Topic and focus within D*

Enoch O. Aboh
University of Amsterdam

1. Introduction

This paper proposes that, even though topic and focus are often considered to be clausal properties, these specifications may be encoded within the nominal structure as well. It is argued that the noun system involves an articulated left periphery, comparable to the clausal complementizer system, which involves discrete topic and focus projections, whose specifiers contain the fronted nominal topic and focus constituents. In some languages, the heads of these designated projections are morphologically realized as elements that are commonly known as determiners or articles. Section 2 sets the stage for the discussion, Section 3 discusses the topic-focus articulation within D, and Section 4 concludes the paper.

2. Setting the stage

The literature on topic and focus constructions often refers to situations where an element of discourse is seen as old information or important, and is marked as such syntactically or prosodically. In the Gungbe example (1), both the topicalized subject Setù and the focused object Mǻri appear within the left periphery where they are marked by the topic and focus markers, respectively. Note that the topic and focus constituents are sandwiched between the complementizer d‘ ‘that’ and the IP-internal third person subject pronoun è.

(1) Ùn nywèn d‘ [Sètù] yà [Mǻri] wè è dà
1sg know that Setu top Mary foc 3sg marry
‘I know that, as for Setu, he married Mary’
Granted that topic and focus may target different constituents (e.g., arguments, adjuncts, adverbs, adjectives, verbs, or propositions, see Aboh 2004), one can conclude that topicalization and focalization are crucially clausal properties of which noun sequences only happen to be a target. There are, however, both empirical and theoretical reasons to suggest that the topic-focus articulation projects within the nominal domain as well. The following data provide partial evidence for such an approach.

Gungbe distinguishes between non-specific (i.e., non-discourse anaphoric) and specific (i.e., discourse anaphoric) noun phrases. Non-specific noun phrases surface as bare noun phrases, as in (2a). In this example, the noun phrase *lɛsi gukmekem fn* is interpreted as definite, but not specific, because it is not pre-established in discourse. The example (2b), on the other hand, indicates that specific noun phrases must occur to the left of the specificity marker *lɛ*. Here, the sequence *lɛsi gukmekem fn* is definite specific because it refers back to some particular rice that is pre-established in discourse or is known to the participants. This reminds us the notion of assumed familiarity as discussed in Enç (1991), or that of specificity as described in Prince (1981).

(2) a. Sɛtɔ no xɔ [lɛsi Gukomem fn]
   'Setu habitually buys the rice from Gukome'

   b. Sɛtɔ no xɔ [lɛsi Gukomem fn lɛ]
   'Setu habitually buys the aforementioned rice from Gukome'

If specificity (or assumed familiarity) and topicality are related in some sense, then the contrast in (2) suggests that the noun sequence in (2b) is marked for topicality. Yet, that sequence can be subsequently topicalized in the clause alongside with (2a), as shown in (3a) and (3b), respectively.

(3) a. [lɛsi Gukomem fn lɛ] yɛ nɔ vivi gbáu
   'As for the aforementioned rice from Gukome, it is very sweet'

   b. [lɛsi Gukomem fn] yɛ nɔ vivi gbáu
   'As for the rice from Gukome, it is very sweet'

The reading in (3a) is that of a specific topic noun phrase (that is, a known/given referent), which is the topic of discussion. In (3b), however, the topic of discussion is expressed by a bare noun phrase that may be interpreted as (±definite) or (±generic) depending on the context. In this regard, the contrast between (2b) and (3b) is that the former represents a nominal topic, while the latter is a clausal topic. These facts lead me to suggest that there is topic specification both within the nominal left periphery and the clausal left periphery. Under this approach, the question arises how does nominal topicalization relate
to clausal topicalization? I return to this issue in Section 3.2 where I propose that nominal topic may license clausal topic.

With regard to focusing within the noun phrase, Bernstein (2001) suggests that the Romance sentences in (4) differ in that the pre-nominal demonstrative in (4a) manifests a more basic word order and triggers an unmarked or neutral interpretation. On the other hand, the post-nominal demonstrative bears main stress and yields a focus reading (4b).

(4) a. Este libro interesante [Spanish]
   this book interesting
   'This interesting book'
b. El libro interesante este
   the book interesting this
   'This interesting book' [Bernstein 2001:2]

This suggests that focus may be determined within the nominal domain. Yet, as the French examples under (5) show, a nominal that involves focus specification (i.e., ci is the French counterpart of Spanish este) can be attracted to the focus field within the clause just as any other nominal constituent.

(5) a. C’est ce livre que Jean a lu qui le rend triste
   it.is the book that John has read that him make sad
   'It is this book that John read, which makes him sad'
b. C’est ce livre-ci que Jean a lu avant de partir
   it.is this book-here that John has read before Prep leave
   'It is this book here that John read before he left'

In the cleft sentence (5a), the definite noun phrase ce livre may be interpreted as contrastive, new information, or presentational focus depending on the context. In example (5b), however, the clefted noun phrase ce livre-ci is necessarily interpreted as contrastive focus because the noun phrase includes the demonstrative reinforcer ci. We therefore reach a situation that is similar to that of the topic whereby the clausal focus attracts a noun phrase that has been assigned focus within D. These facts lead me to conclude that the noun system allows the expression of topic and focus. Granted that topic and focus are the properties of distinct projections within the left periphery, I further propose that the nominal system involves a left peripheral topic-focus articulation.

2.1 The split-D hypothesis

This paper proposes that the facts described in (2) through (5) reflect properties of the nominal left periphery, that is, the D(eterminer)-system. Under this view, the D-system is comparable to the clausal left periphery: the C(omplementizer)-system. Like the C-system, the D-system involves topic and focus projections (TopP and FocP) whose specifiers host topic and focused constituents. TopP
and FocP project between DP, the highest projection of the system, which expresses the interface between the discourse and the nominal expression, and NumP, the lowest projection, which links the D-system to the nominal I(nflectional)-system. As such, NumP encodes the agreement features and certain referential features (e.g., number, deixis) that parallel those of the nominal I-system (Aboh 2004). Following Campbell (1996), I also assume that noun phrases involve covert predication whereby the noun head functions as a predicate of the referent of the noun phrase. This would mean that the nominal I-system (i.e., FP in the representation 6) includes a subject position that may host the possessor in possessive constructions. Several empirical facts support the split-D analysis.

(6) $[\text{DP} \ldots [\text{D} \ldots \text{topic} \ldots \text{focus} [\text{NumP} \ldots [\text{Num} \ldots [\text{FP} \ldots \text{N} \ldots ]]]]]$

2.1.1 $D$ is the nominal equivalent of Force

Szabolcsi (1994) reports the following contrast in Hungarian.

(7) a. (a) Mari kalap-ja
   the Mari(-nom) hat-poss-3sg
   'Mary’s hat’

   b. Mari-nak a kalap-ja
      Mari-dat the hat-poss-3sg
      'Mari’s hat’

   c. [Péter-nek] mindenki csak a kalap-já-t láttat
      Peter-dat everyone only the hat-poss-3sg-acc saw
      ‘As for Peter, everyone saw only his hat (e.g., no one saw his coat)’ [p. 205]

The data under (7) indicate that the Hungarian nominative possessor always follows the determiner, as in (7a), while the dative possessor always precedes, as in (7b). Example (7c) further shows that the dative possessor may be wh-extracted out of the noun sequence, hence the so-called possessor extraction construction. According to Szabolcsi the paradigm in (7b–c) is comparable to subject extraction in the clause, via [SpecForceP]. Under this approach, the possessor moves to [SpecDP] where it enters a Spec–Head configuration with the subordinator a, the nominal counterpart of the clausal complementizer, e.g., that in English (7b). In the possessor extraction construction, however, the possessor subsequently extracts from [SpecDP] giving rise to example (7c). This would mean that the contrast in (7a–b) could be represented as in (8a–b), where DP is the nominal equivalent of CP and IP represents the nominal inflectional domain (see Knittel 1998, and Haegeman 2004 for alternative approaches).

(8) a. $[[\text{DP} [\text{IP} a [\text{IP} \text{Mari}]_{\text{nom}}] [\text{IP} \text{minden} [\text{IP} \text{kalap-ja}]]]]$

   b. $[[\text{DP} \text{Mari-nak}]_{\text{acc}}] [\text{IP} a [\text{IP} \text{I}]_{\text{IP} \text{kalap-ja}}]]$
In terms of this analysis, the determiner in D is a subordinator, that is, a C-type element. On the other hand, the article is an I-type element, which merges under I and encodes nominal inflectional features such as definiteness. This analysis leads to two main conclusions: (i) [SpecDP] is a non-thematic operator position, as indicated by the dative extraction constructions under (7b–c), and (ii) the expression of definiteness is primarily determined within the nominal I-system, but may be reflected in the nominal left periphery (i.e., the D-system) due to definiteness concord (Aboh 2004).

2.1.2 Num0 as the expression of [number] and [definiteness]

I propose that NumP, which delimits the D-system downward, represents the juncture where definiteness concord is achieved. Put differently, NumP encodes the INFL features (e.g., deixis) that match those expressed in the I-system. Under this approach, Num0 is comparable to the clausal Fin0, which has also be shown to express inflection or agreement features that match those of the embedded IP (see Haegeman 1992, Rizzi 1997, Paoli 2001).

The following Gungbe facts on the number marker l’e, which realizes Num0, support this view. In (9a), the sequence àkwékwè àt’sn ‘five bananas’ is indefinite, and may refer to any five bananas (e.g., out of a heap of ten). In (9b), however, the sequence àkwékwè àt’sn l’e is definite and denotes a pre-identified set of five bananas (e.g., each heap is made of five bananas). This contrast indicates that a sequence NP-l’e is necessarily interpreted as [+definite, +plural]. I (2004) propose that l’e encodes a bundle of features including [definite, number].

(9) a. Mì sà àkwékwè àt’sn ná mì
2pl sell banana five for 1sg
‘Sell me five bananas’

b. Mì sà àkwékwè àt’sn l’e ná mì
2pl sell banana five numb for 1sg
‘Sell me the five bananas’

Similarly, the number marker encodes agreement features because expressing numerals inside a specific definite noun phrase requires the presence of number l’e. Recall from previous discussion that the number and specificity markers need not co-occur.

(10) Mì sà àkwékwè (l5) (l’e) ná mì
2pl sell banana det[+specific] numb for 1sg
‘Sell me (the aforementioned) banana(s)’

However, the complex noun-numeral àkwékwè àt’sn ‘five bananas’ cannot be marked as specific if number is absent. Contrast the grammatical example (11a), where the noun phrase includes l’e-l’e, and refers to a pre-established set of five bananas, to the ungrammatical noun phrase (11b), which lacks number l’e.
These examples suggest that Num₀, realized as \( l \), establishes a concord between the definiteness and plurality expressed in the nominal I-system and the D-system. In addition, the ungrammatical example (11b) indicates that DP dominates NumP, since \( l \) must merge before \( \text{Num} \).

3. The topic-focus articulation

The previous discussion suggests that D and Num are the nominal counterparts of the clausal Force and Fin, respectively. In what follows, I argue that the topic-focus articulation projects within these two borderlines. Recall that Gungbe manifests bare NPs. These sequences are interpreted as (in)definite or generic depending on the context. They surface without the specificity or plural markers, but may include nominal modifiers (e.g., demonstratives, adjectives, numerals). In (12a), for instance, the bare NP \( \text{távò} \) ‘table’ is interpreted as indefinite. In (12b), however, the sequence \( \text{távò títán} \) ‘first table’ is interpreted as definite.

\[ \text{(12)} \]
\begin{align*}
\text{a.} & \quad \text{Kókù hǹ́n \text{távò wá xwegbè} } \\
& \quad \text{Koku hold table come house} \\
& \quad \text{‘Koku brought a table home’} \\
\text{b.} & \quad \text{Kókù wè hǹ́n \text{távò títán wá xwegbè} } \\
& \quad \text{Koku Foc hold table first come house} \\
& \quad \text{‘It is Koku who brought the first table home’}
\end{align*}

This contrasts with example (13a), where the sequence \( \text{távò cè} \) ‘my table’ is understood as definite non-specific, unlike the sequence \( \text{távò l} \), which is interpreted as definite specific because it is pre-established in discourse. A similar contrast arises in example (13b) where the sequence \( \text{távò dè} \) is interpreted as indefinite specific as opposed to the sequence \( \text{távò cè} \). The ungrammatical example (13c) suggests that the specificity \( l \) and \( d \) exclude each other because they express complementary sets of features [+specific; ±definite].

\[ \text{(13)} \]
\begin{align*}
\text{a.} & \quad \text{Kókù m̀n \text{távò cè} bò d ṣ̄ m̀n ná xò \text{távò l} } \\
& \quad \text{Koku see table 1sg-poss and say 3sg fut buy table det [+spec, +def]} \\
& \quad \text{‘Koku saw my table and said that he would buy that table’} \\
\text{b.} & \quad \text{Kókù m̀n \text{távò cè} bò d ṣ̄ m̀n ná xò \text{távò dè} } \\
& \quad \text{Koku see table 1sg-poss and say 3sg fut buy table det [+spec, -def]} \\
& \quad \text{‘Koku saw my table and said that he would buy a certain table’}
\end{align*}
The descriptive generalization is that a specific noun phrase requires the markers 1 or δé as the morphological realizations of the features [+specific, +definite], and [+specific, −definite], respectively. This would mean that bare NPs are necessarily specified as [−specific, +definite] or [−specific, −definite]. In previous work, I (2004) proposed that the specificity feature on the Gungbe head noun must be checked against D, which also bears the feature [+specific] encoded by the marker 1. This requirement is achieved by nominal predicate fronting in Gungbe, because the language lacks N-to-D raising. Accordingly, the nominal predicate (i.e., FP in 14) is pied-piped to [SpecDP] (via [SpecNumP]).

(14) \[DP [FP távò] [D 15 [NumP távò [Num 16 17]]]]

But there are good reasons for refining this analysis along the lines of (15a) for the Gungbe-type languages, which involve predicate fronting to [SpecTopP]. Under this approach, the predicate is attracted to [SpecTopP] to check the features [specific] under Top, which hosts the specificity marker. In the English-type languages, however, one may propose along the lines of Campbell (1996) that [SpecTopP] is filled by a null operator that binds the subject of predicate headed by the noun. Campbell (1996:162) further assumes that “the specificity operator is a kind of DP-internal topic, which links the internal subject position (and hence the DP itself) to a referent identified previously in the discourse.”

(15) a. \[DP [TP távò] [TopP [TP távò] [Top0 18 19]]]]
   b. \[DP [TopP 20 21 [TP 22 the [TP 23]]]]

3.1 Question marking and the topic-focus articulation within D

Pieces of evidence for refining the representation (14) as in (15a) come from nominal question marking in Gungbe. This language displays complex question words that consist of a noun phrase and a question marker as indicated by the forms nú-té ’what’ and fi-té ’where’ in (16a–b).

(16) a. [Nú-té] we Kófi xé? thing-q foc Kofi buy ‘What (thing) did Kofi buy?’
   b. [Fi-té] we Kófi yí? place-q foc Kofi go ‘Where did Kofi go?’

The sentences in (17) further show that the questioned noun can be separated from the question word by intervening modifiers.

(17) a. [Távò xóxó té] we Kófi xé? table old q foc Kofi buy ‘Which old table did Kofi buy?’
These data indicate that the complex forms in (16) cannot be analyzed as [X-Y] adjunction structures where the questioned noun (X) and the question particle (Y) are two (lexical) heads. Instead, the bracketed sequences in (16–17) seem to involve structures where the element to the left of the question particle Q represents a phrase that is in the specifier of a Q-phrase, as schematized in (18).

(18) \[[\text{DP} \ldots \text{[Q-phrase XP \{Q, t’\}]]}\]

The question now arises what is the nature of this Q-phrase? A parallel between the clausal and nominal question formation and focus constructions may help answer this question. The Gungbe focus and wh-questions require fronting of the focused constituent or wh-phrase to the left of the focus marker, as in (19a–b). The ungrammatical example (19c) indicates that wh-phrases and focused constituents are in complementary distribution.

(19) a. \([\text{[àkwékwè] w’} \text{Kòfí x’}]\) Kòfí x’
   banana foc Kofi buy
   ‘Kofi bought banana(s)’
   b. \([\text{[é-t’} w’} \text{Kòfí x’}]\) Kòfí x’?
   3SG-Q foc Kofi buy
   ‘What did Kofi buy?’
   c. *é-t’ Kòfí w’ x’?
   3SG-Q Kofi foc buy
   ‘What Kofi bought?’

These facts suggest that focused constituents and wh-phrases target the same position within the C-system. Accordingly, the focus head w’ realizes the feature \[F\] that is checked by focused- or wh-phrases as shown in (20a). By extending this analysis to the nominal domain, I reformulate (18) as in (20b), where the Q-phrase is a focus phrase whose head is realized by the element t’.

(20) a. \([\text{FocP XP|\{F\}|}\text{Foc w’}[\text{FinP \ldots t XP \ldots}]]\]
   b. \([\text{FocP XP|\{F\}|}\text{Foc t’}[\text{NumP \ldots t XP \ldots}]]\]

This would mean that the D-system involves both a topic phrase (TopP) and a focus phrase (FocP), whose heads are realized as l’ and t’, respectively. An apparent counter-argument to this analysis, though, is the incompatibility between the nominal question marker and the specificity markers.

(22) *\([\text{[Távò x’x’5 ls/qé} t’] w’} \text{Kòfí x’}]\)
   table old det \[\text{[spec,def]}\] Q foc Kofi buy
   ‘Which old aforementioned table did Kofi buy?’

In previous work, I (2004) concluded that these facts are expected if we assume that both the nominal question marker and the specificity marker encode D. In
that system, the features [specific] and [interrogative] were amalgamated under D. But this seems counter-intuitive because interrogative features express some kind of force (Cheng 1991), while specificity relates to discourse-linked properties, such as known or shared information. In addition, a fine structure of the clausal C-system suggests that the functional head that hosts the feature interrogative (e.g., in Gungbe and Italian) is different from Force, which encodes clausal-type (Aboh 2004). If we maintain the parallelism between the clausal and nominal left peripheries, we are led to conclude that the nominal topic and interrogative features do not fall under the same head D. Put differently, the specificity and the focus markers ḷ and ṭ do not compete for the same position.

This conclusion is supported by the fact that, in Gungbe, sentence (17a), repeated here as (23a), has a second variant that is given in (23b), where the element dë intervenes between the fronted phrase and the question word ṭ. It appears that speakers choose the strategy in (23b) in order to single out a particular (or specific) referent. Example (23c) further indicates that the question word and the particle dë may co-occur with the number marker.

(23) a. [Távò x̄f x̄f ṭ] ẉ Kòfí x̄f
   table old q foc Kofi buy
   ‘Which old table did Kofi buy?’

b. [Távò x̄f dë ṭ] ẉ Kòfí x̄f
   table old part q foc Kofi buy
   ‘Which one of the old table(s) did Kofi buy?’

c. [Távò x̄f dë ṭ ḷ] ẉ Kòfí x̄f
   table old part numb foc Kofi buy
   ‘Which one of the old tables did Kofi buy?’

It is not clear to me what is the origin of the particle dë, which may arguably derive from the relative marker dë(ɔx) ‘that/who/or the numeral (ɔ)dë ‘one’. Whichever the case, it seems reasonable to assume that this particle is a head that encodes the feature [one], which can be thought of as being discourse-anaphoric, just as specificity. Granted this, I further argue that, in situations such as (23b–c), the fronted element occurs in [SpecTopP] and binds an empty category in [SpecFocP]. Top and Foc are realized as dë and ṭ, respectively (24).

(24) \[p_{[Távò x̄f x̄f dë(ɔx)]} [SpecTopP távò x̄f x̄f dë(ɔx)] [SpecFocP ec(lying) x̄f x̄f] [focutch ṭ [NumP ṭ x̄f x̄f] [Num ṭ x̄f x̄f]]]]]]]]

Keeping the parallel with the clause, such nominal sequences are reminiscent of topic sentences in which the focused constituent is subsequently topicalized as the example (25) indicates. In this case, the topic constituent has been extracted from the focus position, which now contains the third person strong pronoun ɔf. 
Building on the discussed facts, I propose to represent the architecture of the nominal left periphery as in (26).

(26) \[ D \quad \text{TopP} \quad \text{Top0} \quad \text{FocP} \quad \text{Foc} \quad \text{NumP} \quad \text{Num} \quad \text{FP} \quad \text{N} \quad \text{N} \quad \text{]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]]}

Under this approach, I conclude that the incompatibility in (22) is pragmatic/semantic, but I leave it to future research, what pragmatic/semantic properties are at work here.

### 3.2 The interaction between nominal and clausal topics

This section suggests that a nominal topic may license a clausal topic. Gungbe manifests argument versus adjunct asymmetry with regard to topicalization. Unlike arguments, not all adjuncts can function as topic in this language. Note the contrast between the time adjunct and the locative adjunct in (27).

(27) a. \[ Gbójé \quad \text{mè} \quad \text{yà} \quad \text{Kofi \quad nà \quad sén \quad dó \quad xò \quad ló} \quad \text{holiday \quad post[inside]} \quad \text{top Kofi \quad fut \quad paint \quad prep room det[spec,del]} \quad \text{]}

   'As for during the holidays, Kofi will paint the room'

b. \[ Xò \quad \text{kpá} \quad \text{yà} \quad \text{Kofi \quad nà \quad zà \quad flén \quad gbáú} \quad \text{room \quad post[beide]} \quad \text{top Kofi \quad fut \quad sweep \quad there \quad indeed} \quad \text{]}

   'As for beside the room, Kofi will sweep there!'

I (2004) proposed that the contrast in (27) can be explained, if we assume that the topic marker \text{yà} is referential in the sense that it expresses the fact that the element to its left refers to specific member(s) of a set in the mind of the speaker or pre-established in discourse. Under this interpretation, expressions like 'holiday' are referential because they express shared knowledge, and can therefore be topic. This is not the case with the bare sequence \text{xò kpá}, 'room beside', where the speaker might be referring to any side of any room. In this regard it is interesting to notice that (27b) is grammatical if we insert the specificity marker, that is, if we make the room specific, as in (28).

(28) \[ Xò \quad \text{ló \quad kpá} \quad \text{yà} \quad \text{Kofi \quad nà \quad zà \quad flén \quad gbáú} \quad \text{room \quad det \quad post[beide]} \quad \text{top Kofi \quad fut \quad sweep \quad there \quad indeed} \quad \text{]}

   'As for beside the aforementioned room, Kofi will sweep there!'

Under the proposed analysis, that \text{xò ló kpá} 'beside the aforementioned room' is eligible for clausal topicalization is straightforward because it is itself a topic. This would mean that nominal topicalization may favor or license clausal topicalization because D and C interact at the interface level. But such conclusion need not be surprising though if we grant recent developments that C (and now D) is a
strong phase. In addition, the proposed analysis makes strong predictions about the distribution or existence of pure bare nouns, but I hope to return to these issues in future work.

4. Conclusion

This paper argues that the D-system involves topic and focus projections, whose specifiers contain the fronted topic and focus constituents, and whose heads are morphologically realized by elements that are commonly known as determiners or articles. TopP and FocP project between DP, the highest projection of the D-system, which represents the interface between discourse and the nominal expression, and NumP, the lowest projection of the D-system that expresses the interface between the nominal left periphery and the nominal inflectional system.

Notes

* I thank L. Haegeman, H. Zeijlstra, the participants of the TIN-dag, and an anonymous reviewer for their valuable comments, which helped improve this paper significantly.

1. Alternatively, one could propose that the article moves to Top to encode the topic feature. The noun may remain in situ or move to some intermediate position within the I-system.

References