Contextuality of Syntax and Adieu to the A/A'-Distinction

Željko Bošković
University of Connecticut

Abstract: The paper highlights a broader theoretical move toward contextuality in syntax, which has largely gone unnoticed because the contextuality is manifested quite differently in different syntactic domains, making the contextuality connection across them not that easy to discern. In particular, this theoretical move toward contextuality is shown to be manifested in the following domains: structure building, labeling, islandhood, phases, phasal edges, spell-out domains, and the EPP. The contextuality of syntax is also shown to be non-Markovian in nature, hence it argues against treating syntax as Markovian in nature. One Markovian property is, however, maintained: no look-ahead power. Most importantly, it is shown that the contextual approach to the EPP opens the door for eliminating the A/A'-distinction from binding, in a way where all that is needed for binding is c-command: with traditional A'-movement, the issue is that the nominal simply fails to c-command.

Keywords: A/A'-distinction, binding, EPP, islands, labeling, structure building, phases, spell-out domains

1. Introduction

The goal of this paper is to highlight a broader theoretical move toward contextuality/context-sensitivity in syntax, which has largely gone unnoticed for two reasons: it spans different syntactic domains and mechanisms, hence it's not easy to notice if one's attention is focused on a single domain; second, the contextuality is manifested quite differently in different domains, which makes the contextuality connection across different domains not that easy to discern. In particular, we will see that contextuality characterizes recent developments regarding locality of movement and islandhood, structure building and labeling, and the EPP—spell-out will also be argued to be context-sensitive in that phase XP does not always behave in the same way regarding spell-out, depending on its syntactic context.

The paper has two goals beyond highlighting the contextual nature of syntax. First, there are conflicting claims regarding whether syntax is Markovian in nature: Chomsky (2021) endorses a Markovian nature of syntactic derivations, while one of the main points of Chomsky (1957) was that syntax is non-Markovian. This paper addresses the issue from a new perspective, namely, by examining what the contextual nature of syntax discussed here means for this debate.

Second, the paper will explore consequences of the contextual approach to the EPP for binding. In particular, it is shown that the contextual approach to the EPP naturally leads to the conclusion that elements undergoing traditional A'-movement are actually not nominal. The paper will take advantage of that to develop a preliminary approach to binding that dispenses with the A/A'-distinction, where all that is needed for binding is c-command.

The paper is organized as follows. Section 2 discusses the history of research on the locality of movement and structure building, showing that it has consistently emphasized the role of contextuality in these domains. Section 3 discusses the contextuality of spell-out domains. Section 4 turns to the contextuality of the EPP, which will be the focus of this paper, and the more general issue of whether syntax is Markovian in nature. Section 5 shows that the contextual approach to the EPP opens the door for eliminating the A/A'-distinction from the binding theory.

2. Locality of movement and structure building

Consider the history of research on locality of movement/islandhood. In the bounding approach (Chomsky 1973), the trouble-makers for movement were defined rigidly: NP and IP were bounding nodes regardless of their syntactic context. *Barriers* (Chomsky 1986) were quite different but the importance of one difference went largely unnoticed—the contextuality of *Barriers*. One cannot even ask whether e.g. CP in general is a barrier. This depends on the syntactic context where it occurs; in *Barriers,* trouble-makers for movement were defined <u>contextually</u>. Chomsky's (2000) early phase approach went back to the bounding approach in defining the trouble-makers, i.e. phases, rigidly, CP and vP being phases regardless of their structural position. This was soon followed by various contextual approaches, where whether XP is a phase depends on its syntactic context (on a par with *Barriers,* in contrast to the bounding/early phase approach), see Bošković (2014) and references therein. Thus, in Bošković (2012, 2013, 2014), there is a phasal domain, with the highest phrase in the domain functioning as a phase; e.g. DP is a phase in the nominal domain when present; when not, a lower projection is a phase.[1]

---

[1]These works show this approach can capture crosslinguistic variation regarding extraction without parameterizing what counts as a phase, which would involve parameterizing the computational system itself. E.g, while Abels (2003) argues for parameterizing PP-phasehood to capture variation regarding P-stranding within a rigid approach to phases, Bošković (2014) shows this can be done without parametric variation regarding phasehood under the contextual approach to phases, whose structural sensitivity gives it the flexibility needed to capture such crosslinguistic variation.

Furthermore, Bošković (2016b) argues that just like the phase status of a phrase is affected by its syntactic context, *phasal edges*, i.e. the status of a Spec regarding the PIC, is affected by the syntactic context in which it occurs: the highest phrase in a phasal domain functions as a phase, and the highest edge in multiple-edge contexts (α in (1)) functions as the phasal edge.

(1) [$_{XP}$ α [$_{XP}$ β…XP=phase

There has thus been a consistent move toward contextuality in the locality of movement. The scope of the contextuality of syntax is even broader. Chomsky (2013) argues that labeling is also contextual: the same element behaves differently for labeling depending on the context. A phrase behaves differently in phrase-phrase and head-phrase mergers. It also behaves differently in different phrase-phrase contexts, and its labeling status changes during the derivation: α in (2) changes from an unlabelled object (before ~~what'~~ moves away) to a CP (after the movement, see below).

(2) What do you think [$_α$ ~~what'~~ Mary bought ~~what~~]?

Bare-Phrase structure is also contextual: whether α is a head, phrase, or an intermediate projection depends on its syntactic context—its status also changes derivationally: what is a maximal projection after a head and a phrase merge becomes an intermediate projection with another merger. Furthermore, Takita et al (2016) argue that spell-out of a phasal complement removes it from the derivation, which turns the Spec of phase head into a complement.

Note here a parallelism with locality of movement: In the bounding/rigid-phasehood approach, one can look at a node, without paying attention to anything around it, and determine whether it's a bounding node/phase. This is not possible in *Barriers*/contextual-phasehood approaches. In GB structure one only needs to look at a node to determine its phrase-structure status, whether it's a phrase or a head (3). In Bare-Phrase structure, looking at any *like* in (4) doesn't help in determining whether it's a head or a phrase—this is determined contextually (what doesn't project (like$_1$) is a phrase, what's not a projection (like$_3$) is a head).

(3)

(4)
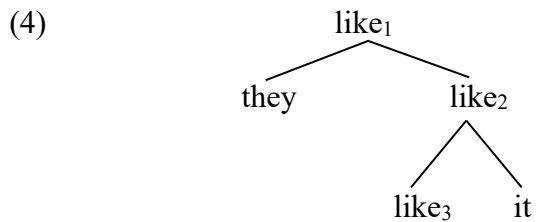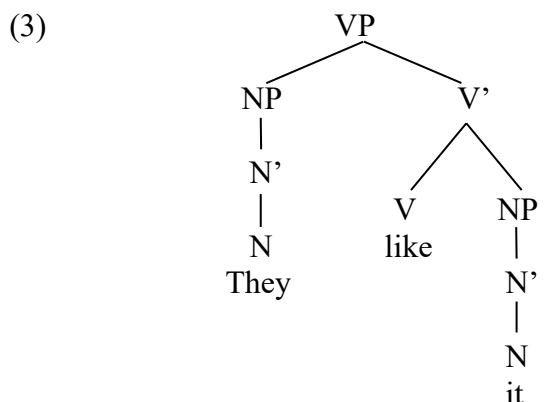
The A/A'-distinction is also now contextual. To illustrate, movement from vP must proceed via SpecvP. The A/A'-status of a SpecvP depends on the nature of that movement: if it's A-movement (the position below/above SpecvP in the chain is an A-position), SpecvP is an A-position (also if it's the landing site of

object shift); if it's A'-movement (as with long-distance-object/adjunct-movement), SpecvP is an A'-position (5); we need to look at the larger syntactic context to determine the A/A'-status of a SpecvP.

(5) How$_i$/Who$_i$ do you [$_{vP}$ t$_i$ think [Mary left t$_i$]]?

Furthermore, Bošković (2015, 2016a, 2018, 2020) provides a uniform account of all island/locality-of-movement effects based on a contextual approach to phases and the labeling theory, which is also heavily contextual, where there are actually no islands as this notion has been traditionally understood—there are no phrases that by their nature, independently of their syntactic context, disallow extraction, extraction being possible from all islands in well-defined contexts.

3. Spell-out

Evidence from numerous phonological/PF phenomena shows that what's sent to spell-out is phases, not phasal complements (see Bošković 2016a), which is actually what is expected since only phases have a theoretical status, phasal complements don't. This, however, raises an issue for successive-cyclic movement (SCM). If what is sent to spell-out is phases, and SCM targets phasal edges, as standardly assumed, an element undergoing movement would get frozen since it would be part of what's spelled-out hence inaccessible. This section shows that the conundrum can be resolved under the labeling theory: spell-out can correspond to a phase and SCM can still target phasal edges. Crucially, spell-out is then contextual in that a phase with a Spec does not always behave in the same way regarding spell-out.

Under standard assumptions, CP is a phase in (6), and the spell-out domain is IP (its complement). Under phasal complement spell-out, Y can interact with X but not Z in PF, while under phasal spell-out, Y can interact with Z but not X.

(6) X [$_{CP}$ Y [$_{IP}$ Z

Bošković (2016a) provides a number of arguments that Y can interact with Z but not X. Consider Bulgarian/Macedonian cliticization:

(7) a. Vera *mi   go*   dade.   Bg/Mac:OK
       Vera me.dat it.acc gave
       'Vera gave it to me.'
    b. *Mi go* dade.   Bg:*/Mac:OK
    c. Dade *mi go*.   Bg:OK/Mac:*

Bulgarian/Macedonian clitics are verbal clitics. Bulgarian clitics encliticize, Macedonian clitics procliticize. Bošković (2001) gives a copy-and-delete account of (7) based on Franks' (1998) (see also Bošković 2001) proposal that a lower copy can be pronounced iff this is necessary to avoid a PF violation. The clitic moves "around" the verb. In Macedonian, the higher clitic copy can always be pronounced, hence it must be. In Bulgarian, it can be pronounced when something precedes it, but not when nothing precedes it (8b). Here, a lower copy is pronounced since higher-copy pronunciation would induce a PF violation. The analysis captures the last resort nature of the V-Cl order.

(8) a. X cl V ~~cl~~   b. ~~cl~~ V cl   (Bg)
    c. (X) cl V ~~cl~~   (Mac)

Note that coordination can support Bulgarian enclitics, hence the impossibility of V-Cl below.

(9) a. I *ti   go*   dade.
       and you.dat it.acc gave
    b. *I dade *ti go*.

3

Consider now wh-C/*li* questions. They involve V-to-C movement, V carrying the clitics along, as in Macedonian (10). Surprisingly, in Bulgarian (11b-c), the conjunction cannot support the enclitics. Instead of expected (11b), we get (11a). Why does lower-copy pronunciation apply in (11a), in contrast to (9a)?

(10) I   *ti*   *go*   dade *li*?
     and you.dat it.acc gave Q
     'And did (s)he give it to you?'

(11) a. I   dade *li* *ti*   *go*?
      and gave Q you.dat it.acc
    b. *I *ti go* dade *li*?

Franks and Bošković (2001) provide a multiple spell-out account, where what's sent to spell-out is crucially the phase. The Cl+V-cluster head-moves to *li*. CP is then sent to spell-out, with lower enclitic copies pronounced since there is nothing in (12) that can support them.

(12) [CP[C ~~ti go~~ dade+li] ti go ~~dade~~]

The conjunction is only then added, deriving (11a)/(13):

(13) i [CP[C ~~ti go~~ dade+li] ti go ~~dade~~]

Now, if phasal complements were sent to spell-out, C would be in the same spell-out domain as the conjunction. This should then lead to higher copy pronunciation even in Bulgarian (14). This apparently doesn't happen, indicating the CP phase is sent to spell-out.

(14) i [CP[[ti go dade] li] ~~mi go dade~~]

Bošković (2016a) gives a number of other arguments for the impossibility of interaction between the edge of phase XP and material outside XP, which favor phase over phasal complement spell-out, as well as for the possibility of PF interaction between a phase head and its complement, which is disallowed under phasal complement spell-out since they don't belong to the same spell-out domain. Thus, D'Alessandro and Scheer (2015), Bošković (2016a) show that raddoppiamento fonosintattico (RF) in Abruzzese, where a word's initial consonant undergoes gemination conditioned by the preceding word in a specific syntactic relationship, is spell-out domain sensitive based on active/passive contrasts. Importantly, in (15)-(16) the complementizer and the pronoun in SpecCP trigger RF within IP, which shows that phasal complements are not spell-out domains. (RF is not possible across CP.)

(15)   Jè mmeje chə vve.
      is   better that come.3SG
      'It's better that he/she comes.'

(16)   lu waglionə   chə   ttu si vistə
      the boy   whom you are seen   (D'Alessandro and Scheer 2015)

PF interaction is thus impossible between the edge of phase XP and material outside XP but possible between the edge of XP and the complement of X. The conclusion is then that spell-out domains correspond to phases. (Pre-phasal approaches to prosodic phrasing (Nespor and Vogel 1986, Selkirk 1986) anticipated phasal spell-out since there CPs correspond to intonational-phrases—the correspondence was with a phase, not phasal complement). But what about successive-cyclic movement (SCM)? If phases are spell-out units and SCM targets phasal edges, the element undergoing SCM would get frozen: it would be part of a spelled-

out unit, hence inaccessible. The issue can, however, be resolved under the labeling theory, whose contextuality was noted above.

Chomsky (2013) argues that when a head and a phrase merge, the head projects, labeling the resulting object. When two phrases merge, there are two ways to implement projection/labeling: feature-sharing or traces/movement, traces being ignored for labeling. (17) illustrates the former: when *what* merges with CP, both have the Q-feature, which determines the resulting object's label.

(17) I wonder [$_{CP}$ what$_i$ [$_{C'}$ C [John bought t$_i$]]]

Consider now SCM in (18). SCM doesn't involve feature-sharing (which follows Bošković 1997, 2007, 2008). Since there is no feature-sharing between *that* and *what*, the embedded clause can't be labeled when *what* moves to its edge (19). After v merges, *what* moves. The element merged with *that*-CP being a trace, it's ignored for labeling, and ? is labeled as CP.

(18) What$_i$ do you think [$_{CP}$ t'$_i$ [$_{C'}$ that [John bought t$_i$]]]
(19) v [$_{VP}$ think [$_?$ what [$_{CP}$ that [John bought t$_i$]]]]

Above, I noted Bošković's (2014) contextual approach to phases, where there are phasal domains and the highest phrase in the domain is a phase. *Phase* is determined when a phasal domain is closed. This is in turn determined when the element with the relevant phasal property (X) no longer projects, which closes the phasal domain—X determines a phase when X no longer projects. WP in (20) then becomes a phase only after merging with K, with K projecting (i.e. with WP no longer projecting).

(20)

```
            KP
           /  \
          K    WP
              (X)
```
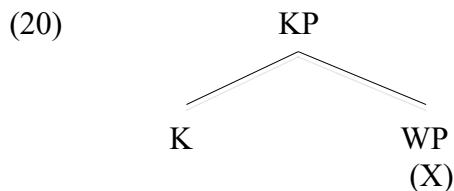
So, the phasal status of WP is determined only after it merges with K.

Consider again SCM:

(21) What do you think that John bought?
(22) [$_?$ what **[$_{CP}$ that [$_{IP}$ John bought]]**]

In (22), C and IP merge, with the resulting object labeled CP. *What* then merges with the CP. The resulting object is unlabeled. Since there was a further merger and CP didn't project, the phasal domain is closed. As the highest projection in the phasal domain, the CP sister of *what* is a phase and sent to spell-out, given that what is sent to spell-out is phases. (Phases/spell-out domains are bolded in (22)/(24)). Once *what* moves away (after merger of additional structure), traces being ignored for labeling, ? is labeled as CP.

We can then have spell-out correspond to a phase and SCM target phases/phasal edges. *What* targets CP (*that*+IP), and this CP is sent to spell-out. *What* is outside the spell-out domain, but not because of the PIC (which is not needed), but because it's not part of the phase that is sent to spell-out.

Consider now agreeing Specs.

(23) Mary wonders what John bought
(24) **[$_{CP}$ what [$_{CP}$ +wh-C [$_{TP}$ John bought]]**]

Here, there is feature-sharing when *what* and CP merge. The resulting object is determined to be a phase after merger with *wonder*, and sent to spell-out.

Traditionally, *what* in (21) and (23) is in the same position, and at the phase edge (i.e. outside the spell-out domain), but in the above analysis these constructions differ regarding whether *what* is in the spell-out domain. What's particularly important is that phase XP with YP in its Spec does not always behave in the same way for spell-out, which indicates contextuality of spell-out: With SCM, the sister of YP is sent to spell-out, otherwise the full XP is sent to spell-out, although both involve phasal spell-out.

4. The EPP

This section broadens the scope of contextuality by including the EPP. What will be relevant is the approach to the EPP from Bošković (2024), where there is an EPP domain, with the EPP satisfied in the highest projection in the domain. This approach gains significance within the broader picture discussed above: it shows broader relevance of contextuality, contextuality also being relevant for the EPP, in fact in the same way as for phases and phasal edges—there is a domain for phases/phasal edges/EPP, the highest phrase in the domain functioning as a phase/phasal edge/locus of the EPP. The approach is based on a puzzle raised by subject wh-movement, which is discussed next.

4.1. *Who left*

There is a debate regarding whether *who* in (25) is in SpecCP, moving via SpecIP (26), or in SpecIP (27).

(25) Who left?
(26) [CP Who$_i$ [IP t$_i$ [vP t$_i$ left]]]?
(27) [IP Who$_i$ [vP t$_i$ left]]]?

There is strong evidence against both. Evidence for wh-movement in (25) argues against (27) (for additional evidence, see Bošković 2024, Messick 2020).[2] Thus, (28) shows *hell*-phrases modify only wh-moved DPs. (29) shows they can modify subject wh-phrases (Pesetsky 1987), indicating there is wh-movement here.

(28) a. *Who left who the hell?   b. Who the hell did May leave?
(29) Who the hell left?

Inverse scope in (30a) shows that object *everyone* can scope over a quantifier in SpecIP. It is, however, impossible in (30b), which indicates that *who* is not in SpecIP in (30b) (Mizuguchi 2014).

(30)   a. Someone loves everyone.   (someone>everyone/everyone>someone)
       b. Who loves everyone?   (who>everyone/*everyone>who)

There is also strong evidence that (25) doesn't involve wh-movement through SpecIP (i.e. against (26)). In many languages with both the SV and VS order, where S doesn't move to SpecIP in VS, the orders have different morphology. Importantly, *who left* has the VS morphology (e.g. Fiorentino). This shows subject wh-movement does not proceed via SpecIP or we would get the S-V morphology. This is confirmed by languages where the agreement associated with subjects in SpecIP is dropped in *who left*. Thus, in Kinande, instead of usual agreement (31a), we get anti-agreement in subject questions (31b).

(31) a. Kambale **a.**langira Marya.
        Kambale agr.saw   Mary
     b. Iyondi yo ***a.**langira/**u.**langira   Marya.
        who   C   agr.saw/anti-agr.saw Mary   (Schneider-Zioga 1995)

---

[2]No *do*-support doesn't mean no inversion in (25). *Do*-support is a last-resort operation to support a T-affix when a phonologically realized element intervenes between it and the V (Chomsky 1957). There is no such intervener in *who walked*, as in *she walked* and in contrast to *what did she buy*. Only the last case then triggers *do*-support.

Consider Icelandic (32). While experiencers block agreement with a lower nominative (32a), an NP-trace doesn't ((32b): *mér* is a quirky subject moving to SpecIP). However, a wh-trace blocks it (32c). If the experiencer in (32c) were to move to SpecIP before wh-movement, the intervener would be an NP-trace. (32c) would then incorrectly pattern with (32b). The wh-phrase apparently doesn't move to SpecIP in (32c)—it moves directly to SpecCP.

(32) a. Það   virðist/*<u>virðast</u> **einhverjum manni** [hestarnir   vera seinir]
EXPL seems/seem$_{PL}$ some   man$_{DAT}$ the-horses$_{NOM}$ be slow
'It seems to some man that the horses are slow.'
   b. **Mér**   <u>virðast</u> **t**$_{NP}$ [hestarnir   vera seinir]
me$_{DAT}$ seem$_{PL}$ the-horses$_{NOM}$ be slow   (Holmberg and Hróarsdóttir 2004)
   c. **Hverjum** mundi/?*<u>mundu</u>   hafa virst   **t**$_{wh}$ [hestarnir   vera seinir]
who$_{DAT}$ would$_{3SG}$/would$_{3PL}$ have seemed   the-horses$_{NOM}$ be   slow
'To whom would it have seemed the horses are slow?'   (Nomura 2005)

Conclusive evidence against (26)-(27) is provided by West Ulster English (WUE). Unlike standard English, WUE allows Q-float under wh-movement (*what$_i$ did he say all t$_i$ that he wanted?*). Q-float is also possible in (33a). However, like standard English, WUE disallows (33b). (33b) shows a subject in SpecIP cannot float a quantifier in the postverbal position in passives. This rules out the derivation where *who* in (33a) moves to SpecCP via SpecIP since *all* would then float under movement to SpecIP, which (33b) shows is disallowed. (33) in fact rules out both (26) and (27).

(33) a. Who was arrested all in Duke Street?
    b. *They were arrested all last night.   (McCloskey 2000)

Who apparently does not pass through SpecIP in *Who left*. How is the EPP satisfied then? Bošković (2024) addresses this issue by arguing that there are two wh-positions, a higher and a lower one, where the lower wh-position is occupied by wh-moved subjects. This means *who* in *I wonder **who left*** is not as high as *who* in *I wonder **who Amy met*** but higher than *John* in *I think **Amy left.***

4.2. Two wh-positions[3]
Consider auxiliary-contraction. Kaisse (1983) shows that there is a one-word-host restriction on contracted auxiliaries hosted by moved wh-phrases (34)-(35). Crucially, the restriction is not found with subject wh-phrases (36). This indicates the wh-phrases/auxiliaries are not in the same position in (35) and (36). (Recall the subject wh-phrase does wh-move.)

(34) a. What's Mary buying?   b. When's dinner?
(35) *Which man's she the fondest of?
(36) Which man's leaving first?

Further, (37) shows only the landing site of non-subject wh-movement is higher than topics, which means wh-moved subjects are lower than wh-moved objects.

(37) a. ?Mary wonders which book, for Kim, Peter should buy.
    b. *Mary wonders which student, for Kim, should buy that book.

Polarity adverbs also indicate different landing sites for non-subject and subject wh-movement, since the

---

[3]Below I only outline some arguments from Bošković (2024), who gives more comprehensive accounts.

two behave differently in this case too.

(38) a. *What under no circumstances should Mary ever buy?
    b. Who under no circumstances should ever hire Peter?
(39) a. *What should under no circumstances Mary ever buy?
    b. Who should under no circumstances ever hire Peter?

Importantly, long-distance moved subjects pattern with objects: aux-contraction in (40) patterns with (35), not (36), which means only <u>local</u> subject wh-movement goes to the lower position.

(40) *Which man's Peter claiming will leave first?

The same obtains with topicalization, (42) being better than (41), and with polarity adverbs, (43)-(44) patterning with object (38a)/(39a) not local-subject extraction (38b)/(39b).

(41) *Mary wonders which student, for Kim, should buy that book.
(42) ?I wonder which student, for Kim, Mary said should buy that book.
(43) *Who under no circumstances should Ann ever say stole it?
(44) *What should under no circumstances Mary ever believe destroyed it?

Furthermore, in many languages C is different in local subject and all other questions.[4] In Norwegian, the relevant head in subject wh-questions is *som*, but it's null in object/adjunct/long-distance subject questions. The obvious conclusion is that the wh-phrases in (45) and (46) don't land in the Spec of the same head.

(45) Vi vet   hvem *(som) snakker med Marit.
     we know who   C   talks   with Marit
(46) a. Vi vet   hvem (*som) Marit snakker med.
        we know who   C   Marit talks   with
     c. Vi vet   hvem (*som) Pål   tror   (at) har kommet.
        we know who   C   Paul thinks that has come.   (Taraldsen 1986)

In Defaka, the relevant head is always phonologically realized, but differently with subject and non-subject extraction. Defaka has focus-movement. (47a) doesn't involve focus-movement. (47b) involves local subject-focalization, (47c) object-focalization. Importantly, (47b-c) have different focus markers, with another difference in V-morphology: there is a morphological reflex of focalization in both cases ((47b-c) differ from (47a)), but the reflex is different. Furthermore, adjunct/long-distance-subject fronting patterns with object-fronting.

(47) a. ì Bòmá ésé-kà-rè.
        I Boma see-fut-neg
      'I will not see Boma.'
     b. ì **kò**   Bòmá ésé-kà-rè.
        I foc   Boma see-fut-neg
      '<u>I</u> will not see Boma.'
     c. Bòmá **ndò** ì ésé-kà-rè-**kè**.
        Boma foc   I see-fut-neg-ke
      'I will not see <u>Boma</u>.'   (Bennett et al 2012)

---

[4] For a different perspective on some such cases, see Pesetsky (2021).

Bùlì wh-movement shows the same pattern, with adverbs showing that short-distance wh-moved subjects/particles are lower than other wh-movements/particles (Sulemana 2019, Pesetsky 2021).

There are also many languages with processes restricted to subject A'-extraction. Dias (2022) shows Brazilian Portuguese *nunca*-ellipsis is possible under subject wh-movement, but not object/adjunct/long-distance-subject wh-movement.

(48) Maria beijou João.   a. Quem nunca [beijou João]
     'Mary kissed John'   who   never kissed John
                                     'Who has never kissed John?'
                                  b: *Quem Maria nunca [beijou t] (cf. Quem Maria nunca beijou?)
                                      who   Mary never   kissed

Hong Kong Sign language (HKSL) exhibits the same exception. Like many sign languages, HKSL has rightward wh-movement—the contrast between (50) and (51)-(52) shows HKSL is not a wh-in-situ language. These examples involve object wh-movement. Long-distance subject wh-movement patterns with objects (53)-(54). Local subject wh-movement is, however, different (55): only locally-moved subjects (see Gan 2022 for evidence they do move) need not undergo rightward wh-movement.

(49) a. AARON LIKE WHAT/WHO?   b. *WHAT/WHO AARON LIKE?
(50) AARON EAT BREAKFAST SHORT-TIME.
     'Aaron ate the breakfast quickly.'
(51) *AARON EAT WHAT SHORT-TIME.
(52) AARON EAT SHORT-TIME WHAT?
(53) KENNY THINK HELP-a AARONa WHO?
     'Who does Kenny think helped Aaron?'
(54) *WHO KENNY THINK HELP-a AARONa?
(55) WHO EAT FISH NOT?   (Gan 2002)

Lewis (2022) shows only short subject wh-extraction is allowed from British English *do*-ellipsis, object/long-distance-subject/adjunct-extraction being disallowed, which is impossible to capture if they all land in the same position (recall there is wh-movement in the former).

(56) I don't know who Martha thinks will leave, but I do know who Emily thinks will (***do**).
(57) Sue wouldn't kiss Peter last night. Well, who WOULD **(do)**?

Branan and Erlewine (2024) show that many languages have local-subject specific relativization strategies. They also show that relativization strategies not involving local subjects land in a different position.

All this indicates that there are two wh-positions, a higher and a lower one, where the lower wh-position is occupied by locally moved wh-subjects. Why is it confined to subjects? Bošković (2024) argues that it is a mixed A/A'-position on the border of A and A'-fields: it's the landing site of wh-movement (creating an Op-variable relationship), but also the position where the EPP is satisfied.

This straightforwardly resolves the puzzle of *who left,* where the EPP appears to be mysteriously voided. The EPP is satisfied: it's satisfied in the lower wh-position, a mixed A/A'-position confined to locally A'-moved subjects. This captures all the patterns discussed above. Norwegian *som,* found only with local-subject wh-movement, is a PF realization of the A/A'-head, subject wh-movement stopping in its Spec. Defaka/Bùlì also have PF realizations of the A/A'-head distinct from C. Local-subject-wh-movement-restricted processes also receive structural accounts: wh-subjects in *nunca*-ellipsis, HKSL questions, *do*-ellipsis are in SpecA/A'P; Kinande anti-agreement occurs with SpecA/A'P, regular agreement with SpecIP.

Branan and Erlewine (2024) analyze subject-restricted relativization as A'-probing for the closest DP, noting this is problematic since it amounts to A'-extraction with the locality of A-movement. Probing for the closest DP is associated with the EPP (topicalization/wh-movement also affect non-DPs and not just the closest DP/non-DP). Under the above analysis, this actually is EPP-probing. It also doesn't land in SpecCP.

Importantly, this leads to a conception of the EPP on a par with the contextual approach to phases in Bošković (2014), where there are phasal domains and the highest phrase in a phasal domain is a phase: there is an EPP domain with the highest projection in this domain being the locus of the EPP.

(58) EPP domain: [$_{CP}$ who-acc [$_{A/A'P}$ *who-nom* [$_{IP}$ *Amy-nom*

Focalized subjects also move to SpecA/A'P, unsurprisingly since wh/focalized elements often pattern together. Consider (59).

(59) a. Only his girlfriend does John give any flowers.
      b. *John gives only his girlfriend any flowers.
      c. Only Mary showed any respect for the visitors.   (Branigan 1992: 84)

The *only*-licensor c-commands the NPI in (59a-b), which apparently cannot be licensed from a purely A-position, hence the ungrammaticality of (59b). (59c) can then captured if the focalized subject moves to the mixed A/A'-position like *who* in (25). Under this analysis, *only Mary* in (59c) is not in SpecIP, hence it can license the NPI, but it is also not in SpecCP. Consequently, it allows inversion.

(60) Did only Mary show any respect for the visitors?

A scope puzzle regarding negative quantifiers also gets resolved. Recall that in contrast to (30b), *everyone* can take wide scope in (30a), indicating the subject in (30b) is higher than (30a). A long-standing puzzle has been that inverse scope is impossible in (61).

(61) Nobody likes everyone.

In many languages such negative elements must undergo focus-movement (see Bošković 2009). The suggestion is then that *nobody* moves to SpecA/A'P in (61), hence the lack of inverse scope.

Consider also overt imperative subjects. There is a controversy regarding their position. Object-drop in imperatives can be used to diagnose it. It's possible with null but not overt subjects (Sadock 1974, Sigurðsson and Maling 2008, Bošković 2011).

(62) a. Open carefully!
      b. ?*You open carefully!
      c. You open it carefully!

Bošković (2023) argues this null object moves to SpecCP (for licensing), evidence for which is provided by parasitic-gap licensing, which requires movement.

(63) Open without closing afterward

Island-sensitivity confirms the movement.

(64) Ask how you can open it with a knife
(65) *Ask how you can open with a knife

Crucially, overt imperative subjects are focalized (contrastively focused). Being focalized, they are located in SpecA/A'P hence block null object movement.

So far we have (58). Bošković (2020) argues for returning to split IP. Consider e.g. coordination. Given that bar-level coordination is disallowed, (66), where the subject is outside and the modal inside the coordination, shows that the subject and the modal are not in the same phrase.

(66) John [travels to Rome tomorrow] and [will fly for Paris on Sunday].

This can be captured in early minimalist structure, which split IP into AgrsP and TP. From this perspective, Bošković (2024) argues that agreeing subjects are in SpecAgrsP, and non-agreeing/quirky subjects (32b) in SpecTP.

There is an ill-understood variation regarding quirky subjects. Taking binding subject-oriented anaphors as a diagnostic for true quirky subjects, Bošković (2024) establishes (67).

(67) Quirky subjects are allowed only in *pro*-drop languages.

Bošković (2024) provides a deduction of (67), the gist of which is that the EPP can be satisfied in SpecTP only in Chomsky's (2015) strong-T languages, which are *pro*-drop languages. Quirky subjects cannot go to SpecAgrsP, which is reserved for agreeing nominative subjects, so they can satisfy the EPP only in pro-drop languages. We then have (68).

(68) EPP domain: [$_{A/A'P}$ wh/A'-moved subjects [$_{AgrsP}$ Amy [$_{TP}$ quirky subjects

What's particularly important for us is that the system accommodates mixed behavior of non-nominal subjects. Consider Locative Inversion (LI). It shows subject properties, e.g. subject-raising (69) and the lack of WCO (70).

(69) On the wall$_i$ seemed [t$_i$ to be hanging some picture].
(70) a. Into every flat$_i$ peered its$_i$ owner.
    b. *Into every flat$_i$ its$_i$ owner peered.

LI subjects also show non-subject properties: they block extraction and disallow inversion.

(71) ?*Who do you think that on this wall hung [a picture of t]?
(72) *Did on the wall hang pictures?

They also disallow inverse scope, like A'-moved subjects and in contrast to regular subjects (30).

(73) On some stage stood every actress.　　　　　∃>∀/*∀>∃

All this follows if they move to a higher subject position than regular subjects, namely SpecA/A'P.[5]

There's a debate regarding whether LIs are subjects or topics due to conflicting data (e.g. (69) vs (71)-(73)). Under the A/A'P-analysis, the conflict is reconciled, without requiring positing a null expletive in non-pro-drop languages like English (which topic analyses need given the impossibility of subject topicalization in English (Lasnik and Saito 1992)).

---

[5]Regarding inversion, the intervening A/A'-head blocks Agrs+T from undergoing it. Inversion targets the higher wh-head, C. (It targets the head with a +wh-feature; in subject questions, this is the A/A'-head—there is no separate +wh-C). Inversion is not blocked with focalized elements (*only*-DPs), which is related to the wh-feature. Under Roberts' (2010) theory of head-movement, through C-Agrs/T association (Chomsky 2008) C gives the relevant feature to Agrs/T, which triggers head-movement. With LI, the intervening A/A'-head disrupts the association. This doesn't happen if the head is +focus given the wh/focus relation (cf. Multiple Agree, where X can Agree with Y across Z if it agrees with Z).

LI can be taken to indicate that non-nominal subjects move to SpecA/A'P, not SpecAgrsP/SpecTP. Clausal subjects provide evidence that this is indeed the case. They also show mixed subject properties, with a similar debate based on conflicting data as with LI ((74)-(75), e.g. Stowell 1981 treats them as topics, Bošković 1995 as subjects). The conflicting data can be reconciled with a SpecA/A'P-subject treatment.

(74) [That Sue resigned] seems likely.
(75) *Is [that Sue resigned] likely?

LI and CP subjects thus indicate that non-nominal subjects move to SpecA/A'P. Interestingly, it has been argued that there is a QP above wh-DPs (Hagstrom 1998, Cable 2010), which essentially makes them non-nominal. Below I'll argue there is a similar projection above all phrases undergoing A'-movement, the right generalization being that non-DP/nominal subjects cannot move to SpecAgrsP. The following then holds:

(76) Non-nominal subjects satisfy the EPP in SpecA/A'P.
(77) Agreeing nominative subjects satisfy the EPP in SpecAgrsP.
(78) Non-agreeing/quirky nominal subjects satisfy the EPP in SpecTP.

As noted above, there has been a consistent move toward contextuality of syntax in many domains, including structure-building, labeling, the A/A'-distinction, locality (islandhood as well as phasehood/phasal edges), and spell-out domains. The EPP is now also added to this picture. In fact, the contextuality of the EPP is the same as that of phases/phasal edges, being defined in the-highest-phrase-in-the-domain terms.

4.3. Markovian or not?

There is a larger point here: The EPP and other contextually-determined mechanisms cannot be considered simple Markovian properties. This actually confirms Chomsky's (1957) conclusion that finite-state devices/Markov processes are insufficient for language.[6] They are characterized by: initial and final state, specifications of transitions from one state to another and a symbol to print under a transition. (79) gives a finite-state-device grammar generating *the man comes* and *the men come*.

(79)
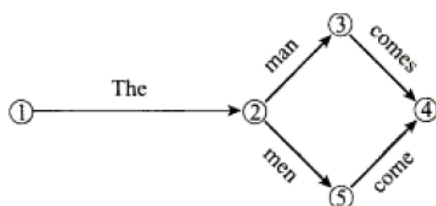
The transition from 1 to 2 prints *the*, from 2 to 3 *man*… The choice of a word restricts the next choice, so we cannot get **The man come*. The hypothesis is that a speaker is a finite-state machine. They can capture infinity. Consider (80): there is a loop at state 4, which yields *the man runs*, *the man runs and runs...*

(80)
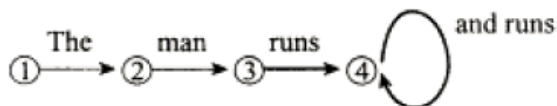
Once it's in a state, the system knows that state, where it can go, and what symbol to print. It knows nothing about where it was—it has no memory; as a result, it cannot capture dependencies that need memory. This is problematic. E.g., Lasnik et al (2000) note (82), which derives missiles, anti-missile missiles, anti-anti-missile missile missiles… The best a finite-state device can do is (81), which doesn't give (82).

---

[6]They are used quite widely, in information theory, thermodynamics, statistical mechanics...

(81)
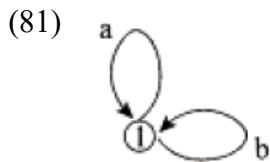
(82) anti^n missile^{n+1}

Non-adjacent/long-distance dependencies, which are quite common (e.g. *the man who fired them is leaving*), are problematic for the same reason. So, a finite-state device cannot capture the nature of language. But is there anything in language that looks like it? There is: compositional semantics. Its tenet is that when computing the meaning we get by combining A and B, we don't know how we got to A or B—we only know A and B and what we get by combining them. In this respect, compositional semantics has no memory, and no look-ahead power—it's a finite-state device. Many PF processes, like affix-hopping, require adjacency, which is Markovian by nature.

A Markovian device is, however, insufficient for syntax. Probably the most important things the theory of syntax needs to capture are infinity and non-adjacent/long-distance relations. A Markovian device can capture the former, but not the latter. Returning to contextuality, contextuality is by its nature non-Markovian. E.g. to determine the phrase-structure status of like$_2$ in (4) one needs to look at all of (4); to determine whether α is a phasal edge in (1), one needs to know that there are no other XP-Specs and that XP is a phase, which is determined only when X merges with another element, with that element projecting. To determine the labeling status of α in (2) one needs to know that ~~what²~~ is a lower copy of a chain. These cannot be done in a Markovian fashion. The above discussion of contextuality thus confirms the non-Markovian nature of syntax, considerably broadening its scope.

One property of Markovian processes has, however, been endorsed in minimalism: no look-ahead. Under the standardly-assumed *local economy* (Collins 1996), when evaluating the best course of action, all we know is where we are and what we can do <u>next</u>, we cannot look-ahead to the steps after the next step, which are in fact Markovian properties.

A question arises, do the contextual processes discussed above require look-ahead? While at first sight it appears they do, this is actually not the case. Bošković (2014) demonstrates this regarding the contextual approach to phases discussed above. Consider now the contextual EPP. Assume heads like Agrs and A/A' have an unvalued EPP-feature and elements that satisfy the EPP have a valued EPP-feature. Agrs merges into the structure, looking for a nominal to attract (only nominal elements undergo internal merge with it). If there is a nominal (*Mary* in (83)) to attract, Agrs attracts it— this merger satisfies the EPP (Agree alone cannot satisfy the EPP, which requires movement—this is the strength property of early minimalism). If there isn't, as in (84) (recall *who* is not a nominal), internal Merge cannot take place. The A/A'-head merges into the structure.

(83) Agrs [$_{vP}$[$_{DP}$ Mary]…]
(84) a. Agrs [$_{vP}$[$_{QP}$ who]…
    b. A/A' [$_{AgrsP}$ Agrs[$_{vP}$[$_{QP}$ who]…]

The A/A'-head and Agrs undergo feature-sharing as in Frampton and Gutman (2002), where two instances of an unvalued feature become one—when one of them gets valued they both do. The A/A'-head can attract *who*; *who* then merges with it, valuing its EPP-feature, which automatically takes care of the EPP-feature of Agrs. Note there is no look-ahead. Furthermore, the propagation/percolation of the EPP requirement is captured without real percolation.

5. Adieu to the A/A'-distinction: Binding

The contextual EPP provides a natural mechanism for eliminating the A/A'-distinction from the binding theory. The A/A'-distinction is a Government-and-Binding remnant. It is in fact quite clear that conceptually, it's as bad as the notion of government, which was abandoned long ago. In the spirit of Chomsky's (2021) duality of semantics, the A/A'-distinction would be natural if it were defined in terms of theta-roles, corresponding to the theta/non-theta-distinction. This is, however, not the case. Case-positions (actually potential Case-positions) are also relevant: A-positions are Case and theta-positions. However, involving Case/theta-assignment is no longer enough to define A-positions. Thus, SpecvP is assumed to be an A-position: the subject is base-generated and object shift lands there. But what about (85), where *how* moves through SpecvP, vP being a phase? SpecvP should be an A'-position here (although SpecvP is a potential Case/theta-position).

(85) How$_i$ do you [$_{vP}$ t$_i$ think [Mary left t$_i$]]?

As noted above, the A/A'-status of a SpecvP depends on the nature of movement stopping there; if it's A-movement (the position below/above SpecvP in the chain is an A-position), SpecvP counts as an A-position; if it's A'-movement (as with adjunct/long-distance-object wh-movement), SpecvP counts as an A'-position; we need to look at the larger syntactic context to determine its A/A'-status.

Interestingly, this means Improper Movement is no longer statable. In the Improper Movement configuration below, t' is now in an A-position, which means there is no A-A'-A movement here.

(86) *Mary$_i$ is believed [$_{CP}$ t'$_i$ that John likes t$_i$].

Case/theta-roles are actually anyway not enough to define A-positions. Consider Japanese scrambling. (87) violates Condition A, with clause-mate scrambling in (88) taken to involve A-movement since Condition A is satisfied. Long-distance scrambling in (89) is taken to involve A'-movement since Condition A is still violated here. However, long-distance scrambling in (90) is taken to involve A-movement since Condition A is satisfied.

(87) *[Otagai-no   hahaoya]-ga [Mary-to Pam]-ni   atta.
    each.other-GEN mother-NOM   Mary and Pam-DAT met
    'Each other's mothers met Mary and Pam.'
(88) [Mary-to Pam]-ni$_i$ [otagai-no hahaoya]-ga t$_i$ atta.
    'Mary and Pam, each other's mothers met.'
(89) *[Mary-to Pam]-ni$_i$ [otagai-no   hahaoya]-ga [John-ga   t$_i$ atta to] omotteiru.
    Mary and Pam-DAT each.other-GEN mother-NOM John-NOM   met C think
    'Mary and Pam, each other's mothers think that John met.'
(90) John-to Bob$_i$-o   otagai$_i$-no   titioya$_j$-ga [$_{CP}$ PRO$_j$ t$_i$ rikaisiyoo to] kokoromita.
    John and Bob-acc each.other's fathers-nom   understand C   attempted
    'John and Bob, each other's fathers attempted to understand.'

It's simply not possible to relate the needed A/A'-distinction for (87)-(90) to Case/theta-roles. There is no sense of Case/theta-roles where the landing site of scrambling involves Case/theta-positions in (88)/(90) but not (89). This is the tip of the iceberg, it's easy to come up with similar paradigms from other languages. There is really no definition of the A/A'-distinction that works in a principled way. Actually, there is: if movement results in a new binding configuration it's A-movement, if it doesn't, it's A'-movement. This is the only "principled" way the relevant cut can be made. Its circularity is quite obvious. (The same kind of

circularity was found with PRO, where the only "principled" criterion for what counts as a governor for PRO became this: if a PRO-sentence is bad, something close to PRO is a governor, if it's good, nothing is.) Note that the A/A'-distinction has been eliminated from relativized-minimality/intervention effects, which are now relativized with respect to the actual features involved. It would thus be extremely appealing if it can be shown that binding effects can be handled without the A/A'-distinction. This section will focus on binding, in a preliminary attempt to get binding condition effects strictly with c-command. The goal is the following: What we clearly need for binding is c-command, how about nothing else. This means that when a nominal is in a traditional A-position it can c-command, but when it's in a traditional A'-position the issue is the lack of c-command: the fronted nominals then do not c-command the anaphor in (91)-(92).[7]

(91) *Which students did each other's mothers meet?
(92) *John and Mary, each other's mothers met.

The basic idea is then rather straightforward: after undergoing traditional A'-movement, the nominal does not c-command hence cannot bind.

Regarding wh-phrases, it has been proposed that there is a QP, projected by a Q-morpheme, above wh-DPs undergoing wh-movement (e.g. Hagstrom 1998, Stepanov 2001, Bošković 2003, Cable 2010). The suggestion is then that there is a similar projection above all phrases undergoing traditional A'-movement (for relevant discussion of topics, see Yoo 2018; for different but relevant discussion see Rezac 2003, Safir 2019). This is in fact why the relevant elements do not move to SpecAgrsP (section 4)—they are not nominal. Importantly, the projections in question prevent the DP from c-commanding out—there is then no binding after traditional A'-movement.

The system discussed in section 4 provides evidence for the proposal that elements undergoing traditional A'-movement are not nominal. We have seen above that non-nominal subjects satisfy the EPP in SpecA/A'P, agreeing nominative subjects in SpecAgrsP, and non-agreeing nominal subjects in SpecTP. That traditional A'-subjects, like *who* in *who left*, pattern with non-nominal subjects in satisfying the EPP in SpecA/A'P then provides evidence that they are indeed not nominal.

So, FocP, TopP, QP, not DP, is what undergoes traditional A'-movement. This raises an intriguing possibility: there is no TopP, FocP in the left periphery (Chomsky 2008); the relevant movements are moving-element driven (see Bošković 2007)—if it's to be topicalized/focalized, then TopP/FocP is projected right above it (what was a feature of the moving element driving movement in Bošković 2007 is a projection now), which then forces movement (see below).

Now, lower down in the structure, the relevant elements do bind/c-command:

(93) John, Mary sold t pictures of himself.
(94) Who did Mary t sell pictures of himself?

Something then must happen during the derivation so that the relevant elements start as nominal but become non-nominal (see also Rezac 2003, Safir 2019). Something like this happening during the derivation is not unheard of—consider distributed coordination constructions (DCC) discussed in Bošković (2022). In (95), the individual conjuncts bear the external theta-roles but agreement is with the ConjP. Bošković (2022) (see also Zhang 2010) argues that the coordination in (95) is formed derivationally. The individual conjuncts are inserted in the respective theta-positions, then a ConjP is formed through sideward merger (Nunes 2004), and inserted into the structure above the theta-positions. (96) confirms the analysis. In (96a), the whole

---
[7]Safir (2019) is an important predecessor of the current work in this respect.

15

coordination binds the anaphor; in (96b) the individual conjuncts do. This indicates that when the anaphor is low, below the position where the relevant elements are base-generated, the conjuncts function as binders <u>separately</u>. When the anaphor is high, they bind <u>jointly</u>. Furthermore, (96c) shows the conjuncts must start separately—they induce a blocking effect separately (t_i and t_j are the closest subjects). The only way to account for (96) is to have the conjuncts start separately, with the ConjP built during the derivation.

(95) [The dog$_i$ and the rooster$_j$]$_k$ were [$_{vP}$ t_i barking] and [$_{vP}$ t_j crowing] yesterday.
(96) a. [Bill and Sue]$_i$ seem to each other$_i$ to be the best candidate in the election and the best nominee for the convention respectively.

    b. [Bill$_j$ and Sue$_k$]$_i$ hired himself$_j$ and nominated herself$_k$ respectively.

    c. *[Bill$_i$ and Sue$_j$]$_k$ seem to be t_i the best candidate in each other's$_k$ campaigns and t_j the best nominee in each other's$_k$ parties respectively.

This is in fact exactly what we need to achieve with derivational structure building in our cases. What we need must then be independently possible, given the behavior of DCCs.

    This, however, raises an issue regarding the cycle. Below I suggest several possibilities for implementing the required derivational structure building (X being the relevant element), which also address this issue (for a different proposal see Safir 2019):

    (a) We are dealing with sideward merger (with X), on a par with what happens with DCCs. Bošković (2022) argues that in DCCs like (95), after being introduced in their theta-positions, the conjuncts undergo sideward merger into ConjP, which is then inserted back into the structure. On a par with this, the wh-DP is base-generated in the theta-position, and undergoes sideward-merger with X, with XP reintroduced into the structure after this sideward merger. (Bošković 2022 argues that the DCC ConjP is not introduced in its final position and that the individual conjuncts undergo regular movement before sideward merger, sideward merger being disallowed from theta-positions (see Bošković 2023)—this will actually be suggested below regarding XP (see (97).)

    (b) With DP already in the structure, X externally merges with it via adjunction, with X (an ambiguous head/phrase) projecting. The DP then ends up being adjoined to XP, but there is no cycle issue, since adjunction is not subject to the cycle (Lebaux 1988). Under Epstein's (1999) approach to c-command, where α c-commands what it merges with, the DP part would not c-command anything from the positions to which XP moves (only XP gains "new" c-command possibilities).

    (c) Saito (2012) proposes that heads can be base-generated in an adjunction structure, in which case one of them excorporates, undergoing head-movement (see also Martinović 2015, where a head is generated with features F1, F2… some can move out, to project a phrase on their own). Thus, for Saito, v is base-adjoined to V and excorporates, projecting its own phrase after that. On a par with this, X is base-generated head-adjoined to D (as in Saito 2012) or it's a feature of D (as in Martinović 2015). X then excorporates, merging with DP, with X projecting (note this is head-phrase labeling); under Saito/Martinović-style treatment, this is in a sense self-merger—it's possible that, as with adjunction, the cycle is not relevant in such cases. Or cycle is phase-based—it can be "violated" within a phase (cf. Richards' 2001 tucking-in).

    All this may be relatable to crosslinguistic variation regarding whether *who* in *who left* undergoes agreement (assuming QP doesn't agree): in Kinande, where it doesn't, Q would then be added earlier than in English, where it does.

    So, FocP/TopP/QP is projected above the DP during the derivation (blocking binding by the DP that was possible before their introduction), when the relevant elements need to undergo the movements in question, since this is what drives them. This separates semantics/information-structure-related functions,

16

in the spirit of Chomsky's duality of semantics: Everything A-related is strictly formal, all formal licensing needs to be done/finished before semantics/pragmatics-driven operations take place. This also provides a new take on the Improper Movement ban (*A-A'-A): traditional A'-movement-driving features are not even present until all traditional A-movement is done, making Improper Movement trivially not possible.

The above captures in a natural way what it means to be e.g. a topic: the relevant marking/indication should be on the moving element, there is little sense to have it in a higher projection when that higher projection need not be there (when there is no topic), and then, what should it attract? The closest what? To implement such attraction there needs to be a marking on the lower/moving element (X) indicating it's a topic, but that may be all that's needed. X then moves where it can be interpreted as a topic (on a par with quantifiers/QR). During structure building, once formal properties are taken care of, if something (X) will be a topic, then it gets a topic feature/projection and undergoes movement. As soon as YP, where XP with a topic feature/projection, i.e. TopP under current assumptions, can be interpreted, is built, XP (i.e. TopP) moves and merges with YP. (This fits naturally with the mapping approach to information structure (Neeleman and de Koot 2008)).

Turning to additional cases, consider (97).

(97) Who seems to himself to be smart?

It's been argued there is a phase with *seems* (Bošković 2015, Legate 2003; cf. Legate's *which of the papers he$_i$ gave to Mary$_j$ did every man$_i$___seem to her$_j$ to dislike $*$*), which requires SCM through its edge. *Who* then stops at that phase edge as a DP in (97), binding *himself*, after which QP is added (cf. the discussion of DCCs in option (a) above, where regular movement must take place before derivational structure building).[8]

Notice the parallelism with DCCs, (98) being acceptable on a par with (97).[9]

(98) [Bill$_i$ and Sue$_j$] seem to himself$_i$ and to herself$_j$ to be the best candidate in the election and the best nominee for the convention respectively.

Regarding (99), the problem is that *who*, which is still nominal at the edge of the *seems*-phase, is too close to the pronoun, causing a Condition B violation. That this is on the right track is confirmed by (100), where the distance is increased by embedding the pronoun.

(99) ?*Who$_i$ seems to him$_i$ to be smart?
(100) Who$_i$ seems to his$_i$ professor to be smart?

There is again parallelism with DCCs:

(101) ?*[Bill$_i$ and Sue$_j$] seem to him$_i$ and to her$_j$ to be the best candidate in the election and the best nominee for the convention respectively.
(102) [Bill$_i$ and Sue$_j$] seem to his$_i$ supporters and to her$_j$ donors to be the best candidate in the election and the best nominee for the convention respectively.

The A/A'-distinction is also not needed for (103)-(104): in (103), the anaphor's nominal part is embedded at the point when the anaphor is higher than *Mary* (cf. fn 8)), so Condition C is not violated (Condition A is satisfied derivationally); in (104), *herself* remains a nominal, causing a Condition C violation.

---

[8]SpecvP through which *which students* in (91) moves may need to be below subject theta-position (this is actually everyone's issue, independent of our concerns).
[9](96a) and (98) together indicate that the derivational ConjP formation occurs after movement to the *seems*-phase edge; ConjP can be merged into the structure in a higher Spec of this phase, or higher up in Split IP.

(103) Herself Mary likes.
(104) *Herself seems to Mary to be smart.

Regarding long-distance scrambling that doesn't create new binding configurations, it may be information-structure feature-driven, as in Miyagawa (1997)—no new binding configuration is then created due to the lack of c-command; see, however, Saito (2003), where due to the way copy-&-delete mechanism applies, only phonological features move into the higher clause in the relevant case—note this analysis also doesn't require an A/A'-scrambling distinction, scrambling being a uniform operation.

   Consider also WCO in (105): The nominal part of the wh-phrase does not c-command *his* so bound variable reading is disallowed.[10]

(105) ?*Who$_i$ does his$_i$ mother like?

Strong crossover (*who$_i$ does he$_i$ like t$_i$*) can be handled as in the standard analysis.
   Note also that under the current analysis, even CP/phase works for clause-matehood for Condition B for (106)—there is still no need for the A/A'-distinction to account for the lack of a Condition B effect: the topic co-indexed with the pronoun is base-generated in SpecCP—since it does not start in a theta-position, it's never a nominal, hence does not induce a Condition B violation.

(106) John$_i$, Mary likes him$_i$.

Consider also BP (107). Islandhood and reconstruction indicate that without the case-marker (*de*), the topic is base-generated; under the movement derivation *de* is needed.[11] This is now straightforwardly captured: under the movement derivation, the topic starts as a nominal—*de* is needed since nominals need case. Under the base-generation derivation, the topic is never a nominal, which voids the Case requirement.[12]

(107) Esse dinheiro/Desse dinheiro, eu vou precisar para o teste.
      this money/of-this money   I will need   for the test   (Bastos-Gee 2011)

Finally, there are problematic cases where QR creates new variable binding possibilities: the quantifier c-commands the pronoun only after QR below.

(108) Some daughter of [every author]$_i$ hated some relative of his$_i$ new wife.   (Safir 2004)
(109) That people hate him$_i$ disturbs [every president]$_i$ (Reinhart 1983)

This is problematic under traditional assumptions, which rely on the A/A'-distinction and the assumption that A'- but not A-movement induces WCO. QR should then induce WCO in (108)-(109). The problem doesn't arise in the current system, where all that matters is c-command: the quantifier c-commands the anaphor after QR. This means QR is not associated with an introduction of an additional non-nominal

---

[10]Regarding (70), (70a) instantiates the well-known fact that some PPs are porous for binding (the discussion of (108)-(109) may also be relevant). In (70b), there is a TopP above the PP, which blocks binding.
[11](i) and (ii-ii) (due to Renato Lacerda) illustrate the islandhood and reconstruction contrast respectively.
(i) Esse/*desse livro, eu conheço um aluno que tá precisando
   this/of-this book I know   a student that is needing
 (ii) Do/??o   seu$_i$ melhor aluno, a Maria acredita que todo professor$_i$ precisa em uma hora difícil.
      of-the/the his best   student the Mary believes that every professor needs in   a   hour difficult
[12]I assume cases like *him$_i$, Mary likes John$_i$ concern competition effects (see Jovović 2024, who shows the effect is <u>not</u> about c-command), which are independent of our concerns. As for Condition C reconstruction effects (*which attack on Mary$_i$ does she$_i$ find unfair* vs *the attack on Mary$_i$ seems to her$_i$ unfair*), the A/A'-distinction here can be restated in nominal/non-nominal terms, or this can be captured in terms of Chomsky's (1995) Minimize-Operator-Restriction preference (the issue is actually not simply about the A/A'-distinction, hence will not be discussed here).

projection, which isn't surprising: the motivation for this projection is feature-checking—I have merely modified standard assumptions by assuming that instead of being a feature of YP undergoing movement, we are dealing with a projection right above YP. QR is standardly assumed not to be feature-checking driven. There should then be no non-nominal projection above the quantifier, hence no c-command issue arises in (108)-(109).[13]

6. Conclusion

The paper has highlighted the contextuality of syntax—there is a broader theoretical move toward contextuality that permeates many domains, like structure-building, labeling, islandhood, phases, phasal edges, spell-out domains, and the EPP. The contextuality of syntax was also shown to be non-Markovian in nature, hence it argues against treating syntax as Markovian in nature. One Markovian property was, however, maintained: the lack of a derivational look-ahead.

We have also seen that the contextual approach to the EPP opens the door for eliminating the A/A'-distinction from binding, in a way where all that is needed for binding is c-command: with traditional A'-movement, the issue is that the nominal simply fails to c-command.

References

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral dissertation, University of Connecticut.

Bastos-Gee, Anna Claudia. 2011. Information structure within the traditional nominal phrase: The case of Brazilian Portuguese. Doctoral dissertation, University of Connecticut.

Bennett, Wm. G., Akinbiyi Akinlabi, and Bruce Connell. 2012. Two subject asymmetries in Defaka focus constructions. *Proceedings of WCCFL* 29: 294-302.

Bošković, Željko. 1995. Case properties of clauses and the Greed principle. *Studia Linguistica* 49: 32-53.

Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge: MIT Press.

Bošković, Željko. 2001. *On the Nature of the syntax-phonology interface: Cliticization and related phenomena*, Amsterdam: Elsevier Science.

Bošković, Željko. 2003. On the interpretation of multiple questions. *Linguistic Variation Yearbook* 1: 1-15.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38: 589-644.

Bošković, Željko. 2008. On successive cyclic movement and the freezing effect of feature checking. In *Sounds of silence: Empty elements in syntax and phonology*, ed. by Jutta M. Hartmann, Veronika Hegedűs, and Henk van Riemsdijk, 195–233. Brill.

Bošković, Željko. 2009. Licensing negative constituents and negative concord. *Proceedings of NELS* 38: 125-139.

Bošković, Željko. 2011. Rescue by PF Deletion, Traces as (non) interveners, and the *that*-trace effect. *Linguistic Inquiry* 42: 1–44.

Bošković, Željko. 2012. On NPs and clauses. In *Discourse and grammar: From sentence types to lexical categories,* ed. by Gunther Grewendorf and Thomas Ede Zimmermann, 179-242. Berlin: de Gruyter.

Bošković, Željko. 2013. Phases beyond clauses. In *The nominal structure in Slavic and beyond*, ed. by Lilia Schürcks, Anastasia Giannakidou and Urtzi Etxeberria, 75-128. Berlin: de Gruyter.

Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45:27-89.

Bošković, Željko. 2015. From the Complex NP Constraint to everything: On deep extractions across categories. *The Linguistic Review* 32: 603-669.

---

[13]I leave for future work cases where bound-variable readings are disallowed. Note that anaphor binding may require a pre-spell-out phi-licensing Agree, which QR could not provide.

Bošković, Željko. 2016a. On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. *The Linguistic Review* 33: 17-66.

Bošković, Željko. 2016b. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 47: 1–33.

Bošković, Željko. 2018. On movement out of moved elements, labels, and phases. *Linguistic Inquiry* 49: 247–282.

Bošković, Željko. 2020. On the Coordinate Structure Constraint, across-the-board movement, phases, and labeling. In *Recent developments in Phase Theory*, ed. by Jeroen van Craenenbroeck, Cora Pots, and Tanja Temmerman, 133–182. Mouton De Gruyter.

Bošković, Željko. 2022. On the limits of across-the-board movement: Distributed extraction coordinations. *Philosophies* 7, 10. https://ling.auf.net/lingbuzz/006164

Bošković, Željko. 2023. Object drop in imperatives and the status of imperative subjects. FASL 32. https://ling.auf.net/lingbuzz/007473

Bošković, Željko. 2024. On wh and subject positions, the EPP, and contextuality of syntax. *The Linguistic Review* 41: 7–58.

Branan, Kenyon and Michael Yoshitaka Erlewine. 2024. Ā-probing for the closest DP. *Linguistic Inquiry* 55: 375–401.

Branigan, Philip. 1992. Subjects and complementizers. Doctoral dissertation, MIT.

Cable, Seth. 2010. *The grammar of Q*. Oxford: Oxford University Press.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, ed. by Stephen Anderson and Paul Kiparsky, 232–286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.

Chomsky, Noam. 1995. The Minimalist Program. Cambridge: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge: MIT Press.

Chomsky, Noam. 2008. On phases. In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133-166. Cambridge: MIT Press.

Chomsky, Noam. 2013. Problems of Projection. *Lingua* 130: 33-49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, strategies and beyond*, ed. by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 3-16. Amsterdam: John Benjamins.

Chomsky, Noam. 2021. Minimalism: Where are we now, and where can we hope to go. *Gengo Kenkyu* 160: 1-41.

Chris Collins. 1996. *Local economy*. Cambridge: MIT Press.

D'Alessandro, Roberta, and Tobias Scheer. 2015. Modular PIC. *Linguistic Inquiry* 46: 593–624.

Dias, Tarcísio. 2022. Hyper-raising and the subject position in Brazilian Portuguese. In *Proceedings of the 13th Generative Linguistics in The Old World in Asia*, 59-74. Hong Kong: The Chinese University of Hong Kong.

Epstein, Samuel David. 1999. Un-principled syntax and the derivationof syntactic relations. In *Working Minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 317–345. Cambridge: MIT Press.

Frampton, John and Sam Gutmann. 2002. Crash-proof syntax. In *Derivation and explanation in the minimalist program* ed. by Samuel David Epstein and T. Daniel Seely, 90–105. Blackwell.

Franks, Steven. 1998. Clitics in Slavic. Paper presented at the *Comparative Slavic morphosyntax workshop*, Indiana University, Bloomington, June 1998.

Franks, Steven, and Željko Bošković. 2001. An argument for multiple spell-out. *Linguistic Inquiry* 32: 174–183.

Gan, Linghui. 2022. Syntactic Structure of Argument Wh-questions in Hong Kong Sign Language. Ms. University of Connecticut.

Hagstrom, Paul Alan. 1998. Decomposing questions. Doctoral dissertation, MIT.

Holmberg, Anders, and Thorbjörg Hróarsdóttir. 2004. Agreement and movement in Icelandic raising constructions. *Lingua* 114: 997-1019.

Jovović, Ivana. 2024. Condition B and other conditions on pronominal licensing in Serbo-Croatian. *Linguistic Inquiry* 55: 402–421.

Kaisse, Ellen. 1983. The syntax of auxiliary reduction in English. *Language* 59: 93-122.

Lasnik, Howard and Mamoru Saito. 1992. *Move α: Conditions on its application and output*. Cambridge: MIT Press.

Lasnik, Howard (with Marcela Depiante and Arthur Stepanov). 2000. Syntactic Structures *revisited: Contemporary lectures on classic transformational theory*. MIT Press.

Lebeaux, David. 1988. Language acquisition and the form of the grammar. Doctoral dissertation, University of Massachusetts, Amherst.

Legate, Julie Anne. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34: 506-516.

Lewis, Beccy. 2022. British English *do*-ellipsis: an argument for phasal ellipsis. To appear *Proceedings of WCCFL* 40.

Martinović, Martina. 2015. Feature geometry and head-splitting: Evidence from the morphosyntax of the Wolof clausal periphery. Doctoral dissertation, University of Chicago.

McCloskey, James. 2000. Quantifier float and wh-movement in an Irish English. *Linguistic Inquiry* 31: 57-84.

Messick, Troy. 2020. The derivation of highest subject questions and the nature of the EPP. *Glossa* 5: 1-12.

Miyagawa, Shigeru. 1997. Against optional scrambling. *Linguistic Inquiry* 28: 1–25.

Mizuguchi, Manabu. 2014. Phases, labeling and wh-movement of the subject. Paper presented at the *32nd conference of the English Linguistic Society of Japan*, Gakushuin University, Tokyo.

Neeleman, Ad and Hans van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *The Journal of Comparative Germanic Linguistics* 11: 137–189.

Nomura, Masashi. 2005. Nominative case and AGREE(ment). Doctoral dissertation, University of Connecticut.

Nespor, Marina, and Irene Vogel. I986. *Prosodic phonology*. Dordrecht: Foris Publications.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge: MIT Press.

Pesetsky, David. 1987. Wh-in situ: Movement and unselective binding. In *The representation of (in)definiteness* ed. by Eric J. Reuland and Alice G. B. ter Meulen, 98–129. Cambridge: MIT Press.

Pesetsky, David. 2021. Exfoliation: Towards a derivational theory of clause size. Ms., MIT.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. Chicago: University of Chicago Press.

Rezac, Milan. 2003. The fine structure of cyclic Agree. *Syntax* 6: 156–182.

Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. Oxford University Press.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals.* MIT Press.

Sadock, Jerrold. 1974. Read at your own risk: Syntactic and semantic horrors you can find in your medicine chest. *CLS* 10.

Safir, Ken. 2004. *The syntax of anaphora*. Oxford: Oxford University Press.

Safir, Ken. 2019. The A/Ā distinction as an epiphenomenon. *Linguistic Inquiry* 50: 285-336.

Saito, Mamoru. 2003. A derivational approach to the interpretation of scrambling chains. *Lingua* 113: 481-518.

Saito, Mamoru. 2012. Case checking/valuation in Japanese: Move, agree, or merge. *Nanzan Linguistics* 8: 109-127.

Schneider-Zioga, Patricia. 1995. Specifier/head agreement in Kinande. *Cahiers linguistiques d'Ottawa* 23:67-93.

Selkirk Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3: 371–405.

Sigurðsson, Halldór Ármann, and Joan Maling. 2008. Argument drop and the Empty Left Edge Condition (ELEC). *Working Papers in Scandinavian Syntax* 81: 1–27.

Stepanov, Arthur. 2001. Late adjunction and minimalist phrase structure. *Syntax* 4: 94-125

Stowell, Timothy. 1981. Origins of phrase structure. Doctoral dissertation, MIT.

Sulemana, Abdul-Razak, 2019. Q-particles and the nature of covert movement: Evidence from Bùlì, *Glossa: a journal of general linguistics* 4: 99.1–21.

Takita, Kensuke, Nobu Goto, and Yoshiyuki Shibata. 2016. Labeling through spell-out. *The Linguistic Review* 33: 177–198.

Taraldsen, Knut T. 1986. *Som* and the Binding Theory. In *Topics in Scandinavian Syntax*, ed. by Lars Hellan and Kirsti Koch Christensen, 149-184. Dordrecht: Springer.

Yoo, Yong Suk. 2018. Mobility in syntax: On contextuality in labeling and phases. Doctoral dissertation, University of Connecticut.

Zhang, Niina Ning. 2010. *Coordination in syntax*. Cambridge: Cambridge University Press.