(28) É dù nú **ǎ** **wɛ̀** **à**? [Fongbe]

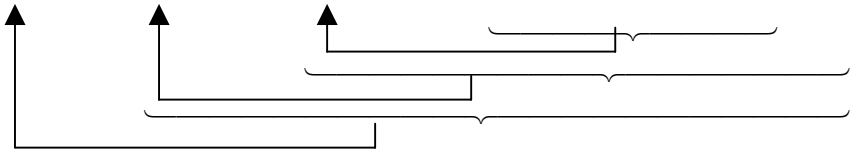
3sg eat thing Neg Foc Inter

'Is it that s/he has not eaten? [Lefebvre & Brousseau 2002: 135, 485]

These examples are particularly illustrative of various facts: (i) some of the left peripheral markers described in (22-24) (e.g., topic and focus) can also occur to the right. (ii) When they co-occur to the right, they appear in the mirror image of that in (24). In (27c), for instance, we observe the sequence (*wɛ̀>yà*) as opposed to (*yà>wɛ̀*) in (24). Notice from example (27d) that preserving the order in (24) in sentence-final position leads to ungrammaticality.

I have proposed elsewhere that this mirror image effect is the result of successive pied-piping (i.e., snowballing movement) which rolls the clause up to the specifier of the interrogative phrase, as described in (29), (Aboh 2004a, b, d).

(29) [$_{ForceP}$ [$_{Force°}$ɖɔ̀ [$_{InterP}$ [$_{Inter°}$ ∅ [$_{TopP}$ [$_{Top°}$ **yà** [$_{FocP}$ [$_{Foc°}$ **wɛ̀** [$_{FinP}$ Kòfĩ ní xɔ̀ mótò]]]]]]]]]



On the basis of these examples, we can say that the peripheral markers (i.e., Topic, Focus, Interrogative) form a paradigm where they mark discourse properties. Since Chomsky (1977), and more recently Cheng (1991), it is unanimously accepted that question particles head a functional projection (ForceP/InterP) in syntax. If we look at the Gungbe facts from this perspective, it is reasonable to conclude that the topic and focus markers are similar to the question particle in that they project in syntax.[12] Put another way, nothing in Gungbe suggests that interrogative force is a particular information-structure-sensitive feature that

[12] See Aboh (2004a) where it is shown that both C and I involve markers (i.e., free morphemes) that are the expressions of distinct functional projections.

requires a different treatment from topic and focus. If this is so for Gungbe (and generally Gbe), then the burden of proof is on the adherents of the view that Romance and Germanic are different and that interrogative force in these languages requires a special treatment.

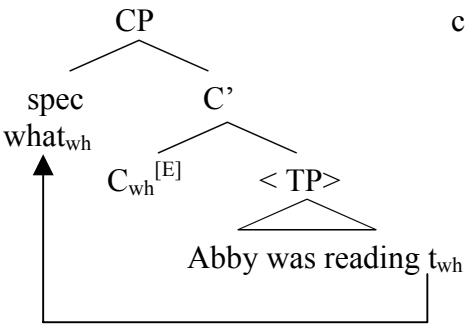
In continuing our discussion, I now turn to fragment answers in Gungbe, which I show support the view advocated in this paper.

3.2.2. Ellipsis

A particularly insightful conclusion reached in Merchant (2004: 664) is that subsentential answers can function as propositions (i.e., correspond to assertion) because (i) they are syntactically generated as parts of full sentences, (ii) they are subject to a movement operation that evacuates them from the non pronounced part of the sentence. Under this view therefore, sentence (30a) can be assigned the representation in (30b) (Merchant 2004: 670).

(30) a. Abby was reading something, but I don't know what <Abby was reading *t*>

b.



c. syntax of E: E [*u*wh, *u*Q]

In the representation (30b) the feature E, which marks the ellipsis site, has the syntax in (30c), forcing it to attach to a C head bearing the relevant wh-features, which are checked by the fronted wh-phrase. Merchant (2004) further argues that while the feature E may attach to different heads cross-linguistically, its phonological and semantic properties are uniform across languages: it instructs non-pronunciation of its complement and the unpronounced complement must be given. What emerges from Merchant's analysis is that ellipsis derives

from a formal syntactic feature, which in English and some languages, attaches to a head within C, whose complement remains unpronounced.

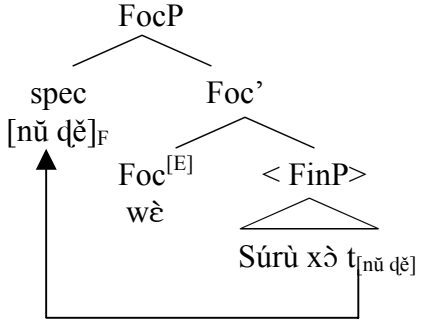
Given the inclination of Gungbe to spell out all the heads within the left periphery (24), we may hope to find evidence in this language where the covert stranded $C_{wh}^{[E]}$ in English will get pronounced. Finding such cases will not only provide strong empirical support to Merchant's analysis of ellipsis, but will also support the point made here that the functional elements represented in (24) are real syntactic objects endowed with formal syntactic features (see also Baltin 2006). Once again, Gungbe makes life easy for us. Though ellipsis is very restricted in this language, the following examples are possible:

(31) a. Súrù xɔ̀ nǔ ɖé, àmɔ́n má nyɔ́ nǔ ɖě *(wɛ̀)
Suru buy thing Indef but 1sg.Neg know thing that Foc
'Suru bought something, but I don't what'

b. Mɛ̀ ɖé wá, àmɔ́n má nyɔ́ mɛ̀ ɖě *(wɛ̀)
person Indef come but 1sg.Neg know person that Foc
'Someone came, but I don't know who'

Compared to their English translations, these Gungbe examples show that: (i) the locus of the feature E in Gungbe is the focus head (i.e., Foc realized here by the focus marker), (ii) ellipsis is not possible in the absence of the focus marker. Adapting Merchant's analysis to Gungbe, I partially represent (31a) as in (32a), where the feature E has the syntax in (32b).

(32) a.



b. syntax of E: E [*u*Foc, *u*Q]

This analysis sheds light on the fact that the element that fronts in Gungbe ellipsis is not a simple wh-phrase of the English type, but a complex phrase involving a dummy noun phrase *nú* 'thing' and a Q-morpheme (see section 5). This clearly relates to the point made in this paper that the feature at work in these Gungbe constructions is a focus feature that is anchored under Foc, the focus head.

Under this view, the only difference between English and Gungbe is that the latter, but not the former, spells out the focus head whose specifier is occupied by the fronted wh-phrase. Given our minimalist standpoint that the role of the computational system C_{HL} is to map the numeration onto a converging linguistic expression, I conclude from the discussed facts that Gungbe involves discourse particles endowed with formal features that are introduced lexically in the derivation as part of the numeration. Accordingly, Gungbe is a discourse-configurational language because the syntax of discourse particles determines the surface structure of the sentence. We therefore have a clear case where information structure, encoded by discourse particles, has direct access to syntax. The discussed examples do not only support the viewpoint defended here, but they also indicate that any analysis departing from this view will violate the inclusiveness condition in (5). [13]

If information structure is indeed a property of the lexicon, we expect to find its effects in other domains besides the clausal ones. Using data form the Gungbe noun phrase, the following section shows that the DP is another field where we find information-structure-sensitive functional items.

[13] Indeed, assuming that the markers are introduced at PF (as one could imagine) is equivalent to saying that lexical items are added to the derivation once completed by C_{HL}. A seemingly possible way out could be to propose that these discourse markers are prosodic cues comparable to stress-assignment in Indo-European. Though a possibility, this view makes these languages even more exotic than they already are because they will be the only cases known to me where prosody translates into words! Therefore the burden of proof is on those who embrace this view.

4. Some parallels between the sentence and the noun phrase

Previous discussion shows that Gungbe (and Gbe languages in general) include discourse particles (e.g., Topic, Focus, Inter) that are introduced in the derivation as part of the numeration. In this regard, a remarkable property of these languages is that we find similar particles within the DP (Aboh 2004c).

4.1. The activation of Topic within D

Gbe languages distinguish between specific and non-specific referents. Specific referents are defined as follows:[14]

(i) A specific definite noun phrase is strongly D(iscourse)-anaphoric and represents a unique referent assumed to be known to both speaker and hearer, and which the speaker intends to refer to.

(ii) A specific indefinite noun phrase need not be D-linked. It represents an existing referent that the hearer may not know about, but which the speaker has in mind and intends to refer to (Aboh 2006b: 224).

In Gungbe, a noun phrase without a determiner (referred to here as a *bare* DP) may be interpreted as (in)definite or generic depending on the context.[15] The sequence *távò títán* in (33a) is understood as definite (because it is unique in the discourse setting) but non-specific because it was not previously mentioned in discourse. On the other hand, a noun phrase including a determiner (referred to as a *full* DP) must be interpreted as specific. When these DPs have been explicitly mentioned in previous discourse or represent entities, referents or events which both the speaker and the hearer know about or have some shared experience of, they are interpreted as definite specific. Accordingly, the sequence *távò lɔ́* (33a) is interpreted

[14] See also Ionin (2006) for discussion on specificity.
[15] See Aboh (2004, chapters 3, 4) for arguing that such bare noun phrases are DPs.

as definite specific (because it refers back to *távò títán* 'table first'). On the other hand, indefinite specifics (e.g., *távò ɖé* in 33b) refer to a referent that is known to the speaker but not necessarily to the hearer. Observe that this specific reading is absent in the sequence *távò cè* (33b) which is only definite. In addition, example (33c) shows that *lɔ́* and *ɖé* are mutually exclusive, though they can co-occur with other DP-internal markers (e.g., demonstrative, possessive) (33d). This is taken as a piece of evidence that *lɔ́* and *ɖé* target the same position, and should be distinguished from other nominal modifiers such as demonstratives (see Aboh 2004a, chapter 4, 2004c).

(33) a. Kɔ̀kú mɔ̀n **távò** **títán** bò ɖɔ̀ émì ná xɔ̀ **távò** **lɔ́**

Koku see table first and say 3sg Fut buy table $Det_{[+spec, +def]}$

'Koku saw the first table and said that he would buy that specific table'

b. Kɔ̀kú mɔ̀n **távò** **cè** bò ɖɔ̀ émì ná xɔ̀ **távò** **ɖé**

Koku see table 1sg-Poss and say 3sg Fut buy table $Det_{[+spec\ -def]}$

'Koku saw my table and said that he would buy a certain table'

c. *Kɔ̀kú mɔ̀n **távò** **lɔ́** **ɖé**

Koku see table $Det_{[+spec, +def]}$ $Det_{[+spec, -def]}$

d. Kɔ̀kú mɔ̀n **távò** **cè** **títán** **éhè** **lɔ́** bò ɖɔ̀ ….

Koku see table 1sg-Poss first Dem Det Coord say

'Koku saw this first table of mine and said….'

It appears from this brief description of the DP that a Gungbe non-specific noun need not be marked while a specific noun phrase must necessarily take the marker *lɔ́* or *ɖé* (see Aboh 2004a, chapters 3, 4 for discussion). The Gungbe definite specific DPs recall Li & Thompson's (1975) description of *ba*-constructions illustrated in (10b), or Prince's (1981)

assumed familiarity. A common aspect to these DPs is that they interact with discourse topic. Thinking of the parallels between CP and DP (Aboh 2004c), a hypothesis that we may formulate is that specificity marking within DP matches topicality at the clausal (i.e., CP) level. That this is the right characterization is suggested by an interesting contrast between focalization and topicalization in Gungbe.

Both bare DPs and full DPs can be freely focused and fronted to the left of the focus marker. In addition, the process is not sensitive to definiteness, as indicated by the various options in (34).

(34) a. Távò wɛ̀ Kófí sà bò mɔ̀n kwɛ́
table Foc Kofi sell coord find money
'Kofi sold TABLE(S) and found money [i.e., got rich]'

b. Távò cè wɛ̀ Kófí sà …..
table 1sg-Poss Foc Kofi sell
'Kofi sold MY TABLE …..'

c. Távò lɔ́ wɛ̀ Kófí sà ……
table Det Foc Kofi sell
'Kofi sold THAT (TYPE OF) TABLE…..'

d. Távò ɖé wɛ̀ Kófí sà …..
table Det Foc Kofi sell
'Kofi sold SOME/A (TYPE OF) TABLE…..'

With regard to topicalization, however, only definites and/or specific constituents can

undergo topic movement to the left of the marker *yà*.[16] Consider the ungrammatical examples (35a-b) and contrast them with the grammatical ones under (35c-d):

(35) a. *Távò yà Kófí sà è bò mɔ̀n kwɛ́
table Top Kofi sell 3sg Coord find money
'As for table(s) Kofi sold (it/them) and found money [i.e., got rich]'

b. *Távò ɖé yà Kófí sà è…..
table Det Top Kofi sell 3sg
'As for some/a certain table(s) Kofi sold it…..'

c. Távò cè yà Kófí sà è…..
table 1sg-Poss Foc Kofi sell 3sg
'As for my table, Kofi sold it …..'

d. Távò lɔ́ yà Kófí sà è..…
table Det Foc Kofi sell
'As for that specific table, Kofi sold it….'

Similarly, adjuncts that have a unique reference (e.g., time adjuncts) can freely topicalize, but other adjuncts (e.g., place adjuncts) can only front if they are made definite or specific. Contrast again example (36a) with the time adjunct to examples (36b-c) where only the specific-definite adjunct *xɔ̀ lɔ́ kpá* 'beside that specific room' is allowed as topic.

(36) a. Sɔ̀ yà, Kòfí ná zà xɔ̀ lɔ́ kpá] gbáú !
tomorrow Top Kofi Fut sweep room Det Post$_{[beside]}$ necessarily
'Tomorrow, Kofi has to sweep beside the room!'

[16] Topics, unlike focus constructions require a resumptive pronoun in Gungbe except in the case of referential time adjuncts (e.g., today, tomorrow). Contrast the examples (34c-d) to (35a) below. See Aboh (2004a, chapters 7, 8) for discussion.

b. [Xɔ **lɔ́** kpá] yà, Kòfí ná zà flɛ́n gbáú !
room Det Post[beside] Top Kofi Fut sweep there indeed
'As for beside the aforementioned room, Kofi will sweep there!'

c. *[Xɔ kpá] yà, Kòfí ná zà flɛ́n gbáú !
room Post[beside] Top Kofi Fut sweep there indeed
'As for beside room(s), Kofi will sweep there!'

Given that both bare DPs and full DPs can be focused (34), we cannot attribute the contrast between (34) versus (35) and (36) to some structural deficiency making bare DPs unsuitable for movement operations. Instead, it seems as if definiteness (however encoded in this language) or specificity marking with *lɔ́* is a pre-requisite for clausal topicalization. This is not surprising since topics are discourse-anaphoric (or D-linked) by definition. The interesting question instead is this: Why would we have this matching between the topic head and the DP it attracts if the interaction at stake were just a pragmatic one?

Keeping to minimalist assumptions, the observed matching appears a classical instance of the probe-goal interaction whereby the Gungbe topic head probes over a definite or specific (i.e., discourse-anaphoric) constituent, which is attracted to its specifier position and interpreted as topic. Needless to say that C_{HL} properly computes this operation because the topic marker *yà* and the specificity marker *lɔ́* are introduced in the derivation as part of the numeration. These elements are then selected and merged with other elements (i.e., D with N, and Top with IP) to form new syntactic objects: DP and TopP.

4.2. The activation of focus within D

A similar conclusion can be reached on the basis of wh-questions in Gungbe. This language does not have English-type wh-words. Instead, question words consist of a noun phrase and a

question particle, as in the forms *nú-té* 'what' and *fí-té* 'where' in (37a-b).

(37) a. [Nú-tɛ́] wɛ̀ Kòfĩ xɔ̀?
Thing-Q Foc Kofi buy
'What did Kofi buy?'

b. [Fí-tɛ́] wɛ̀ Kòfĩ yì?
Place-Q Foc Kofi go
'Where did Kofi go?'

As example (38) shows, the noun, in these complex wh-words, can be separated from the question particle by intervening modifiers. This is partial evidence that Gungbe wh-phrases are complex phrases where a phrase is associated with a Q-morpheme.

(38) [Fí jɔ̀-fĩ tɛ́] wɛ̀ Kòfĩ trɔ́n bò ɖɔ́ kú ɖíó mɔ́n?
Place well-known Q Foc Kofi come and has die dirt so
'What possible important place could he come from to be so dirty?'

Let us compare these wh-phrases to the format of focused constructions and wh-questions.

(39) a. [[é-**tɛ́**] **wɛ̀**] Kòfĩ xɔ̀?
3sg-Q Foc Kofi buy
'What did Kofi buy?'

b. *é wɛ̀ Kòfĩ xɔ̀ ?
3sg Foc Kofi buy

b. [[àkwékwè] **wɛ̀**] Kòfĩ xɔ̀
banana Foc Kofi buy
'Kofi bought BANANA(S)'

Here, we observe a striking parallel between wh-questions and focus expressions. The sentences under (39) indicate that, in wh-questions and focus constructions, the attracted DP is immediately adjacent to the focus marker *wɛ̀*. That Gungbe resorts to the same focus marker clearly shows that fronting in both cases involves the same functional head. Looking

closely at wh-questions, we further observe that the fronted wh-phrase must contain the Q-morpheme *té*.[17] The proper nature of this morpheme is not clear to me and it could be considered a pure question particle or a general operator sensitive to both focus and question operations, which I tentatively labeled here as Q. Given this description, the pattern in (39) is reminiscent of the data discussed in the context of topicalization, where we have shown that a marker inside the DP allows the latter to be attracted at the clausal level by a topic marker. Therefore, just as the topic head probes over a DP marked as definite or specific, the focus head, in relation to interrogative, appears to probe over question phrases that are marked by *té*. Accordingly, we have here another instance of a matching between a probe and its goal (see footnote 19 for the discussion on long distance agreement).

What these Gungbe facts tell us is that DPs (or else the relevant constituents) do not end up being topicalized or focused by chance, but because they embed the relevant Topic and Focus features. These are spelled out on the probe by lexical material in the form of the discussed markers.[18] The next question is what type of features are Topic and Focus

[17] On a purely methodological side, the pair in (39) shows that it would be highly undesirable to assume that question (39a) is a pure product of syntax, while (39b) is leftover to pragmatics even though the two derivations seem to be dragged by the focus marker.

[18] Assuming the view developed here, another relevant empirical domain to explore with regard to interrogative, topic, and focus features, is that of long distance Agree. Indeed, it has been shown in the literature that formal features can be checked under long distance Agree between the probe and the goal. The analysis put forward here suggests that the same should hold of information-structure-sensitive features. Though further work on this issue is in order, certain facts from Standard Arabic (SA), as discussed by Ouhalla (1993), suggest that focus may be checked under long distance Agree. Ouhalla (1993: 279) reports that SA allows both in-situ or fronted focused expressions as indicated below:

(i) a. ʔallaf -at Zaynab-u **qasiidat-an**
write(Perf)-3F Zaynab-Nom poem-Acc
'Zaynab has written a POEM.'

b. **qasiidat-an** ʔallaf -at Zaynab-u
poem-Acc write(Perf)-3F Zaynab-Nom
'It is a poem Zaynab has written.'

According to Ouhalla, (ia) encodes new information focus while (ib) expresses contrastive focus (Kiss 1998). In addition, it is possible to express contrastive focus with an in-situ constituent thanks to the focus markers *ʔinna* and *la*. The former occurs in sentence-initial position while the latter attaches to the focused in-situ constituent.

(ii) **ʔinna** Zaynab-a (**La**)-shaa'irat-un (laa riwaaʔiyyat-un)
FM Zaynab-Acc FM poet-Nom not novelist-Nom
'Zaynab is a POET (not a novelist).'

In terms of Ouhalla *ʔinna* heads a focus phrase, while *la* Ouhalla is comparable to "a tonic accent (focus stress) in that it serves the function of identifying the constituent which is being contrastively focused. This element can therefore be considered a constituent focus marker (p. 281)." This characterization recalls my own

compared to other (traditional) formal features (e.g., tense, aspect, mood, case)?

5. Topic and focus features

It seems clear from the Gungbe facts that nothing in principle distinguishes between Topic and Focus features and other formal features such as Tense, Case or ϕ-features.[19] We have already shown in previous sections that the licensing of the Topic or Focus head implies a matching operation that is necessitated by Probe-Goal relations. In this section, I take the comparison a step further.

An interesting fact about Gungbe (and Kwa in general) is that, in addition to discourse-related markers (e.g., Topic, Focus, Inter), they also involve INFL-related markers that encode tense, mood, aspect, negation, etc. (Aboh 2004a). Let us consider the category Tense in Gungbe. This language has a future tense marker *ná* only, illustrated in (40): here we see that the only difference between (40a) and (40b) is precisely the presence of this marker in the latter but not in the former. Thus, (40b) is interpreted with future reading while (40a) is read as perfective. The ungrammatical examples (40c-d) indicate that the subject cannot remain to the right of this marker (e.g., in [spec VP]), recall that Gungbe is SVO.

---

description of DP-internal topic and focus in Gungbe. In current Minimalism framework, Ouhalla's (1993) proposal can be seen as an instance of probe goal relation, where the feature focus under the probe is checked by long distance Agree between *ɲ̀nna* and *la*. If this characterization turns out to be the right one, then we may have to change our perspective on languages where there seems to be no segmental marking at all for topic and focus. Indeed, the fact that assignment of in situ focus often correlates with main accent/sstress in many languages (e.g., English) can be seriously taken as an instance of long distance agreement between the focus probe or the topic probe within the left periphery and the relevant goal within IP. This means that English is just a less obvious realization of the Arabic situation. This description would also suggest that ex situ versus in situ focus distinction boils down to the choice between Move and Long distance Agree in these languages.

[19] This need not mean that these features form a uniform class. While Case and ϕ-features relate to nouns only, topic and focus apply to all categories (see Aboh 2004). Similarly, Rizzi (2004) proposes that topic, focus, (and interrogative) are criterial features, but this does not apply to Case and ϕ-features. Further work is need for understanding the intrinsic properties of these features and how they relate to the observed differences. What matters for the present discussion though, is that the CHL manipulates these two classes of features the same way (I thank Anna Cardinaletti for commenting on this issue).

(40) a. Kòfí xɔ̀ xwé lɔ́
Kofi buy house Det
'Kofi (has) bought the house'

b. Kòfí **ná** xɔ̀ xwé lɔ́
Kofi Fut buy house Det
'Kofi will buy the house'

c. ***ná** Kòfí xɔ̀ xwé lɔ́
Fut Kofi buy house Det

d. *xwé lɔ́ **ná** Kòfí xɔ̀
house Det Fut Kofi buy

Aboh (2004a, chapter 5) demonstrates that the marker *ná* encodes T in Gungbe. In the context of the present discussion, and following Chomsky (1995: 237, ff), this amounts to saying that the Numeration that produced sentence (40b) contained *ná* as expression of T.

Now, let us contrast the sentence in (40b) to the topic and focus sentences in (41a-b). Examples (41c-d) show that these markers must necessarily be associated to some constituent immediately to the left.

(41) a. Kòfí **wɛ̀** xɔ̀ xwé lɔ́
Kofi Foc buy house Det
'KOFI bought the house'

b. Kòfí **yà** é xɔ̀ xwé lɔ́
Kofi Top 3sg buy house Det
'As for Kofi, he bought the house'

a. ***wɛ̀** Kòfí xɔ̀ xwé lɔ́
Foc Kofi buy house Det

b. ***yà** Kòfí xɔ̀ xwé lɔ́
Top Kofi buy house Det

Structurally, the future marker *ná*, the topic and focus markers *yà* and *wɛ̀* parallel in all respects. They all seem to be endowed with a strong feature (say, an EPP-feature) that forces adjunction of another category that they match with. Semantically, these markers have an effect as well. When we contrast the sentence in (40a) to those in (40b), (41a) and (41b), ignoring preposing operations, we realize that the merge of the relevant markers correlates with distinct semantic effects: future in (40b), focus in (41a) and topic in (41b). These

examples show that there is no flexibility in the displacement operations leading to Topic and Focus constructions in Gungbe. Movement is mandatory in the context of the topic and focus markers and the interpretation of the formed expressions is invariably: XP-*yà* = topic, and XP-*wɛ̀* = focus. In addition, the sentences provide us with further evidence that Tense, Topic and Focus functional categories are present in the lexicon in the form of markers with the intrinsic feature (e.g., Tense, Topic, Focus), which are selected from the numeration.[20] Yet, this conclusion does not answer the question of how the attracted element/constituent, that is, the goal, gets assigned the Topic or Focus feature.

This question takes us back to the comparison with Case because it is in a sense reminiscent to the issue of how a DP attracted to [spec TP] gets nominative case and is interpreted as subject. The question is whether the attracted nominal is listed in the lexicon with its Case (e.g., nominative vs. accusative) and ϕ-features or whether these are assigned arbitrarily as the noun enters the numeration. In the case of Interrogation, Topic, and Focus as discussed here, we have shown that some syntactic operations (including feature-matching) were involved that indicate that the goal must be marked for the relevant features on the probe. Adopting Chomsky's (1995: 235ff) view on Case, it seems hard to assume that lexical categories (e.g., N) are listed in the lexicon with corresponding Topic and Focus features. In terms of Chomsky (1995: 236):

> "a lexical entry should not indicate that [a noun N], has Case and ϕ-features; that follows from its being of category N and the fact that Case and ϕ-features have to be assigned [….] derives from general principles, and nothing intrinsic to the lexical entry […] tells us that a particular occurrence is singular or plural, nominative or

[20] The same could be said of the DP where the determiners (i.e., the specificity markers) pair with the Q-particle inside D as illustrated by the following examples.

(i)	a.	Dáwè	ɖé	b.	Dáwè	lɔ́	c.	Dáwè	tɛ́
		man	Det		man	Det		man	Q
		'A certain main'			'The specific man'			'Which man'	

Here again, we observe that, in addition to having a clear semantic effect, these markers have an EPP feature that translates into fronting of the NP-complement to their left (see Aboh 2004a, chapters 3, 4, 2004b).

accusative".

The same, I claim, holds of Topic, Focus, and Interrogative features. These are optional formal features that are added as lexical items enter the numeration.[21] Chomsky (1995: 236) further proposes that the choice of a lexical item, say X, is a two-steps process: step (1) forms the numeration containing X and its index, and step (2) introduces it in the derivation. It is further proposed that the optional features (e.g., Case and ϕ-features) are most probably added by step (1). Given my stance that Topic, Focus, and Interrogative are optional formal features, nothing need be added. Following Chomsky (1995) therefore, I assume the null hypothesis that, similarly to Case and ϕ-features, core features of Information Structure (e.g., Topic, Focus, Interrogative) are arbitrarily added as the relevant categories are introduced in the numeration. The main advantages of the proposed analysis are that:

- It is compatible with the fact that topicalized, focused constituents, and wh-expressions show the same distributive restrictions cross-linguistically (Cinque 1990, Rizzi 1997, Aboh 2004a)—thus reminding us of Vergnaud's guidance on case.
- It does not further enrich the theoretical apparatus with any new device that is not firmly rooted in robust empirical data.
- In addition to rigorously complying with the inclusiveness condition, this analysis supports the view, often disputed in the literature, that "even formal features typically have semantic correlates and reflect semantic properties (Chomsky 1995: 381, footnote 14)."

[21] This idea is not totally new. See for instance, Rizzi (2004) for some recent formulation. In terms of Pesetsky and Torrego (2004) and subsequent, what this means is that the relevant DPs being valued for the features topic and focus, provide a value for the unvalued Topic and Focus probe within the clause.

6. Conclusion

This paper argues for the strong minimalist view that Information Structure starts in the numeration in the form of discourse-related lexical items which drive the derivation. In many languages, these discourse-related particles, which encode among other features Interrogative force, Topic, or Focus, show very specific syntactic behavior indicating that they project in syntax. The analysis further shows that these features are comparable to other optional formal features (e.g., Case, ϕ-features) that are added arbitrarily when the lexical item enters the numeration.

References

Abney, Paul S. 1987. The English Noun Phrase in its Sentential Aspect. Ph.D. dissertation, MIT.

Aboh, E. O. 2004a. *The morphosyntax of Complement-Head Sequences. Clause structure and word order patterns in Kwa*. New York: Oxford University Press.

Aboh, E. O. 2004b. "Left or Right? A view from the Kwa Periphery". In D. Adger, C. de Cat and G. Tsoulas eds., Peripheries. Syntactic Edges and their Effects. Dordrecht: Kluwer. 165–189.

Aboh, E. O. 2004c. "Topic and Focus within D". *Linguistics in the Netherlands* 21: 1–12.

Aboh, E. O. 2004d. "Snowballing Movement and Generalized Pied-piping". In A. Breitbarth and H. van Riemsdijk eds., Triggers. Berlin: Mouton.15–47.

Aboh, E. O. 2006a. "Complementation in Saramaccan and Gungbe: the case of c-type modal particles". *Natural Language and Linguistic Theory* 24,1:1–55.

Aboh, E. O. 2006b. "The Role of the Syntax-Semantics Interface in Language Transfer". In L2 Acquisition and Creole Genesis, Dialogues. C. Lefebvre, L. White, and C. Jourdan eds., Amsterdam: Benjamins. 221–252.

Aboh, E. O. & R. Pfau. 2006. "What's a wh-word got to do with it?". To appear in Benincà, P. & N. Munaro eds., *The Cartography of CP*. New York: Oxford University Press.

Amha, A. 2001. *The Maale Language*. Leiden: CNWS publications.

Baltin, M. 2006. Doubly-filled comps are not doubly-filled comps. Ms. New York University.

Bernstein, J. 1997. "Demonstratives and Reinforcers in Romance and Germanic languages." *Lingua* 102: 87–113.

Bernstein, J. 2001a. "Focusing the 'Right' Way in Romance Determiner Phrases." *Probus* 13,1: 1–29.

Bošković, Ž. 2001. *On the nature of the syntax-phonology interface. Cliticization and related phenomena*. Amsterdam: Elsevier.

Bošković, Ž. 2002. "On multiple wh-fronting". *Linguistic Inquiry* 33: 351-383.

Cheng, L. 1991. *On the Typology of Wh-questions*. Ph.D. dissertation, MIT.

Cheng, L. & J. Rooryck. 2000. "Licensing wh-in-situ". *Syntax* 3: 1-19.

Chomsky, N. 1971. "Deep Structure, Surface Structure and Semantic Interpretation". In D. Steinberg & L. Jakobovits eds., *Semantics*: *An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge: Cambridge Univerisity Press. 183–216.

Chomsky, Noam. 1977. "On Wh-movement." In Peter, W. Culicover, ed., *Formal Syntax*. Academic Press, New York, San Francisco.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, N. 2005. On Phases. Ms, MIT.

Cinque, G. 1990. *Types of A'-Dependencies*. MIT Press, Cambridge, Mass.

Cinque, G. 1999. *Adverbs and Functional Heads, A Cross-linguistic Perspective*. New York: Oxford University Press.

Dyakonova, M. 2004. “Information Structure Development: Evidence from the Acquisition of Word Order in Russian and English”. In K. Bentzen ed., Nordlyd 32, 1: *Tromsoe Working Papers in Language Acquisition*: 88–109.

Fanselow, G. 2006. “On pure syntax (Uncontaminated by information structure)”. In Brandt, P. & Fuss, E. eds., *Form, Structure and Grammar*. Berlin: Akademie Verlag.

Fanselow, G. 2007. “The Restricted Access of Information Structure to Syntax—A Minority Report”. In G. Fanselow and M. Krifka eds., The Notions of Information Structure. Interdisciplinary Studies on Information Structure 6: 205–220.

Frajzyngier, Z. 2001. *A Grammar of Lele*. Stanford, CA: CSLI.

Gundel, J. K. 1974. *The Role of Topic and Comment in Linguistic Theory*. Ph.D. dissertation, University of Texas, Austin.

Ionin, T. 2006. “This is definitely specific: Specificity and definiteness in article systems”. *Natural Language Semantics* 14: 175–234.

Jakendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass:. MIT Press.

Kiss, K. E. 1998. “Identificational focus versus information focus”. *Language* 74: 245-273.

Li, C. & S. A. Thompson. 1975. “The semantic function of word order: A case study in Mandarin” In C. Li ed., *Word order and word order change*. Austin: University of Texas Press.

Li, C. & S. A. Thompson. 1981. *Mandarin Chinese*. *A functional reference grammar*. Los Angeles. University of California Press.

Longobardi, G. 1994. “Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form.” *Linguistic Inquiry* 25: 609–665.

Marantz, A. 2000. “Case and Licensing”. In E. Reuland ed., *Arguments and Case*: *Explaining Burzio’s. generalization*. Amsterdam: Benjamins , 11–30.

Merchant, J. 2004. "Fragments and Ellipsis". *Linguistics and Philosophy* 27: 661–738.

Ouhalla, J. 1993. "Negation, focus and tense: the Arabic maa and laa." *Rivista di Linguistica* 5: 275–300.

Pesetsky, D and Torrego, E. 2004. The Syntax of Valuation and the Interpretability of Features. Ms. MIT and UMass/Boston.

Prince, F. E. 1981. "Toward a Taxonomy of Given-New Information". In P. Cole, ed., *Radical Pragmatics*. New York: Academic Press. 223–255.

Rizzi, L. 1997. "The Fine Structure of the Left Periphery". In L. Haegeman ed., *Elements of Grammar*. Dordrecht: Kluwer. 281–337.

Rizzi, L. 2004. On the Form of Chains: Criterial Positions and ECP Effects. Ms. University of Siena.

Samek-Lodovici, V. 2006. "When Right Dislocation meets the Left-Periphery. A Unified Analysis of Italian Non-Final Focus". *Lingua* 116: 836–873.

Szabolcsi, A. 1987. "Functional Categories in the Noun Phrase." In István Kenesei ed., *Approaches to Hungarian.* JATE, Szeged.

———. 1994. "The Noun Phrase." In Ferenc Kiefer & Katalin E. Kiss, eds., *Syntax and Semantics. The Syntactic Structure of Hungarian* 27: 179–274. Academic Press, New York.

Szendroi, K. 2001. *Focus and the Syntax-Phonology Interface*. Doctoral dissertation, University College London.

Valduvi, E. 1990. *The Informational Component*. Ph.D. dissertation, University of Pennsylvania, Philadelphia.

Zubizaretta, M. L. 1998. *Prosody, Focus and Word Order*. Cambridge, Mass.: MIT Press.

Zwart, J.W. 1998. “Where is Syntax? Syntactic Aspects of Left Dislocation in Dutch and English”. In P. Culicover and L. McNally eds., *The Limits of Syntax*. *Syntax and Semantics 29*. New York: Academic Press. 365–393.