Lectures in the Minimalist Program*

Syntheses & Exegeses

<>

Joseph Galasso

Readings List:

(i)   (C)homsky, N. (2002) On Nature and Language. Cambridge University Press (CUP).
          · Chomsky's Chapter 4 'An Interview on Minimalism' (p. 92)
(ii)   (R)adford, A. (2009) Analyzing English Sentences: A Minimalist Approach. CUP.
(iii)   (G)alasso, J. (2019) Recursive Syntax. LINCOM Studies in Theoretical Linguistics, 61.

Abstract: This paper **'***Exegeses & Syntheses of the Program' ('ESP-paper') attempts to broadly sketch out the leading tenants of Chomsky's 1995 Minimalist Program (MP). The paper comes to consider the progression of 'Merge to Move', beginning with the principles of locality which operate over an array of Binding constraints, taking as the first instance Combine members (a, b) (an external merge), and then on to establishing an unordered Set {a, b}, and then to a local Move operation (internal merge) which establishes an ordered Pair <a, <a, b>>. From these sequences of external to internal merge-operations, an array of syntactic phenomena come into view, each of which enters some form of an explanatory equation, as argued for by minimalist pursuits. Other topics include Merge over Move, Phase-base theory, Light verb constructs, VP-shells, Principles of economy of movement, and Reasons for movement. The ESP paper was written as a graduate-student guide to issues surrounding MP.*

Finally, as a broad sweeping 'pedagogical device', we peer into myriad aspects behind Lasnik's 'Anti-locality' Condition. What does 'locality' exactly mean here (c-command)? How is it that adjacency is banned from recursive syntax (X-bar)? The condition stipulates that If an item gets displaced (internal merge), it cannot move into its existing phrase, but rather must expand a higher/functional phrase. How does this condition effect movement (e.g., wh-movement, head-to-head movement) regarding 'Merge over Move', as well as notions of transfer/spell-out involving phrasal projection? Lasnik & Saito: If head movement doesn't enhance and achieve any new configuration, or is too short and superfluous, then the movement is barred. (See p. 19 herein (P-10)).

Contents of Main Lectures

•'The Fibonacci series shows up all over the place' (C, p. 142).

•An opening note *On Merge and Move*: (sister-sister vs mother-daughter hierarchy, (G, 67-68).

a. Recursive hierarchy, X-bar theory, C-command, Binding. (R, pp 69-75), also see link in (c). (G, Ch. 4 p. 43, Ch. 6 p. 67)
https://www.academia.edu/6021575/A_Note_on_C_Command_weeks_3_4_cont

b. The 'Four sentences' (G, Chapters 3 & 8). (R pp 69-75 (C-command, Binding))
https://www.academia.edu/42204238/Working_Papers_and_Reflections_on_Syntax_no_3_Lecture

c. 'Fascinating vs Celebrating' typologies (G, 13, also see p. 13 of link below).
https://www.academia.edu/42204216/Working_Papers_and_Reflections_on_Syntax_no_1_Lecture

d. Child Language Note: Skinner v Chomsky interpretation of language design, associative/connectionism vs rule-based procedures, so-called 'vertical v horizontal' platforms of network architecture (AI-based language platforms: See G, 162). (R, p. 42 (Binary principle), 69-75 (syntactic relations), 70-75 (C-command)

a. vP/VP-Merge (= Case) (G, p. 125), VP is Semantic-Thematic [+Interpretable] (G, 125)

b. Number (plural {s}) has [+Interp], subject-verb agreement {s} as [-Interp]

c. Redundancy in language: 'These books' : These [+pl] books [+pl] (C, 111)

d. Theta-marking (e.g., Agent, Patient) (G, 24, 34, 125, 130), Control v Raising Predicates (R, 266).

e. Dual Probe-Goal relation (R, p. 14 (PF, LF), 119 (Case marking), 398, 402 .

f. CP-Move (=AGR), CP is syntactic-Formal [-Interp] (G, 127).

a. Dynamic Antisymmetry (DA) (labelling account, Moro). (G, 112-119ff, 129).

b. Merge creating lexical categories (G, pp 18, 32, 129, 148-153)

c. Compounding & Derivational morphology (G 10, 13, 103)

d. Root/non-move vs Synthetic/move) Compounds (G, 20)

e. vP (light verb) straddles the 'lexical/functional spilt' and Case as duality of semantics (e.g., oblique/Dative vs structure/configurational).

f. Locality constraints on Binding ('Anti-locality' conditions).(G ,51).

a. Why Agree, Why Move? (cf. Miyagawa), Merge over Move hypothesis

b. TP (Tense: affixal vs Featural), and hands over Case feature to vP (*vP marks Case).

c. CP (Force, and hands over Agreement to TP (*CP marks Agreement): Hence the two *Phases: vP, CP (while TP could be adjunct in nature). Assumption on Split VP-shells (Ergative predicates).

d. Kayne: 'all items move at least once—all right branching structures end in trace').

e. Phase-based theory (Marantz). Probe-Goal Union/relations (Miyagawa). Final Notes on Phase.

Organization & Prologue: Opening Remarks 'What is hierarchical design (sister/mother-daughter)?'

Let's begin by asking a fundamental question: What is the nature of *Number* [+Pl] in English?

Sure, we can 'spell-out' the plural rule as: [N + [s]] = pl, [[book]s]…(though of course, even here, at this very simple surface spell-out, all linguists are not on-board (cf. a 'singular pathway' model shows as undecomposed [books]) [1]. But even with such as rule, Chomsky asks: Aren't we just restating the problem? So, how do we attempt an explanatory account of e.g., *Number* in English, its processing, storage & retrieval, its conception?

⇨ Preliminary discussion! 'In the beginning, there is Merge'. While all things must be either 'singular or plural' (**category**), not all things must be a 'table, chair, nightstand' (**item**). Hence, Move spreads a singular-categorical rule across an array of items (C, p. 111).

```
            PL                    (where Move constitutes an 'Edge'-position)
          /    \
      Move    Merge        e.g. {furniture {table, chair, nightstand}}

      [ {s}      [N] ]            [Category    [items]]  [2]

                 [[table]s]
```
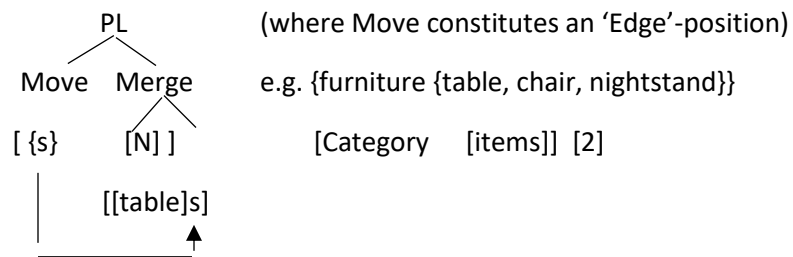
Merge & Move is the kind of natural distinction we find in the world. This displacement of 'item over category' (Merge & Move) is quite normal in the world. While Items must assume a Saussure-style 1-1 sound-to-meaning association in order to establish the concept (save of course homophones and ambiguity found in language), categories are free to adjoin to any predetermined class of items insofar that no actual singular itemized description must be part of a prescribed condition: viz., a plural marker can as easily attach to an unknown entity (a nonce word cf. Berko's 'Wugs test') as attach to a known one.

•The '**Four sentences**': an omnibus tour leading to the MP. (C, Notes on 'Editors' Introduction').

(i) Can eagles that fly swim?   (1955)   (See G, Chapters 3 & 8 for full discussion)
(ii) Him falled me down   (1960s)
(iii) a. The horse raced past the barn fell. (Garden-path sentences of 1970s)
   b. The boy Mary asked to speak to *John thinks he is smart. (anti-adjacent binding 1980s)

(iv) I wonder what that is ___ up there.   (See Ex. no. 11 pp. 131ff)

      a. *that's
      b. that is

⇨ Discussion on hierarchy in Language: How is language not a 'Beads-on-a-string' theory? [3]

•Prologue: Exegeses on Chomsky's Chapter 4 'On Nature & Language'.
   ⇨ Rationale of research program: The MP seeks out the design of language, the language and brain corollary, and sets out to understand a (L)anguage's traits:

•The Four Main Lectures ('Form Merge to Move': investigations on how principles of economy work on syntactic movement operations—Implications into 'phrase/phase' structure analyses).

Opening Remark & Analyses: (C, Notes on 'Editors (E)-Introduction').

Much of linguistics prior to Noam Chomsky dealt in traditional concepts of grammar of the Saussurean school, i.e., 'Language as a social object. Here, language was seen as something out there, an object, to be collected, partially mastered, as if words and patterns of language were *seashells* thrown onto the beachhead, where prospective individuals could gather up each shell, and perhaps classify and store based on color, shape, size, etc. (a real 19th century Darwinian exercise of piecing sound-to-meaning objects together). The idea that language was biologically based had no import in this model. This Saussure to Chomsky progression takes us from 'language as social' to 'language as biological'. (See Chomsky 1959). One of the first assertions that comes out of the Chomskyan model is the notion that language is 'creative'. *Creativity* of language up until this point had never been noticed before, since language had been defined as that mere collection of social objects (seashells) that get stored and retrieved based upon when an utterance needs expression. Chomsky asks: How is it that we can process a sentence we never heard (or uttered before)? Consider the creative sentence in (E-1) below:

(E-1) Yesterday, I saw a pink and yellow elephant roller-skating down Mulholland drive.

Now, I have never spoken those words in that precise pattern/order before, and yet we can understand it, and parse/process the sentence. There must be a certain amount of creativity going on here. Of course, one way to handle this creativity (perhaps what Saussure would say) is to suggest that a speaker has built-up over time an analogy of over hundreds of thousands of sentences which gives the speaker the mere approximate impression/resemblance of the sentence in (E-1). For instance, consider sentence (E-2) (and see if you find some 'resemblance' to (E-1)):

(E-2) Tuesday, I saw a 1965 Mercedes driving along the 405.

Now, if you do *intuitively* feel some resemblance between (E-1) and (E-2), ask yourself exactly what it is, where is it coming from? Clearly the *words* (which counter-juxtapose 'one to one') are very different. Well, if we strip-away the words (the colorful seashells), what we are left with is the under-governing skeletal structure, what we call a 'template syntax' devoid of words, but into which certain words can slot. This was an entirely new way of thinking about language. Secondly, it wasn't just creativity alone that moved Chomsky to challenge behaviorist theories of the time, but also the nature of *errors* child made developmentally: (*Him fell me down, I goed to the park,* etc.). Hence, the two-prong analysis of this new approach would be formed and supported by *creativity and errors.* What this new approach required was a model that incorporated *recursiveness* into the design of language—i.e., syntax is expressible by precise grammatical models endowed with recursive procedures. (*See the 'Four sentences' for discussion (G)).*

(Note decomposed parsing of [[fall]ed], [[go]ed] (as discussed in following section)).

Parsing: The *intuitive* resemblance is a residual trace of how the mind (i) first creates unfilled syntactic categories/slots and then (ii) fills-in each category appropriately (open syntactic slots filled with phonological material). Consider below one example of how a given syntactic slip (a mistake called 'a slip-of-the-tongue') can inform our understanding of how underlying parsing governs our surface syntax:

Slip: *What about taco_ tonights?*   **Target utterance***: What about tacos tonight?*

Note that the parsing of the plural {s} got detoured from N+s=PL, [[taco]s] and ended up as an erroneous suffix attached to the adverb 'tonight' [[tonight]s]: (a word/adverb which doesn't even avail itself to number in English grammar). But also note how the underlying recursive parsing of [[ ]] correctly functions as a viable morpho-syntactic underpinning for the word [tonight]: e.g., [[tonight]'s] (as possessive {'s}), or verbal clitic of {is}). Likewise, we can compare the brackets-[ ] for sentences E1 and E in this way. In fact, a quick look at the parsing of the two sentences (E-1 and E-2 above) reveals the same underlying parsing of slots. Let's flesh out our *Slip* each word in turn (showing brackets as part of a potential parse): 'What about tacos tonight?':

*(i) [[What]]...**(ii) [about]...(iii) [[taco]]...(iv) [[tonight]].

Where the underlying morphosyntactic slots appear accordingly: (Also see lecture 1i).

[[ ]]...[   ]...[[ ]]...[[ ]].

*(i) (Historically, wh-words were indeed decomposed as [wh]-prefix with matrix [at, en, ere, each, stems]: with a **wh-**question and **th**-response analogy: [wh[at]] => [th[at]], [wh[en]]>[th[en]], [wh[ere]]>[th[ere]]).

**(ii) (*About* is an undecomposed preposition [about]. The idea that 'about' may host a plural as in 'where-abouts' amounts to the fact that the phrase [where-abouts] is only an idiomatic expression, and number [Plural] is not productive in 'abouts': [abouts], and not [[about]s]), similar to [news]).

Now, we can expand on the '*creative & infinitive expressiveness'* of language by theorizing that language is not as a 'beads-on-a-string' theory (chez Behaviorism, which hold that all phrasal strings proceed undecomposed-[ ] and enter via 'full-listing' in the vocabulary, etc..), but rather that language is recursive in nature by maintaining a parsing-template mechanism whereby each slot can be filled via appropriate lexical/functional categories (with Nouns/Adjective, Verb/Adverb, Prepositions serving to fill the **lexical** category slot-[ ], and Determiners (A, The) and Auxiliary verbs (Do-Be-Have) serving to fill the **functional** category slot [[ ]]: e.g., [DP These [Pl] [[book]s]], [Aux Can [V speak]], [T [pres] [[V speak]s]], etc.

('These books are on the table': parse: [ ]...[[ ]]...[   ]...[ ]..[ ]..[ ].

('John speaks French': [ ]...[[ ]]..[ ]).

So, in this manner, as an exercise, see how the parsing of sentence (E-2) compares to (E-1): *Tuesday, I saw a 1965 Mercedes driving along the 405*... (compare it to E-1 and see if you feel the *intuitive resemblance*):

(E-2): [Adverb],..[Prn]..[V..] [D [Adj] [N]]..[[V]ing]..[prep]..[DP [N]].

One of the earliest issues was to tackle exactly what this recursive procedure was. This led (in chronological order) to transformation rules, phase- structure rules, X-bar theory, and eventually to *move-α* (move anything anywhere) of Merge/Move sequencing.

Part of the early model which would make up aspects of a species-specific yet universal 'Universal Grammar' (UG) including the research questions:

    (i)   What is knowledge of language?
    (ii)   How is it acquired?

The first taking as its research question the 'biological endowment of language', the second 'its developmental, maturational onset'. For the latter, questions quicky emerged on the nature of the very input children received (the so-called 'Poverty of Stimulus'—i.e., that the input children receive is so impoverished, incomplete, partial, fragmented, ambiguous, as well as incalculable in other respects which would quickly outstrip any potential language theory which took brute memorization as its core component (viz. Behaviorism).

Poverty of Stimulus

Let's take as a first example the fact of how speakers intuitively know how pronouns can and can't be referred to: (where underlines words co-refer/ *marks ungrammaticality):

(E-3)   <u>John</u> said that <u>he</u> was happy.

(E-4)   *<u>He</u> said that <u>John</u> was happy.

    Question: what kind of knowledge allows a speaker to intuitively know that (E-3) is grammatical but (E-4) is not? It can't be knowledge which is based on a linear strategy since (E-5) below is correct but mimics the word-order we find in (E-4).

(E-5)   When <u>he</u> plays with his children <u>John</u> is happy.

    So based on mere linear word order, E-4 and E-5 should be equally grammatical. But they are not. And this knowledge seemingly shows up both quite early in child language (say, after 4 years of age), as well as universally across all the words languages (for what we can tell).

    The only way we can explain this seemingly universal knowledge is to evoke a hierarchy-principle which governs the underlying syntactic-template (a recursive design, perhaps a by-product of the human brain).

        Of course, the acceptability of (E-5) comes to light once we see that movement has ensured (returning us to the true word order found in (E-4):
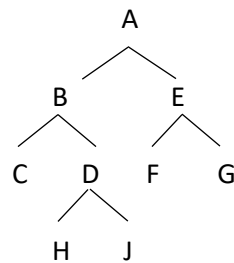
(E-5') [When he plays with his children [John is happy ~~when he plays with his children~~]]

       [...he$_j$ [John   ~~he...$_j$~~]]

(Note: for further discussion related to the Poverty of Stimulus', see Gordon's 'Rat-eater' experiment. (§1, (19) below).

But, let's take a deeper look into this intuitive knowledge regarding pronoun referencing (to be recapped later). Some brief notes: first, perhaps due to pragmatics, pronouns (He) must be FREE (they can't be bound to coreference within their **domain**. This is a form of 'Government and Binding' theory (G&B) of the utmost Chomskyan persuasion (Chomsky 1981).

Domain

So, what's a **Domain**? Let's say that: 'a Domain of the pronoun is the pronoun and everything that follows thereafter'—hence [PRN.....Domain). If PRNs must be FREE within their domain, then they can't coreference (Lasnik's (1976) Principle of Non-Coreference).

Domain of Pronoun: [Pronoun X...

                    ...domain of PRN X]

Once this recursive scheme is considered, the ungrammaticality of (E-4) comes into view: Let's delimit a domain by brackets [ ]:

    (E-3')   <u>John</u> said that [he $_{j/k}$ was happy]

        (PRN 'He' is free with its domain): it may refer back to 'John' (as underline/subscript j shows) but it may also refer to another John (with no reference to the John, as subscript k shows). In other words, the Pronoun is FREE.

    (E-4')   *[<u>He</u> said that <u>John</u> was happy]

        Here, the name 'John' is within the domain of 'He' while still trying to maintain its bound and forced coreference to the pronoun; hence its breaking of Lasnik's principle.

Likewise, this same recursive syntax shows up in wh-formations Lasnik's principle below:

(E-6) Which picture of <u>himself</u> does <u>John</u> prefer__?   (see C, p. 42)

So again, why is (E-6) above acceptable (with [Himself + John] surface ordering) given the unacceptance of (E-4)?

Well, we must note that the wh-phase 'which picture of himself' has been fronted, yielding the right domain structure.

(E-6')   a. <u>John</u> prefers which picture of <u>himself</u>
          b.* <u>He/himself</u> prefers which picture of <u>John</u>?

          c.   [John$_j$ said that [he$_{j/k}$ was happy]] = [He was happy]: PRN 'He' heads domain.

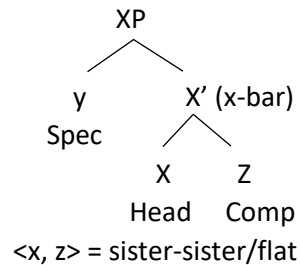   d.   *[He $_j$ said that John $_j$ was happy] = PRN 'He' must be free within its domain.

NB: So, again, none of this can be explained by sheer linear word order, via any simple linear principle, but rather only by hierarchical design (where domains can be created which instigate recursive structures). Speakers of all languages possess this principle and speakers seem to apply this automatically even in evaluating newly creative-based sentences. It's part and parcel of an intuition which we call the 'faculty of language' (C, p. 7).

Domain of analysis: (cf. 33)

C-Command   **Specifier>Head>Complement**

```
           A                                    XP
          / \                                  /  \
         B   E                                y    X' (x-bar)
        / \ / \                             Spec    / \
       C  D F  G                                   X   Z
         / \                                      Head Comp
        H   J                        <x, z> = sister-sister/flat

                                  <y < x>> = mother-daughter/recursive
```

C-command: We say that A is the mother of B, E (which themselves are daughter to A, and sisters between themselves). B is the mother of C, D, etc. Sisters c-command each other (both ways). Mothers don't c-command. So, A doesn't c-command (since A has no mother). B c-commands E, F, G. F c-commands only G, and H only J. E c-commands B (since sisters c-command themselves both ways).
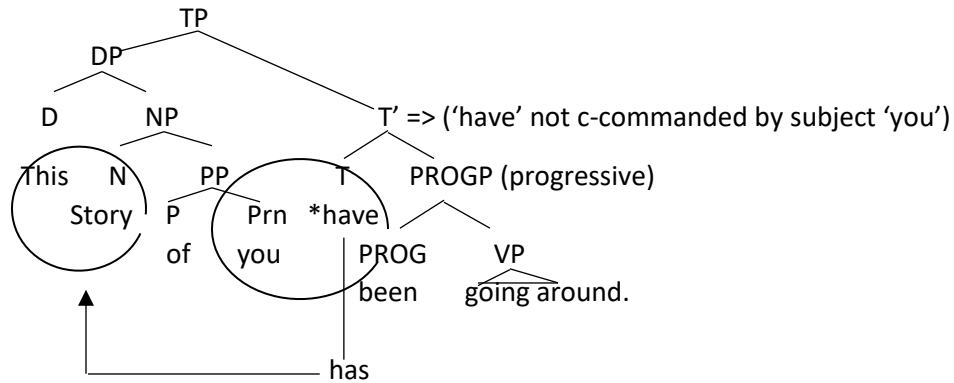
D contains H, J. B contains C, D, H, J. The Domain of (c-command domain) a Head H is the set of constituents c-commanded by H—for example its sisters and all the constituents contained within its sisters. The domain of C (CP) includes its TP (tense phrase) complement and all constituents of the TP. E is said to be in the Domain of A: Consider A = CP, E=TP below:

[CP [TP..]] Domain of CP…
```
          / \
         [...TP>vP>VP] = domain of CP
```
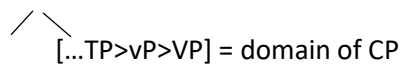
Let's flesh out a bit more 'Domain of' (c-command) below: (cf. 32b). Consider why 'have' in the sentence below would be ungrammatical when considered within a 'domain analysis':

(E-7) This story of you ***have/has** been going around (cf. 32b)

```
                        TP
            DP
      D         NP                    T' => ('have' not c-commanded by subject 'you')
   This    N      PP          T          PROGP (progressive)
      Story   P      Prn  *have      
              of     you        PROG      VP
                                been    going around.

                          has
```

   i. This story   of [*you have]…
   ii. [This story] of   you [has]…

Note that the verb must be spelled out as 'has' [3 person/singular] since the **c-command relation** is with the subject 'This story' [3P, sing] and not with the adjacent previous word 'you' (as would be found in the superficial linear word order). The DP 'The story' is the subject of its matrix verb 'has', and You is not the matrix subject of 'have'.

Note: This same type of binary branching (sister // mother-daughter) that we find in Domain analyses is also found in (the syntax of) phonology. Consider the horizontal spreading of Assimilation (a phonological rule): (see G, p. 67). Compare phonemes /_rs_/ in 'car' vs /rs/ in 'carson'.

Recall, *sister relations* of are the Merge flavor   ∧   'cars'

          [r]   [s] => /z/ (assimilation applies)

While *mother daughter*s are of the Move flavor:

                  ∧   'Carson'
         [r /\
           [s]] ….=> /s/ (no assimilation applies)

```
      ∧
   α     β            {α, β} is a 'sister-sister' relation. (flat structure)

      /\
   γ     /\           { γ {α,}} is a 'mother-daughter' relation. (recursive hierarchy)
        α    β
```

Active-Passive Move

Let's bring two different phenomena together under 'one roof' (of movement). Consider how 'active to passive' derivations work similarly to what we find regarding Pronoun-Domain expressions (C, p. 15):

(E-8)   a.   __ was washed the car (by Bill)

b. The car was washed___   (by Bill).

Note how in passive morphology the (active) subject in (E-8a) (Bill) (indicated by [_] can't have assigned its theta-role (θ-role) (viz., the first word/subject 'Bill' marked as [__] can't have assigned its theta-role AGENT/ACTOR, its role being optionally diverted via the 'by-phrase'). This is due to the mere fact that the displacement (of Bill) leaves an empty slot [__] behind, and non-arguments (including vacuous arguments) can't be theta/Case assigners in this way. When a speaker tries to parse [__ was washed 'the car'], the object/argument 'the car' is not in a viable position to obtain a proper theta-role (nor Case). Hence, by 'dethematizing' the deep-level specifier position (subject) in a passive construct (marked by __), the object (the car) must move upward into the leftover vacuous Specifier-subject position in order for this Spec position to be a Case/theta-role assigner—in accordance to EPP that all spec positions must be filled—so that the spec-subject argument can receive its theta role PATIENT. By dethematizing θ-roles, Case is also affected. (Nb. Case and θ-roles do work often in tandem, both being somewhat semantic in nature, e.g., Inherent case, Oblique case (e.g., by him/*he)). In this sense, Case theory is what's behind movement.

Theta roles: Bill = AGENT (actor of action), The car = PATIENT (the object undergoing/receiving an action). The vacuous Spec position in (E-8a) must be filled (viz., an EPP effect—extended projection principle that all declarative sentences (TP) must have a subject). The empty spec slot in TP must be filled, thus motivation the object to move into that subject slot to receive Case/Theta assignment. In this sense, the morphology of possessive 'intercepts' the assignment of the external thematic role (Agent) of the subject position and diverts it to the 'by-phrase' Object (it too being a case/theta assigner, as in oblique Case which comes after the Preposition (with whom/him/them/ with *who/*he/*they). So, while the subject of the passive is the Spec of TP 'The car', it being a non-acting PATIENT, the true active subject (AGENT) is 'Bill'. (See Lecture §3 'Merge: Local Move' (49) on Theta-roles.

    We are reminded of the same movement operation with so-called Ergative structures, where theta-marking/Case is involved (also see (34) and Split-VPs (52)). (See A for further discussions on ergative/unaccusative structures):

(E-9)   a.   The window broke__(by John)   (ergative) // (note: potential passive 'by-phrase':

b. John broke the window__.   a. OVS passive to b. SVO active)

c. There came a man.   (unaccusative: showing verb raising of 'came')

d.   A man came__

Regarding unaccusatives (such as [There came [a scream (~~came~~) from the woods]] we can see how arguments get displaced ('A scream came from the woods') showing expletive pronoun 'There' fronting and verb raising. Likewise, in (E-9a), 'The window' can't break something as an active AGENT. But 'John' can break a window. Hence, 'John' is AGENT, 'the window' is PATENT. (See Theta roles §3 below).

So, in order to get these arguments into the right theta-role order of assignments, John must move into Spec-TP:

(E-10)   (i)   [TP The window broke (by John)]

         (ii) [TP John   broke the window ~~broke~~ (by ~~John~~)]

This analogy has been advanced by Chomsky as part of economic considerations, via., 'move as last resort'.

    a. ___   is [a man in the garden]

    b. There is [a man in the garden] (EPP, Spec-TO must be filled)
    c. A man is [t___   in the garden]

    Now consider further raising constructs (See C, p 38 ex. (44):
    d. ___   seems [__ to be [a man   in the garden]]]
    e. A man seems [t' to be [t   in the garden]]] : [A man$_j$ seems [PRO$_j$ to be in the garden]]
    f. There seems [t   to be [a man   in the garden]]]
    g. *There seems [a man to be [t   in the garden]]]

Comparing (E-10a) to (E-8a), note how (E-10a) too has a leftover empty [ ] spec position, and so (due to EPP) either: (i) the subject of the lower preposition phrase ('a man') must move upward to fill the position, or (ii) the direct insert of the pronoun expletive 'there' fills the slot. If direct lexical inert of There' is going to be processed, it must be processed at the very earliest stages of the derivation in order to secure Merge over Move, and not get the ungrammaticality of (g) where it seems both merge and move have been simultaneously applied (in violation of principles of economy of movement). (e) shows big PRO.

'Merge over Move'
Note how in (f), 'a man' remains in its base-generated position, (as subject of preposition). 'a man' selects NOT to move upward, and instead the expletive 'There' is directly inserted to satisfy EPP. While in (g) 'a man' moves upward *creating a violation of Merge over Move). It seems that the language faculty prefers 'Merge-first' (as first instance/as default) over 'Move-as last resort', since merge of expletive 'There' comes at less of computational cost: i.e., 'There' is the result of a direct insertion into the derivation, directly pulled from out of the lexicon with no computational costs incurred by Move. This Merge over Move as a benefit to cost ratio is similar to what we find regarding 'Locality' (as discussed earlier). Short-local movement is preferred over Long-distant movement, and Merge (an external move of simple combine of [a] + [b] (as seen in 'There insert') being preferred over internal Move [a [~~a,~~ b]].
Note: For a fine illustrating of 'Merge over Move' (or 'Merge-first-then-Move-second' in terms of sequence of derivation) see (47) **Root vs Synthetic compounds.** For the same **'wine bottle' (merge) vs 'bottle of wine' (move)** analogy, see (50ff). (i) Merge-α [wine] + [bottle] , (ii) Move-β: [bottle of [wine ~~bottle~~]].

The principle of economy also overlaps with what we know about Locality and Phases:

(i)   'Locality: where movement is constrained from crossing more than two adjacent boundary-nodes. (Economy of movement),

(ii)   'Phase': where computational limitations require chunks of structure to get immediately sent to spell-out in communicative chunks (at phases: vP, CP), so to lower computational costs. (Economy of processing)

Locality.

For movement-based constituents [....X....Z....Y] (C, p. 40), it is said that no relation can hold between X and Y if there is an intervening item (Z) with the same structure as X, Y): hence, movement is maintained cyclically at a 'minimal distance'. (Note: unlike (quantum) physics where 'action at a distance' is observed (C, 52), this locality constraint seems to be more Descartes-like 'mechanical' in nature, e.g., observing Newton's extended near bodies of action, movement, etc. (Though, of course, many such non-trivial properties of language indeed go against this naïve 'mechanical account', as observed in recursive nesting and abstract structures, both of which move away from adjacency principles and are rather tethered to hierarchy (e.g. see 'Four-sentences] discussion, G, p78).

Let's flesh-out this locality principle below (**wh-movement, 'wanna'-contraction)**:

(E-11)   a. [How do you think [(how)$_j$ he solved the problem (how)$_j$]]? (where traces remain local)

b. *[How$_j$ do you wonder [who$_k$ solved the problem __$_j$]]? (trace can't skip boundary node)

Notice that only in (a) is the cyclic movement up the tree unhampered (the two open/upward slots are available to host cyclic raising of 'How' (in keeping with locality). In (b) 'How' can't raise up from being the complement of 'problem' to inserting in fronted position (of wh-question: so-called Wh-movement. This is due to the fact that an intervening word of the same structural type (i.e., a wh-word 'Who') occupies that position, thus blocking 'How' wh-movement. Perhaps the best illustration of this locality principle is to examine so-called *'wanna-contraction'* constructs: (showing both Wh-movement and Auxiliary inversion of' do') (See G, p. 69):

(E-12)   a. [Who do [you (do) want to help who]]? = (Who do you want to help?) (See Exercise 9, [30-31]).

b. Who do you wanna help? (wanna -contraction is fine).
c. [Who do [you (do) want who to help you]]? (= Who do you want to help you?)
d. *Who do you wanna help you? (wanna-contraction is blocked)

Note how in (a) want to are adjacent allowing for 'wanna' contraction at PF (PF being sensitive to surface/adjacency). In (d) [want // to] is disconnected by the intervening wh-word 'who', hence blocking surface PF 'wanna contraction' (moved who (in c) leaves an **empty category** behind).
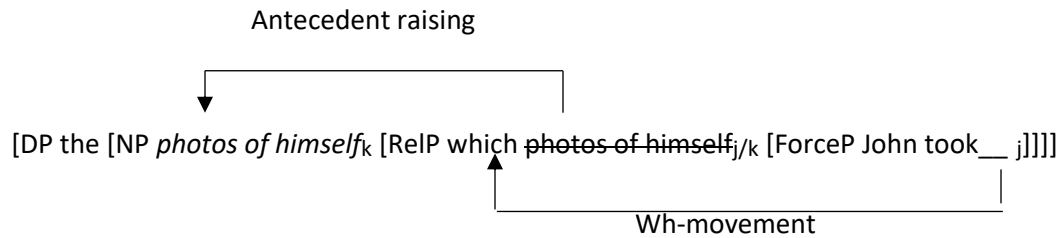
Note: also see 'sister-sister' local relations versus 'mother-daughter' non-local relations, (G., 67, also see note in (E-6') above).

Prologue (P): Exegeses on Chomsky's Chapter 4 'On Nature & Language': *Interview Minimalism***.**

(P-1) (p. 92). **Relative Clauses.**   One of the first MP problems: Providing rigorous rules for seemingly an infinite amount of structure: Paradox: How to give an accurate Descriptive account in MP-reduced Explanatory terms: *'For instance, Relative clauses look different from interrogative clauses…'*

⇨   Let's flesh out the steps of derivation in the relative clause below:

i.   '[The photos of himself which John took]… are great'.   (Radford, 2019, p. 47)

Antecedent raising

[DP the [NP *photos of himself*$_k$ [RelP which ~~photos of himself~~$_{j/k}$ [ForceP John took__ $_j$]]]]

Wh-movement

Trace-1 (j) Wh-movement

Trace-2 (k) Antecedent Raising

(Also see (49) for how theta-markings of a 'Wh-arguments' (as found in ForceP above*) can be assigned twice: (i) first as object of verb (VP), (ii) as an Interrogative element with a wh-feature (CP). *(ForceP reduces to vP/VP in the earlier steps of the derivation). (Radford here follows Douglas 2016 re. ForceP).

In this sense, Relative Clauses (RC) incorporate the kind of movement we already find in Wh-Interrogatives. So, we can say that movement is not provided as a uniquely rule for a specific grammar, but rather can be universally employed via (Move-α) (move 'anything anywhere') under Condition-β.

(P-2) (p. 94).   Chomsky expands on this by adding the principle of '*move anything anywhere'* under fixed circumstances yields differing surface strings: *There was a long debate over, say, 'John is expected to be intelligent'. For instance, is this a passive sentence like 'John was seen'…or is it a raising construct like 'John seems to be intelligent'? Well, the right answer is that there aren't any constructions anyway, no passive, no raising: there is just the option to dislocate something somewhere else under certain conditions. In one case you get a passive and in the other case it gives you a question, etc.* (See (49) Movement and Scope/Theta-marking)

(P-3) (p. 93).   *Poverty of Stimulus*: You can estimate the amount of data they (children) have quite closely, and it's very limited (e.g., see E-4 vs. E-5 above). (See G, p. 193). For example, Gordon's 'Rat-eater' experiment is a prime example of how children go beyond their input: children tacitly know not to allow inflectional infixes (between lexical stems). Plurals inside compounds are so extremely rare that children are likely to have never heard any; their inference thus in some sense probably goes beyond their data. (NB. Just because children may never hear given structures doesn't necessarily

prevent them from potentially generating them. Hence, there's some innate constraint: (i) [ ]+[ ] merge, (ii) *[[ ] INFL] + [ ]: (e.g. The response to the question: 'What do you call a person who <u>eats rats</u>?' is 'rat-eater', while 'rats-eater' is unattested—referencing the deletion of INFL when inserted between two lexical items: with [ ] + [[ ]s] being permit ([rat]-[eater[s]]) but [[rat]s]-[eater] banned). *Note implications to the full-listing, single pathway model here—if rats' is memory-based, as a full-listing model suggest, how do we explain the deletion of INFL is certain conditions? Gordon's experiment rather suggests a dual mechanism model between (i) rote-learned irregular/derivational items such as plural Noun [mice] (cf. mice-eater) as opposed to (ii) rule-based regular/inflectional items such as plural Noun [[rat]s] (rat-eater).

(P-4) (p. 93-4)   *Genetic endowment*: If you abstract those principles from the rules and attribute them to the genetic endowment of the child then the systems that remain look much simpler. Note how with X-bar Theory a multitude of possible structures can get reduced to a single mechanism. So, for example, X-bar was an attempt to show that Phrase-Structure (PS-rules) actually don't have the array of complexity once assumed. Eg. [+/-Head initial] now places divergent word orders within a unity context. (See G, A-2). There is no such thing as the VP in Japanese, or Relative Clauses in Hungarian, just the principle of 'Move anything anywhere' under fixed conditions (parameters). (cf. P-1 above).

(P-5) (p. 95)   *Economy* considerations must play a larger role than earlier assumed as motivated by '*Problem of Learnability*' (in Children)—which gave rise to P&P. As seen above in (E-11), any notion of 'learning'   which could lead to such abstract locality constraints must come from a UG endowment (where brackets-[ ] represent constituency markers—a theory-internal assumption not at all obvious in PF input).

   a. [How do you think [(<u>how</u>)ⱼ he solved the problem (~~how~~)ⱼ]]? (where traces remain local)

   b. *[Howⱼ do you wonder [<u>who</u>ₖ solved the problem __ⱼ]]? (trace can't skip boundary node)

   Some kind of a parameter must provide this knowledge: e.g., while 'How' and 'Who' might seem to be substitutable (of the same category), regarding cyclicity of trace vs boundary node their Head specificity creates boundary nodes, etc. (See G, 121 'Heads as Computational Atoms').

(P-6) (p. 96)   *Can Language be Perfect, Optimally Designed*? => Discussion. Given the above locality constraints, might there be broad, General-Cognitive (GC) (problem-solving) considerations for how cyclicity works: (i) Combine Member (SET/PAIR), (ii) Merge>local, (iii) Merge>distance, (iv) Move… which yields VP>vP, vP>TP, TP>CP. (From this, a design of locality constraints apply to Move-α).

   One way to think about how GC might playout is to consider how (i) first-order and prosaic MERGE schemes get bootstrapped onto latter (ii) second-order MOVE operations, an order which was suggested by Piaget as following a cognitive developmental scheme (notwithstanding the fact that language/formal syntax is not entirely tethered to a Piaget-like cognitive scaffolding (cf. Piaget vs

Chomsky debate of 1975, Abbaye de Royaumont/Paris). This same 'merge>move' progression found in the classic developmental progression of lexical category>functional category mirrors what we now know of incremental brain maturation of specific brain regions: Temporal-lobe/Wernicke's area>Frontal Lobe/Broca's area.

Optimality certainly is obtained when we consider how language makes use of **item vs category**. For example, Chomsky (C, 111) expresses the optimality view that e.g., pluralization as a rule which can be spread over an array of items (Nouns) speaks to economy and proficiency: 'So, in fact, plurality on nouns is rather like different words, just as you have 'table, chair', you have 'singular', 'plural'—and there are sensible reasons why plural should be an inflection and 'chair' shouldn't. Namely, everything has to be singular or plural, but not everything has to be a *chair.*

⇨ The idea behind 'item vs category' comes to us for free' as part and parcel of an optimality theory of language: Inflection is the result of how 'the spreading of rules' (generically) can equally be applied across an array of very different items (though of the same class).

⇨ The case of the above plural v singular optimality not being applied (as a spread of INFL) shows up in irregulars where indeed new words have to be introduced (child>children, foot>feet). The default rule of applying the INFL rule shows up in over-regularization (often found in child language), e.g., *childrens, foots, mouses*, etc). The child's (default) intuition that the rule N+s=PL is most efficient way to mark number exemplifies optimality.

(P-7) (p. 98) The *Galilean* tradition of dispensing with data in favor of maintaining theory.

⇨ *Theory comes first* (and if the data refute the theory, perhaps the data are wrong).
Consider 'wanna' contractions in this context. Data seem to support the general notion that a given infinitive 'want to__' phrase simply has the option of speech-reduction: i.e., 'want to' > 'wanna', and that there is nothing more beyond this mere surface, phonological-reduction phenomenon. However, as we now know, there are hidden, silent theory-internal considerations at play which don't necessarily show up in the surface data…One might simply deduce the data and give an overly simplistic account that says: 'the phase 'want to' can reduce to 'wanna' in fast, spontaneous speech, and that's all there is to it'. Of course, we now know otherwise: once a theory constraining such reduction was found, we quickly discovered that there was much more than mere surface phenomenon, viz., that there were hidden (non-phonological) empty-categories which went well beyond the surface and intervened in the deep-structure analysis (See E-12). (Also see 'empty categories/wanna contraction' found in [§§30-31] of Exercise 9 'Case Marking and Null Constituents'.

(P-8) (p. 106) One could also argue that language, as an abstract computational capacity, is less than 'optimally adapted' to the human performance system (with memory limitations, garden-path sentences, etc).

⇨ Chomsky argues here that even though **'communicative optimality'** does suffer under the dual stresses of *ambiguity and memory* limitations, even at times catastrophically, despite

this, there is some 'method to the madness', at least in terms of **'formal language optimality'** utilizes **recursiveness** as language's essential & defining attribute. The notion he provides is that the reason 'communicative' optimality suffers degradation is due to other 'formal' interface conditions which must be satisfied. For example, that Movement (perhaps 'non-optimal' in its own right) is employed in order to **check-off** certain [-Interp] features from a lexical item, etc. That some sacrifice in communicative optimality must be inherent to the formal system in order to satisfy other interface conditions. (See **Faculty of Language-narrow** (FLn) vs (**broad**) (FLb) in this context—with the former being exquisitely coopted by recursive syntax (hence, 'Language' (FLn) as narrowly defined as **recursive in nature**), and the latter (FLb) more **cognitive/motor-control-based** leading to general problem-solving skills. (See Fitch et al. 2005).

　•FLb = cognitive processes (lower-level brain function), substantive in nature,

　　(e.g., chimp communication, bird calls, child lexical stage-1, 1-1 frequency-item learning)

　•FLn = recursive syntactic (human-specific brain function), abstract, hierarchical nature.

　　(e.g., recursive & abstract, child functional stage-2, non-frequency/categorization)

⇨ For an Artificial Intelligence (AI) account leading to putative dual platforms, see G, Note-4.

⇨ One could ask: why don't they [speakers of a language] assume a strategy of a simple linear principle? Why is not a beads-on-a-string' theory an appropriate stratagem for a syntactic computation? It would seem likely that *linearity* would be highly ranked in the pegging-order of an 'optimality' theory. Yet, children innately know from a very early age that this is not so, and that a more abstract hierarchical system must be employed. While hierarchy may not be 'communication-optimal' regarding surface saliency, it fulfills on the backend the open-endedness required of language—viz., that language *be recursive, creative, and infinitively expressive.*

⇨ (p. 107) '*Maybe the whole architecture of the mind is not well designed for use*' (i.e., communicative purposes, but rather is designed to meet interface conditions).

　Telepathy: For instance, image the humans had telepathy—we wouldn't need phonology since the interface condition of sound-meaning is already satisfied. Also, consider what it is which required speech-sound/phonology to occupy the physical world of 'space and time'. This too is an interface condition: the fact that a word, say 'CAT' must be *strung horizontally* through space and time as /kæt/ (/k/ + /æ/+ /t/) and that any attempt to pronounce all three sounds all at once (as *stacked vertically* on top of each other) would go against satisfying PF *sensorimotor* conditions (p. 116). A Telepathic version of CAT would otherwise be optimal.

⇨ (pp. 109-115) '*Morphology is a very sticking imperfection. If you were to design a language you wouldn't put it in*'. Inflectional morphology, mainly (since derivational is substantive word-building). So, Inflectional: why even have it? (It's abstract, redundant, and doesn't at all carry interpretable communicative properties....Take for instance the INFL distinction as found in *structural* CASE (I vs Me). Why have it? (Sure, for non-configurational/free word-

order languages, CASE might be communitive [+Interpretability] for theta-marking/accusative aims, as is found in **Semantic—Inherent/Oblique/Ablative Case** (e.g., 'with knives' (instrument), 'with them' (theme/accusative)). But for Fixed word order languages (SVO) certainly **Syntactic—Structural Case** seems redundant. Consider child language in this context: a young two-year-old child saying 'Mommy, me do it' certainly serves its communicative aim. (So, what kind of a 'communicative -pressure' forces the child to move from 'Me' to 'I' (other than the fact that it appears in the input)? And of course, there's the cited *redundancy of morphological agreement*: viz., 'These books': why pluralize the noun 'books' if the determiner 'these' is already plural, or why 'She speaks': mark a singular verbal {s} in 'speaks' given that 'she' is already marked for singular number, etc.). (See (P-6) above regarding discussion of 'items vs categories').

⇨ This split of Case distinctions between '**semantic vs syntactic' properties** of Case gives rise to the **Duality of Semantics**:

   o (i) **[+Interpretable] Merge-based semantics**/LF (Inherent case), and
   o (ii) **[-Interp] Move-based/edge-related displacement** (LF>PF) (Structural case).

A Theory: There are surface-related properties (Edge), and there are LF-related properties—Displacement distinguishes the two *per se* (in consideration of interface conditions).

   (See Miyagawa who suggests expanding this duality to encompass phase-projections—(i) VP/vP (THETA-ROLE), (ii) CP (AGREE) distinction of phase projections). (See Radford 1990, Galasso 2016 for similar trajectory in child language development).

(p. 114) **Edge**: Surface-level (PF) shows up at the 'Edge' (for example VP-subject movement to TP (cf. EPP/VPISH)* [edge-subject [VP ~~subject~~]] (moving upward: e.g., words don't get pronounced in-situ (in their original position). The Edge position is characteristic of Displacement/MOVE, while MERGE is non-dislocated.

*(EPP *Extended Projection Principle* stipulates that all declarative/TP phrases must have a filled Spec-position/subject [TP Spec-subj [vP...]]. VPISH is the *verb phrase internal subject hypothesis*, showing, theory-internal claim that all subjects begin within VP for theta marking).

(P-9) (p. 114-115) **What drives Displacement?**

   The duality of semantics (as presented above) might explain the reason for Movement/Displacement. Chomsky considers the following:

   So, there are (i) the **LF-related properties/Merge**, and there are (ii) the more **surface-related (PF)/Move** properties. PF is expressed at the dislocated Edge (the 'left periphery' of the built-up syntactic tree—but note, that such displacement up the tree is driven by formal [-Interp]-features which must be checked-off. These LF **semantic properties** are more local, non-dislocated, and non-edge (such as

Local/merge for Theta-marking, etc). *Hence, there are two types of information the Faculty of Language (FL) is going to have to search for: **one Edge-related, the other Local:***

[Edge, PF…

[…Local, LF]].
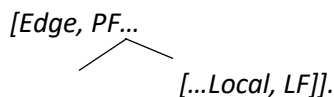
An expression will somehow have to distinguish these kinds of information, and in fact an optimal way of doing it would just be to resort to dislocation: expressions are phonetically interpreted at the edge (Move), even though they are semantically (thematically) interpreted at the local position, the position of Merge.

⇨ **Edge-features** (p 114). The *edge* of an expression typically involves 'specificity', viz., elements move leftward to the Edge in expressing some sort of 'specificity' (Force, etc.). Consider Object Shift in this manner: (e.g, *John gave Mary flowers> John gave flowers to Mary* (where the former is typically viewed within MP terms as the default structure (IO, DO order). *Flowers* (Direct Object) is seen as moving/shifting (leftward up the tree) to the Edge to gain specificity. (This also shows up in **Active to Passive** movement, where old information (*Mary*) is moved leftward to the Edge to show some specificity: *John kissed Mary > Mary was kissed*).

(P-10) (p.133) **Bare Phase Theory.** The main distinction here, theoretically, is between (i) First Merge Complement (V merge > N), and (ii) second Merge Specifier, a dislocated/elsewhere position which creates <Spec <V, N>>….with possible third merge, and so on…creating multiple Spec positions up the syntactic tree: $Spec^2 > Spec^1$ > vP>VP (noting here that TP would simply be the result of subsequent Merge-operations in order to create a Spec position).

Spec-2

Spec-1   vP….

multi-Spec

Spec   Head   Comp
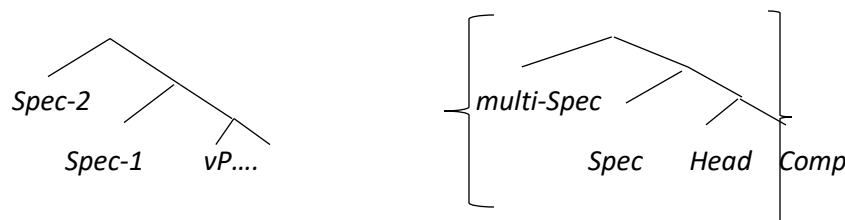
Note: **The Edge is composed on (Multi) Specifier + Head** (two positions which can host moved elements: e.g., movement from Comp of VP into Head of C (e.g., Wh-movement), Head-to-Head movement (e.g., French-style 'verb raising'), Spec-to-Spec movement (e.g., subject raising (VPISH)), etc. Note that a Complement is inherently *Frozen* in place (and otherwise gets sent immediately to transfer) unless displacement (Comp-raising) actives to expand a new/higher phrase. (See Galasso 2015 regarding *immediate transfer* for child language syntax).

In Chomsky's talk at *UCL, if I recall properly, in the Fall of 1994 (or Spring of '95), where we hashed out type-o's and a few minor mistakes as found the 1995 'Minimalist Program' manuscript, the first and only diagram he ever drew on the chalkboard that afternoon was the above reduced 'Spec>Spec>vP..structure'…Indeed, the true spirit of Minimalism.

Child Language Note (p. 134). If a child hears a Head (as pulled from the lexicon), and it's the first Head, the child's only option is First Merge (Head>Comp), or Verb>Noun. UG provides this algorithm. Head Merge with Complement (Heads being Verbs, which make up a two-word phrase. Otherwise, all individual words are potential Heads. One theory holds that all single words carry their own *atomic-features* which make up their (potential) Head status. Of course, one could argue that the defining aspect of 'Head' is that it must first work in tandem with a Complement (in this sense, its is the Complement of the Head which defines and provides Head status). Recall, that all lexical items can only be defined once they enter into a structural/syntactic relation (thus defining their +/-Head status).

Anti-Locality (Lasnik & Saito 1992): Let's consider the Edge as part of the Anti-locality condition—viz., If head movement doesn't enhance and achieve any new configuration, or is too short and superfluous, then the movement is barred.

(Comp to Spec allowed expanding Phrase XP to YP)

*(Comp to Spec movement banned within same XP)
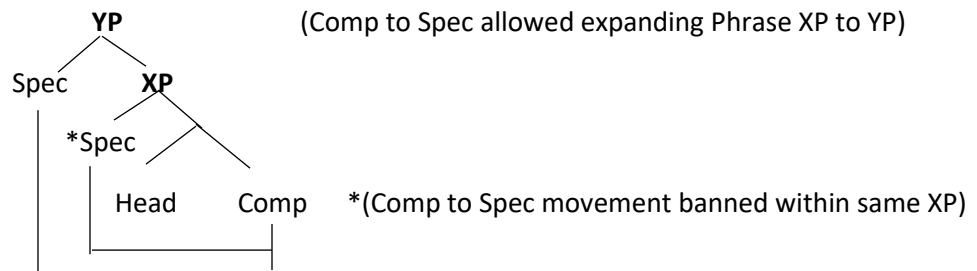
(i) **Internal Merge** (Merge) is allowed with XP (=Merge $\{\alpha,\beta\}$) for:
   a. Compounding [Adj Black] + [N bird] = [blackbird]
   b. Derivational morphology ($\sqrt{}$[V Teach] + {er}= [N Teacher]
   c. Lexical Phrasal V + N = VP [VP [V drink] [N milk]]

(ii) **External Merge** (Move) expands XP to YP (Move $\{\alpha, \{\alpha,\beta\}\}$) for:
   a. Inflectional morphology ($\sqrt{}$[V Teach] + {s} =(She teaches) [T {s} [V teach] s]
       (affix lowering from Head YP to Head XP)
   b. Functional Phase:
       Head-to-Head (French verb raising):

           e.g., [TP Je (ne) [T parle$_j$ [NegP pas [V ~~parler~~ $_j$]]]]
       Wh-movement:

           e.g., [CP Where$_k$ [C does [TP John [T does$_j$ [VP live where$_k$?]]]]]
           j) Head-to-Head Movement, k) Wh-movement

   c. Focus Phrase: (see Relative Clause (P-1) above)

Lecture 1. Introduction: Opening Remarks

To understand a 'Language' trait, we should:

(i) (p. 23) **Seek out the mechanism**:
a. Structure
b. Thought/Logic

(ii) (p. 36) **Seek out the genetic, environmental factors** (the ontogenetic perspective):

a. FL (Faculty of Language: innate)
b. Parameters (variations: environment)

(iii) (p. 51). **Seek out its functionality**—How does Language function? (Functionalism vs Formalism): Problems of function: redundancy, communicative factors, interface with PF/LF, e.g., garden-path sentences (see G, sentence #3), ambiguous structures, root v synthetic compounding operations, etc.

⇨ Question: It seems language is NOT optimal for function, but rather is optimal (perfect?) to meet and satisfy interface conditions (at PF, LF). Consider omission and commission as found in child language, or adult ambiguous structures.

(iv) (p. 56) **Seek out its evolutionary history** (phylogenetic: pongid/hominid split 6MYA (million years ago), formations of a putative Proto-language, Brain/Gene evolution (Broca's, FoxP2), etc., (e.g., ontogeny recapitulates phylogeny in young child syntactic development).

The Minimalist Program (MP)—begun in earnest by Chomsky 1993, 1995, and has since become more fully articulated in a myriad of recent works, e.g., Chomsky 2000-2006, et al— is precisely that, 'A Program' (by which several spin-off topics and proposals can be investigated: the Brain-to-language corollary, Child language, Neurolinguistics, Language evolution, etc. Chomsky emphasizes the fact that the MP is a theoretical framework, even an 'umbrella program' under which many diverging and converging models can come together (assuming certain universal principles understood to be inherent in the MP-framework—principles which underwrite language design). The central aspects of the project that we shall address herein are quite clear-cut: (i) To explain why and how language can arise only via an innately endowed human cerebral template (a language 'design'), as provided for by an account of Universal Grammar, and, to the best of our current understanding, to begin to delimit such a brain-to-language mapping of the template in ways which address the UG requisite—as language cannot be acquired by brute statistical leaning, as suggested in e.g., Baayen models of language learning. (ii) To attribute language usage in real scientific terms—and not by any naïve theory of language learning which solely relies on imitation, correction, reinforcement and memorization. This scientific tact harkens us back to the Skinner v. Chomsky reviews (1959) and contemporary spin-off debates— rule-based models vs. connectionism e.g., Marcus vs. Elman (1999) respectively—which could actually be revisited here as an instructive pedagogical device in highlighting and differentiating the issues at stake. (iii) To describe the cited brain-to-language corollary in ways which jibe with the data upon which meta-synthesis studies are brought to light on the theory.

The First Chomskyan Period: Descriptive Adequacy (Chomsky 1965)

The Chomskyan framework can be partitioned roughly into two general periods, with the trajectory of the research paradigm progressing from linguistic introspects of 'Description' to that of 'Explanation'.

The first thirty-years or so, say between 1958-1988—namely, the 'John Locke' lectures (Oxford, 1969) through to the 'Managua' lectures (Managua, 1988). This period was mostly concerned with *describing* language structure in such a way that would naturally lead to the devise of powerful models upon which Universal Grammar (UG) was grounded, and later upon which Principles & Parameters (P&P) of UG and even the earliest chapters of the fledging Minimalist Program (MP) were drafted. This period of time saw the field of linguistics turn from being a *humanities-driven* discipline to that of a *soft-science* discipline, at least in the sense that now models and computations were being used as heuristic tools, similar to what one would find in mathematics and biology. This new approach to linguistics began to give a more active hand to biology precisely because UG was to be defined as a 'species-specific' algorithm without antecedents to its kind found anywhere else in nature: language is a human exclusive activity. The new terms now drawn in this first period of the Chomskyan linguistic program (of the second half of the last century) gave way and saw its natural path evolve in the new century with what we might call the second Chomskyan period, whereby new pursuits would follow in asking how a UG-based language computation might be physiologically realized via a brain-to-language mapping. This new turn of linguistics brought together neurologists, geneticists, cognitive sciences, and biologists along for the Chomskyan ride, where many new insights would flourish in what we today call Linguistic Brain Science (referred to as Language Faculty (LF) Science).

In this second period, roughly starting from the 1988 Managua lectures period (which gave birth to P&P) and onward, and with the introduction of fMRIs, etc., being now utilized in language studies—as fMRIs were hitherto often relegated to the exclusive (non-linguistic) examinations of tumors and sports-based knee injuries—LF scientists are now in a viable position of evaluating how previously known brain regions such as Broca/Wernicke's areas might be extended and/or redefined in ways which implement the workings of UG/P&P. This second period, now well into its third decade, will be the phase in LF science where language meets the mind/brain. We are well on our way!

The Second Chomskyan Period: Explanatory Adequacy (Chomsky 1988)

The general lines of inquiry drawn forth from this second period are confined to that which might otherwise today be construed as quite natural of any linguistic investigation. While we now may take such lines of reasoning for granted, for several generations (and hundreds of years) prior to the Chomskyan revolution, such linguistic lines of inquiry had never been appreciated nor had they been fully pursued. The Chomskyan lines of inquiry are hence two-prong:

(i)   Period 1: What is the system of knowledge a speaker has that makes-up a particular language, in our case English of any given variety (developing, normal, variant, abnormal)?

(ii)   Period 2: What can we take away (neurologically) from the physical underpinning of the mind/brain that provides for such knowledge: what is the nature of this biological endowment of language? As a result of the advancement of neuro-imaging devices made over the past

few decades, a fruitful scientific research program has emerged seeking to establish a working brain-to-language relationship insofar that we can begin to account for the distinctions made between how the design of the (internal) brain identifies and partitions various (external) linguistic structure from out of the naturally occurring language input, what we will call the ambient language stream.

In the first period, culminating to the Principles and Parameters Theory (1981, G&B Theory), it became apparent that the tools for describing the world's languages were becoming too multi-facetted, as well as too cumbersome, with many paradoxes.

On one hand (the descriptive hand), linguists were finding that their elaborate rules to capture the language phenomena were increasingly becoming too enriched, more and more apparatus had to be assumed, theory internal. This was seen as a dilemma insofar that the whole aim behind a universal grammar was to reduce and abstract away superficial elements of languages (plural) in obtaining a core (singular) property. This goal of reduction was increasingly becoming out of reach. With every new empirical discovery, more elaborate and sometimes ad hoc schemes had to be invented. What tended to happen was to say, well, this is what happens to the verb in Japanese, as opposed to English, or that this new analytical tool seems only to apply in Finish, etc.

On the other hand (the explanatory hand), it seemed a robust 'restrictiveness' was needed since language (in its full recursive form) must be a unique (by)-product of the human brain/mind—since there is only one human brain, it becomes expedient to show how 'uniform singularity' can capture this brain-to-language corollary. This was what a singular and robust theory of *Merge* did for linguistic theory (MP)—viz., move anything anywhere (move-α), but under various tight constraints (e.g., binding and locality, movement based on [-Interpretability), the checking-off of formal features (Case, AGR), probe-goal relations, etc.). So, while the P&P model began to see how descriptive adequacy was on a run-away course, the MP's major role was to both rein-in the overpowering description and reduce it to a core minimum, hence the name Minimum Program. But it is sure that this 'minimalization' of description can trace back its roots to the P&P Model, where a reduction of most phenomena could be taken as a mere binary choice of a parameter. This would also greatly reduce the burden of language acquisition by children, the so-called *learnability problem*. The ease now of simply 'passively' setting the binary parameter 'one way or the other'—as there is no conscious role played by the child in determining such setting —co-opted whatever Piaget-like cognitive active learning the child was thought to have (See Piaget vs Chomsky debate of 1975, Abbaye de Royaumont/Paris). Here, rather, the parameters are seen as a part of the faculty of language (LF), the human endowment of language. The mere setting of the parameters is reduced to the subconscious selection of two switches (Off, On)—much like how our biological/genetic procedure is set-up (binary and not fully on-line, but rather part of our passive *procedural* knowledge. This is what guides our intuition about language structure.

 Structure

• What is the nature of recursive design found in natural language?

The Fibonacci sequence seems to turn up everywhere in nature whenever conditions on design are imposed: The Fibonacci series (0,1,1,2,3,5,8,13…) shows up all over the place. Must be a result of physical laws/properties inherent in the conditions of design which must be met (much like how Interface conditions on PF, LF must be met).

• Out of the Fibonacci series comes binary branching of Merge (as 'sister-sister' relation) >then displacement of Move (as 'mother-daughter' relation) as found in recursive syntax.

Structure as Fibonacci sequence

merge means add one item with another.

move means displace an item from an original position.

Adjacency means two items which sit next to each other.

The Fibonacci Code

The very idea that the way we humans string words/items together may have ancestral links to spiral formations found in shellfish is nothing short of stunning. Yet, the 'golden ratio' of Fibonacci holds: 1,1,**2,3,5,8**,13,21,34… etc.… for our language design. (Or, if you prefer to read the ratio as a binary rule, then [0 = 1], [1= 0, 1].

(Merge (add) first two numbers (adjacent) of the sequence to get the third number…and keep going: 1+1=**2,** 2+1= **3,** 3+2= **5,** 5+3 =**8**, 8+5=**13…**

(0) 'Fibonacci Spiral Formation' (like shellfish, snails).
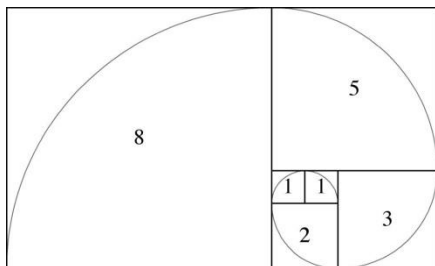
Here is the rule that **explains everything ever designed by nature** (the Fibonacci code):

0=1, 1=0,1. See how it works below through design:

Tree Diagram: 'Top-down' design: merge + move: 0= 1, 1= 0,1…

(1)

```
            0
            |
      1_____1
        /\                Merge [0, 1]
     0      1                 1
     |      /\          Move [1] out of Merge/set [0,1] => [1, [0,1]]
     1    0    1             2
     /\   |   /\
   0  1  1  0   1            3
   |  /\/\   |  /\
   1 0  1 0  1 1  0    1          5…., 8, 13, 21, 34,
```
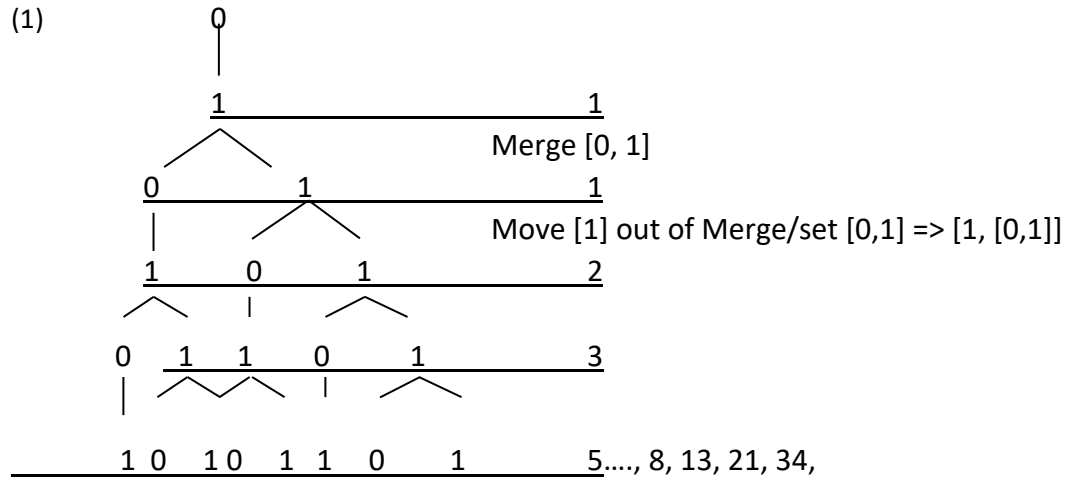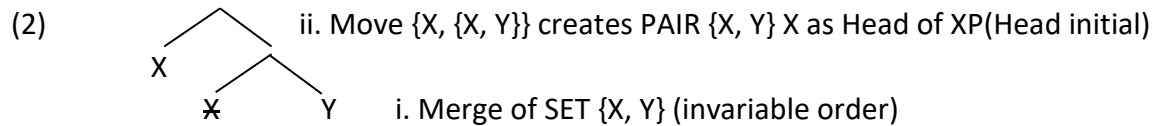
    (i)   Take the first sequence (combine *members/items* {X} and {Y}

    (ii)   The second sequence give us a *Pair* Move: (*Pair* yields word order).

(2)   ii. Move {X, {X, Y}} creates PAIR {X, Y} X as Head of XP(Head initial)

    X

       X̶   Y   i. Merge of SET {X, Y} (invariable order)

(Nb. Keep in mind our earlier discussion of sister-sister//mother-daughter relations—e.g., prior to Move of X̶ to higher position, {X, Y} were sisters.

This is the beginning of formal syntax—a beginning which yields: X-bar theory, Phrase, Word order, C-command, Root vs synthetic compounding, phonological constraints (e.g., on Assimilation), developmental child language acquisition (L1), distinctions in Autism, Specific Language Impairment (SLI), issues in Second Language Learning (L2), discussion on platforms governing Artificial Intelligence (AI), etc.

⇨ Discussion: Labelling algorithm comes out of PAIR via internal Move. In the MP, there are no longer Phrase Structure Rules which generate phrases. Phrases are a product of Merge: a Head moving to create a
Label {y}: { y { a, b}}. This label {y} by definition must be asymmetric (see Dynamic Antisymmetry, e.g., Moro, see (12) below).

At this preliminary stage, let's just take a closer look at how flat 'sister-sister' (non-hierarchy) get supplanted by recursive hierarchy, using both phonological and morphological processes as a model:

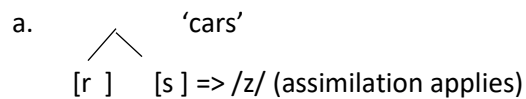(2')   ii. Move {X, {X, Y}} creates PAIR {X, Y} X as Head of XP(Head initial)

     X

         X̶   Y   i. Merge of SET {X, Y} (invariable order)
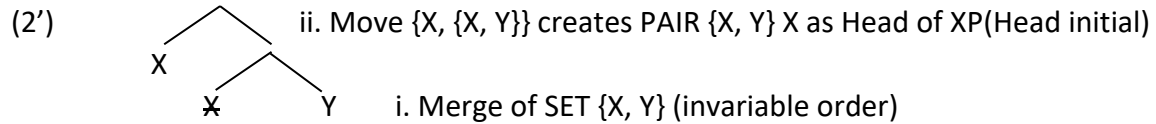
Hierarchy in Phonology. Perhaps the simplest illustration of how hierarchy works in language is the example of phonological assimilation. Note below how only flat sister branching of two adjacent phonemes allows for Assimilation (whereby the /r/ affects its neighboring sound/phoneme /s/, voices it, and changes it to a /z/ (a). Note how mother-daughter relation in (b) violates an *adjacency condition* placed on assimilation (whereby no voicing assimilation applies across the hierarchical branch:

S*ister relation*: 'Merge/Combine'   a.   'cars'

                                    [r]   [s] => /z/ (assimilation applies)

M*other-daughter* relation: 'Move/Displace'

                          b.   'Carson'

                              [r

                                  [s]] ….=> /s/ (no assimilation applies)
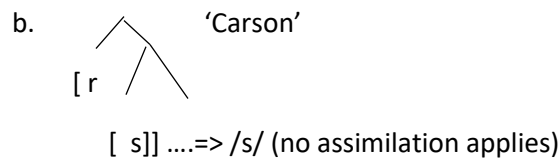
Hierarchy in morphology. Here, the hierarchy of language provides a dual pathway (dual mechanism model behind morphology). We note that 'books' must be stored as [stem] + {affix} [[book]s] (two separate modules), and not (as Behaviorism/connectionist models would have us assume), as a single module [books].

                     PL

             Move   Merge
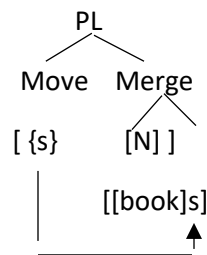
          [{s}   [N]]   Note: a dual pathway of [[book]s], not *[books]

                    [[book]s]

Merge and Move are products of Brain design:

(3) So there are two ways the brain processes information via design:
   (a) **Linearly** [ ]: where adjacency counts: [ ] + [ ] + [ ] etc. simply add adjacent objects/words together [x] [y] [z] where x affects y and y affects z (a domino effect). For example 'Ben is riding a unicycle' (five words sit next to each other).
   (Note: Early child language theory shows that small children's utterances are 'flat', a beads-on-a-string' theory without recursiveness).

   (b) **Non-linearly**: [[ ]]: where two things don't have to sit next to each other: [x [y]z] where x affects z but not y). (See 'Four-Sentences' discussion found in G. Lectures 3 & 8).

   This non-linear stuff is **very strange**. All computer languages, games, etc. depend on bits of information that sit next to each other (like binary code of 0s and 1s for computers)

Computer language of zeros and ones (0, l) depend on adjacency.

But not language (language is very strange…things can affect other things from a distance. For example:

(4)   '[Ben] [is] riding his unicycle'

        Question: Is Ben   ~~is~~ riding his unicycle?

To make it a question you invert the word 'is' which sits next to the word 'Ben': so, [Ben] + [is] then invert [is] to get [Is] [Ben] ~~[is]~~…? So fine, **this is linear**. But now look at this sentence:

   i.   'Ben who is my friend's son is riding his unicycle'

      ⇨   If you invert the closest/adjacent [is] you get the wrong structure:

   ii.   Is Ben who ~~is~~ my friend's son is riding his unicycle?

   iii.   *(Is Ben who _ my friend's son is riding his unicycle?) (* = ungrammatical).

Did you notice here that 'closest adjacency' just doesn't work: the closest [is] cannot invert. We must rather turn to the tree diagram in (1) for some structure:

Embedded structures within structures look like this: [[ ]]…not just linear [ ].

        (your grammar class calls them relative/embedded clauses)

(5)   [Ben [who is my friend's son] is riding his unicycle]

    a.   Is [Ben [who is my friend's son] ~~is~~ riding his unicycle]? = correct question

    b.   (Is Ben who is my friend's son _ riding his unicycle?)

Even though 'is' (in 4) is *distant* in closeness (adjacency), the correct 'is' is *closer* in structure.

Here's the structure just like our tree diagram in (1):

[Ben [who is my friend's son] is riding his unicycle]. [[ ]]…

It looks like this [[ ]]…

…and not like this [ ] .

[Ben who is my friend's son is riding his unicycle].

Chomsky draws attention to this as early as 1957 when he presents arguments surrounding the nature of the question in (6) below:

*(Chomsky 1957, p. 22: It is clear, then, that in English we can find a sequence a + S_1 + b, where there is a dependency between α and β, and we can select as S another sequence containing c + S_2 + d, etc.)

(6)   [Can eagles that fly swim]?

      (If we assume the wrong 'Flat' structure-[ ] (as shown in (6), it doesn't provide the recursive embeddedness which delivers the correct answer to the question: It is the second verb 'swim' we are asking if eagles can do, not the first verb 'fly' even though one might think that adjacent/closeness is how speakers peg subject to verb. This adjacency principle is not correct; speakers of natural language rather rely on structure-dependency driven by recursive design. Mere sister-to-sister relations for parsing a sentence doesn't suffice. So the question is: Where does this capacity to break with 'closeness of adjacency' and rather turn to 'closeness of structure' come from?

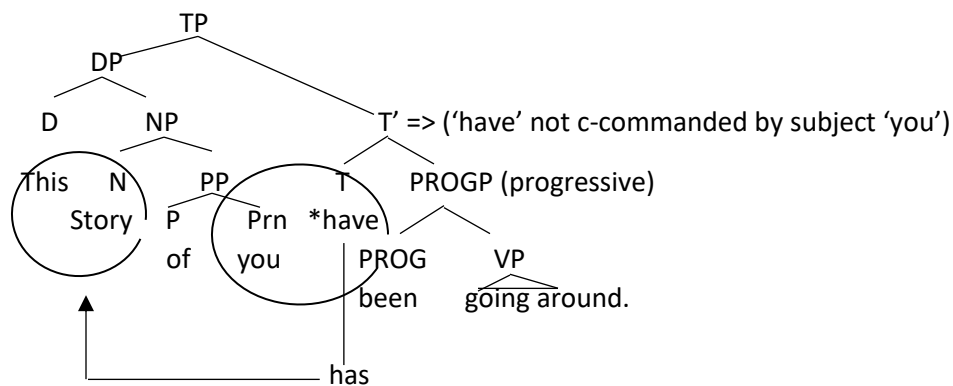   Here's proper recursive breakdown of the sentence in (7) below:

(7)   [Can eagles [that fly] swim]?

Now we see that in fact 'a closeness principle' of a kind is applied, but not closeness in terms of a flat sister-sister relation of merge-based items, but rather closeness in term of a mother-daughter relation: flat [ ] vs recursive [[ ]].

Also, note our discussion herein on how such recursive embedded nesting can affect our understanding of C-command, whereby the structural subject of the verb 'has' is not the adjacent Noun/pronoun 'You', but rather is the DP 'This story'…note how one can remove the nested Prepositional Phrase (PP) [of you] and still maintain the sentence structure:

a. [This story [ ] has been going around]:   [This story [of you] has been going around]
b. *[This story of [you have been going around]]

This story of you **have/has** been going around (cf. 32b)

TP
DP
D   NP   T' => ('have' not c-commanded by subject 'you')
This   N   PP   T   PROGP (progressive)
        Story   P   Prn   *have
                of   you   PROG   VP
                                    been   going around.

has

i. This story   of [*you have]…
ii. [This story] of   you   [has]…

_____

*(Note: Also see Merge-based [fascinating] vs. Move-based [[celebrat]ing] types in 1.ii below).

In that discussion, note:

--'Merge-based/combine' DERivational morphology: [V fascinate] + {DER-affix ing} => [Adj fascinating],

(a sister-sister relation of [stem] + [affix] (= derivational morphology)
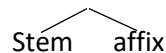
Stem   affix

--'Move-based/displace' INFLlectional morphology of [V celebrate] + INFL-affix ing} => [[V celebrat]ing],

(a mother-daughter relation of [stem + [affix]] (= inflectional morphology).

stem

affix

Modeling C_{HL} (Computation Human Language):

(8) Models of C$_{HL}$ (Computation Human Language):

Thought/Logic (Formal)   Two potential Model:
        |
    (point A)   a. Model 1 ↓   (point A)
        |
    CHL array of lexicon   b. Model 2 ↓↑   (point A)
      /   \
   PF   LF

Model 1: Forward-only: point A to PF, LF (look-ahead model)
Model 2: Forward-backward (fountain model).
In Model 1, PF and LF are not connected. In Model 2, point A is derived by a prior PF, LF
derivations: viz., all lexical items have already gone through CHL and have returned to lexical
array (i.e., a lexical item/Head leaves a residual trace of itself behind as found in prior constructs
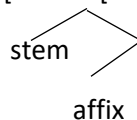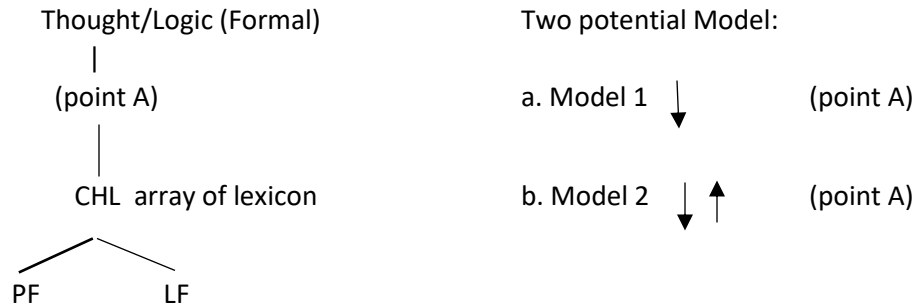which have made its way through PF, LF. The notion that Heads contain treelet structures will be
discussed as 'institutions of something missing in a derivation. (See the problem in (9) below).
Discussion: Imagine you pull out from the lexical array the item 'DIG'….
Question: Is the item DIG just a generic phonological shell, or is the item already replete with
'features' henceforth derived (at a deep structure)?

For instance, when you pull DIG from the lexicon, the 'intuition' is that something is missing—
which YOU or THIS could provide—e.g.,
(a) You dig?
(b) Dig this!
So that 'dig' either is {a}, or {b} of the set {a, b}, and thus must either become the Head/Target
of the projection (dig this), or the Complement of the Head (You dig) (where Head is <~~you~~ : <dig
this>> (TP-max prog), and comp is <~~you~~ ~~dig~~ : < dig this>>> (VP-min prog):

(9)   Dig   (TP Max projection)
        /   \
    You   dig   (VP Min projection)
             /   \
          dig   this
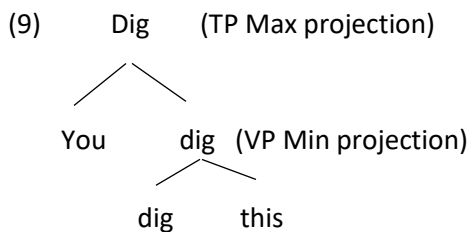
Consider DP 'This ditch' as both a min and max projection (a dual status):

29

(10)   DP

```
        DP
       /  \
      D    NP
    This  ditch        'ditch' is both min and max projection
```

Problem: 'Ditch' has a dual status:
  (i)   Min projection of DP (X') X-bar.
  (ii)   Max projection of NP (XP). (Discussion of Model 2)

        {this {this, ditch}}

'ditch' as V-bar: [V' [V ditch] [N this]], X-min proj = Head….Or 'ditch' as XP max projection:

(10')   XP (DP)

```
        XP (DP)
       /    \
     Spec    X' (v')
     This   /  \
         Head  Comp/NP
        (This)  ditch
         Ditch  this
```

⇨ The Problem: When we select DIG from the array/lexicon, there's an intuition that somethings missing: where does this intuition come from? : via (Logic, LF) formal… an a prior derivation via CHL. (?). Could Heads found at LF already be replete with this intuition (of missing features) that only PF surface-level structure could provide?
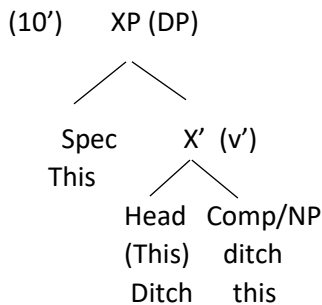

Model 2 suggests that singular lexical items are equipped with features as if items {a, b} carry with them 'treelet structures' based on the atomic features of the Head. Consider a second example: item DITCH. DITCH can either be pulled out of the lexicon as (i) DITCH this! [VP [V Ditch] [N this!]] (a min projection) or (ii) 'This DITCH', a DP Max projection [DP [D This] [N ditch]]. The computation comes out of the relation of Merge.

Model 2. (forward/backward). The 'items' out of the lexicon/array come both as Heads (X-min) and projections (X-max).

⇨ Atom features (of Heads) are:
  (i)   Heads (X-min)
  (ii)   Projections (X-max)….Hence, Heads/words are defined by their surface-level PF structures. It's no use talking about structure only at LF since for our interface to work, CHL (Computation of Human Language), LF must manifest itself and be accessible via PF to determine structure.

Note: Consider the double semantic theta-roles of 'which books' below (see ref. in G, p. 35):

'John read which book?'
       (i)   [Which book] first starts out as {a}: <object of 'read'> (read what?)
       (ii)   [Which book] then moves to take over a semantic role of {b} <interrogative operator> (which book?)

                Theta role {a}. Object/Theme
                Theta-role {b}: Interrogative

Showing full 'dual semantic role' derivation: [Which books did [John (~~did~~) read ~~which books~~?]]

Question: In order for these semantic roles (an LF interface requirement) to materialize and be met, mustn't they first be pronounced at the surface phonology (PF) in order to secure sequence of role mapping? Might a treelet structure, as residual trace of 'intuitive intent' be somehow encoded as an inherent element of the atomic feature?

CHL provides for these 'Atomic Features' as mad apparent by deep structure LF, but they must then be spelled-out at PF interface: viz., Atomic features are inherently Head driven since all items start out at LF as potential Heads from out of the array: the end result of Spec and Comp merely being a PF artifact byproduct of the number of merges: first merge = Head and Comp, second merge gives you Spec, etc.). (Note how 'number of merge-sequences' can only be obtained via PF interface). In Bare Phrase structure theory, the distinction between Complement and Specifier disappears, there is no difference: it's just first Merge, second Merge, and so on…if it's attached to a head, we call it Comp if it's first merge…etc. (See C, p. 133).

    So, as shown above, when the item DITCH is selected as (i) 'Head', then X-min (V-bar) is provided at PF [V' ditch this]/predicate…when selected as (ii) 'Projection', then X-max (DP) [DP This ditch]/subject…

    Also note that it seems LF is not Functionally optimal since only the upper-most part of the syntactic tree gets pronounced. (Once items [-Interpretable/Formal features have done their job (have got check-off), they can't be revised to do it again…they are banned from entering a computation later on. For example, 'once structural Case has been satisfied, you can't satisfy it again somewhere else' (See C, p116). Since only the upper-most part of the tree gets pronounced, this might be an external/outside world constraint (a phonological motor control skill or performance constraint.

    E.g., @Representation of Thought (LF): 'I like Plato'
    @ Derivation/PF: Plato, I like__.   Derivations just might be PF factors.

Thesis: Syntax (in MP terms) is defined as the interface between PF & LF: (CHL = syntax).

(11)   Thought/Logic (Formal)

          Interface> | | ⟶   computations:

                        CHL   ╲   i. representations

                     ╱   ╲   ii. derivations
              PF   LF
        Outside world   Inside world
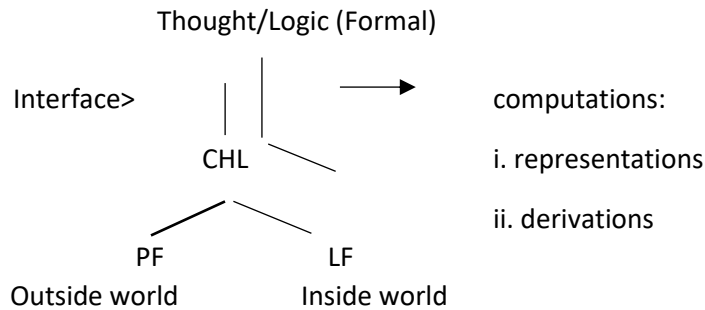
Computation starts with an array of lexical items (a lexicon), Head have their own 'Atomic features'. Thesis: Projections (Head & Max-projections) flow up and down between CHL and LF (as indicated by second line/pathway leading 'Thought' directly to 'LF').

Verb: [+V, -N] => 'DITCH this'… [V' DITCH this]/= as predicate
Noun: [-V, +N] => 'This DITCH'…[DP This DITCH]/= as subject

CHL operates on: Heads (X-bar-min) and Projections (XP-max). It's the feature of one of two items which determines projection.

        i.   First Merge = [Head Comp] sister-sister flat relation
        ii. Second Merge = [Spec [Head-Comp]]: mother-daughter hierarchical elation

Chomsky states: After the first merge, when the second expression comes along, the child is confronted with the question: does this (item) have the same semantics of one of the positions of the hierarchy, because it has some interpretation…if so, then the child should postulate a new head. (C, p. 135). This suggests some 'look-ahead', some residual trace of atomic features already assigned to H.

(12) Labelling (dynamic asymmetry: Moro):

'Y-Label' requires movement. In order to break flat sister-sister relation (of two Heads), movement must be employed (referred to as 'Dynamic Asymmetry' (DA) cf. Moro 2000):

        a. {House Boat} is ambiguous as {H+H}: there is yet no hierarchy:

                    {a, b}   (= 'a kind of a house /and, or/boat': <ambiguous>)
        First merge: ╱   ╲
                    [a   b]
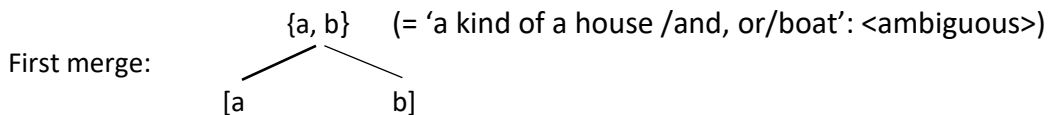
{a,b}: sisters, no hierarchy: Head-Head (one of which could be the complement of the other)
(Note: word order deviations in child syntax support arguments for a flat first-merge stage).
It's first a complement which merges with a Head: in the case of (a) above, either {a, b} could be
Head or Comp. It's the merging of a well-defined Comp which yields a Head. (cf. 'look ahead'
Model 2).

b.   [House [house boat]] (= only 'a kind of a boat')

{a, {a̶, b}} => 'Y' is labelled (as phrase-Y)

Second merge:

[a'   [a'] is probe
                                                 [a] is goal (of Probe-Goal relation)
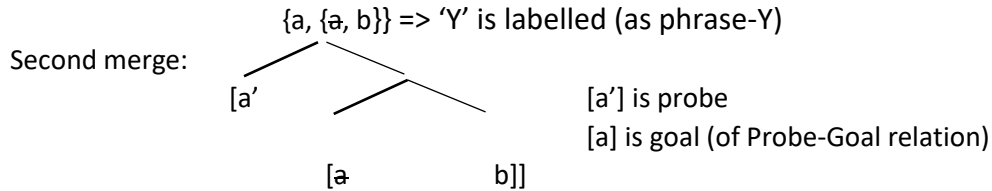
[a̶   b]]

[a'] targets [a] below to raise forming a spec position (creating DA).

{a,{a, b}}: mother-daughter hierarchy/creates a Specifier position.

Merge of members [a, b] yields a 'SET': Set = {a, b} with No order. PAIR derives order via movement [a, [a, b].

Progression-classification of an item:
1.   Member: 'individual item'-level (pulled from lexicon/array)
2.   SET: level of 'two items' (no order of items)
3.   PAIR (label, phrase) (order): where 'Y' is labelled.
*(See Note 1 (end of text) on the notion that 'all right-branching structures must end in a trace'—viz., that everything must move at least once (Kayne)).

Y-phase' doesn't come from the lexicon, it's a result of computation:
(i)   First, lexical item (Head)
(ii)   Second, Combine (which defines Label)*
         *Label (Y) is defined via the manner of Combine/Move.
         Merge happens at CHL prior to PF, LF, but (as suggested by Model-2) there is a back-and-forth flow between CHL and LF. The possibility here suggests that we must determine and define what Y is before PF, LF.

                                                          PF (PF to LF: both where properties get expressed)
(13)   Lexicon ⟶ CHL ⟨
                                                          LF (LF to PF: the classic 'sound-meaning' association)
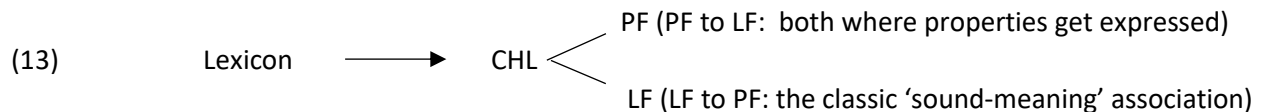
But there is a disconnect between LF and PF (they are two separate modulars of the interface).

(14) Note on C-command. (Also see Domain, C-command as presented in Opening Remarks)

One of Chomsky's axioms is that structures found at LF may not even get pronounced at PF (See G, p. 30):

    a.   Which picture of himself$_t$ does John prefer ___$_{trace}$? : @PF
    b.   (John prefers which picture of himself) : @LF.

    c.   *This picture of himself demonstrates that John is really sick.

(* marks ungrammaticality)

To explain why (c) is ungrammatical, even though the same surface-structure items <John, himself> are grammatical in (a), we must turn to hierarchical C-Command (see G, p. 50).

Answer: There is movement of a) not c).

(a') Which picture of himself does John prefer ~~which picture of himself~~ ?

Reconstruction at PF:
a)   Shows movement so that the base-generated 'trace'-phrase keeps binding at LF

b)   C-Command holds at LF (unpronounced at PF).

(15) Let's flesh this out more fully, consider (see C, p. 6):
a)   <u>John </u>said that <u>he</u> was happy.   (where underlined words coreference).
b)   *<u>He </u>said that <u>John</u> was happy.

(Notice in (b) that the pronoun (he) is force to coreference with the name (John) within the same domain (banned by the Pronoun coreference theory of Lasnik 1976 (Principle of Non-coreference 9see c. p. 7)). This goes against Lasnik's theory that pronouns must be Free within the same domain (the same bracket) such that John & He can't be the same person within the domain—viz., NB. everything that follows the pronoun creates its own bracketed-Domain).

(i)   *[<u>He</u> said that <u>John </u>was happy] (single domain= [He…..John]): Not OK.
(ii)   <u>John</u> said that [<u>He</u> was happy] : single domain = [John said that [He…]] OK
           Pronoun 'He' is Free (it can refer to John or to any other John).

Recap: Why is coreference between name (John) and pronoun (he) possible in a), but not b)? How might a child dela with this data?

Also note that there are plenty of examples as found in the input which would predict that such positioning of b) would be permitted (hence, the fact that children have tacit knowledge

that b) is not allowed proves that <u>children go beyond their input in processing underling structure (at LF), despite possible PF similarities</u> of e.g., adjacency, positioning of items within the string, etc. Often this is what's behind the idea that 'children move beyond their input' in processing hierarchal language structure (also see 'poverty of input' arguments). (Note: see (19) below).

Here's an example where order is reversed while remaining a grammatical string (He>John):

c)   [When <u>he</u> plays with his children, [<u>John</u> is happy__]]. (note the recursive [[]] structure, leaving trave behind).

(while noting how a flat-[] string in (c') doesn't capture trace theory):

(c') [When <u>he</u> plays with his children <u>John</u> is happy].

(Imagine the string if it were flat-[ ] structure, with no trace operation. This string if it were flat would be ungrammatical, e.g., *[<u>He</u> said <u>John</u> was happy], as we compared (i) vs (ii) above.

Question: How is it that c) is allowed but b) is blocked, despite the same surface-level PF (viz., the position of items within the string)? How do children know this?

For starters, the Faculty of Language (FL), c) provides move-based operations, leaving a trace behind in the LF processing (young children as early as 4 years of age—once the begin to operate move-based schemes in their syntax (a maturational development)—tacitly know this).

c') [When <u>he</u> plays with his children, [<u>John</u> is happy ~~when he plays with his children~~]]

(with wh-movement of wh-clause).

Theory: <u>Pronouns must be free:</u> a pronoun introduces its own domain. Let us say that the 'Domain' of an element A is the phrase which immediately contains A: [Pronoun A ➤ A domain…].
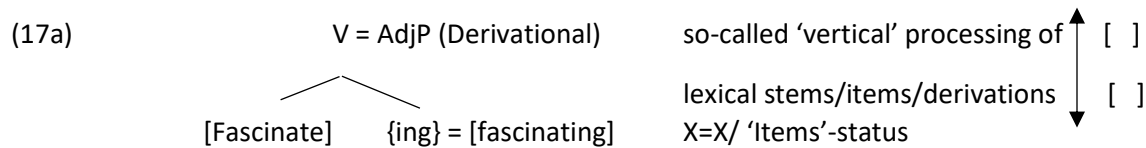
(See also C-command (C, p. 6))

.

1. II.   Genetic/Environmental Factors

(16) Genetic: 'The Brain to Language corollary' (The Dual Mechanism Model (cf. Clahsen, Galasso, Gordon, Marcus, Pinker).

Items vs Categories

(i)   Vertical processes sensitive to 'Frequency of **item**' (e.g., derivational morphology)

(ii)   Horizontal processes sensitive to the 'Rule of **category**' (e.g., inflectional morphology) (See G, 12ff)

Vertical/Item-based: affix {ing} of derivational morphology must search an appropriate stem Verb for attachment—this is 'frequency sensitive' which amounts to word/lexical learning:

(17a)   V = AdjP (Derivational)   so-called 'vertical' processing of ↑ [   ]

                                                                          lexical stems/items/derivations ↓ [   ]

         [Fascinate]   {ing} = [fascinating]   X=X/ 'Items'-status

         'This is a [AdjP [Adj fascinating] [N class]]'

Horizontal/Category-based: the affix {ing} in inflectional morphology is decomposed:

Following the rule (progressive): Be + [verb+ [ing]]   [[   ]] (spreading of rules)

                                                                          ←————————————→

e.g., Mary is [celebrat[ing]] her birthday.   So-called 'horizontal' processing of
                                                                    functional words/categories/INFL
                                                                    X+Y=Z / 'Category'-status

(17b)   V = VP (Inflectional)

         [Celebrate   [{ing}]] = [celebrat [ing]]- remains Verb
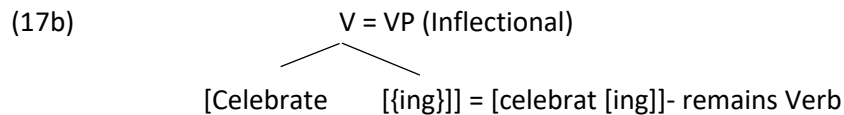
         (See G, p. 12).

What we can do here is assign two modes of processing to the two distinct areas of the brain (namely, vertical goes with temporal lobe X=X processing (Wernicke's area), which is rote-learned and frequency sensitive, while the horizontal (spreading of rules of the X+Y=Z type is attributed to Broca's area/ Left Front Hemisphere). (Item-based is largely *lexical insert* vs Category as *spreading of rules*).

The Dual Mechanism Model of the Brain delivers the idea that there are two fundamentally different areas of the brain (Wernicke's area (Limbic-system/Temporal-Lobe) & Broca's area (Front Left Hemisphere), which reflects that nature of the two aforementioned morphological typologies, along with regular vs irregular word formation, and frequency of item vs abstraction of category storage & processing of language.
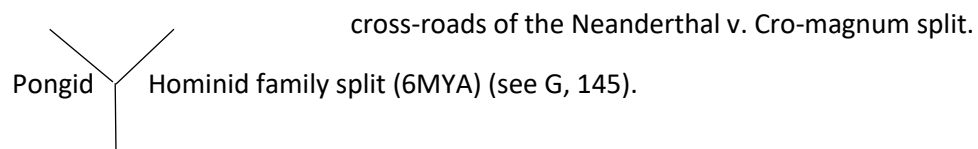
Vertical processing is solely based on brute-force, rote-learning mechanisms which heavily rely of frequency of item, while horizonal is abstract, categorical and rule-based. There are both ontogeny as well as phylogeny relevancy—viz., early hominid (Cro-Magnum) prior to the 40-60Y split with Neanderthal presumably had exclusive rights to horizontal processing, which availed Cro-magnum with such quantitative symbolism which out-paced that of Neanderthal (much to the latter's eventual dismiss).

Discussions/arguments for a Genetic mutation-account for the fundamental difference in these two parts of the brain: FOXP2, Basal Ganglia, etc.).

(18)   Temporal Lobe   Frontal Lobe (FL becomes active at @40-60KYA) at the

cross-roads of the Neanderthal v. Cro-magnum split.

Pongid   Hominid family split (6MYA) (see G, 145).

This split seems highlighted by the Temporal Lobe v Frontal Lobe cut which demarks species separation.

Spin off topics here include: Proto-language, a merge over Move developmental theory of child language acquisition, properties of autisms across the spectrum, animal communication v human language, etc.

This 'Great Leap Forward' takes place @40-60KYA (thousand years ago) leading to Cro-magnum>early modern man's ability to process symbolism (abstraction >categorical language).

A Note on Autism as placed on the spectrum.

This split also between family 'Pongid' (monkey line) and 'Homo' (human line) also creates a spectrum cline of sorts between vertical processing of X =X (represented by pongid processing) vs. horizontal processing of X+Y=Z (characteristic of human categorical thought). This same spectrum, as its name entails, could also overlap with autism spectrum disorders, moving from vertical processing as seen in Asperger's syndrome vs Willliams' syndrome which moves on the cline towards horizontal.

Proto-Language (Bickerton) suggests that this same spectrum can account for early proto-language (Homo erectus) as well as what is accounted for in Pidgin language.

This shift from item-based/X=X memory schemes (consistent of Behaviorism) to variable/categorical X+Y=Z schemes (consistent of Generative Grammar) offers us one of two accounts:

(i) 'function shapes form'—viz., as environmental/software pressures for memory became overloading, (i.e. became to overload the hardware-form/brain), hardware functionality processing had to advance and keep-up develop. This is the software to hardware computer analogy as a way to account for exponential brain growth just prior to this period, say @100KYA. Of course, the flip-side account is also available as a second factor: this second account might explain this growth as taking

(ii) 'form as shaping role of function': i.e., advance in innate hardware/brain capacity pushed the limits of software/functionality, hence 'form defines function'). (See G, paper: https://www.academia.edu/42275126/The_Myth_of_Function_defines_Form_as_the_Null_Biological_Adaptive_Process_and_the_Counter_Linguistics_based_Response_Accumulative_Lecture_and_Topics_for_Research

The most interesting fact about this duality of item vs category is that it maintains the flavor of ontogeny recapitulates phylogeny, at least in two ways: (i) developmentally (ontogeny of a child's language development), and (ii) over the range of species (phylogeny). (Haeckel). Similar brain overlaps are obvious: Limbic system where frequency of item reigns paramount, vs Frontal Lobe where algebraic equations are the order of the day.

For discussion of Brain Analogies, see G. (p. 168). Brain as 'steam engine>clock>computer' (then becoming abstract) brain as 'ghost in the machine' (Descartes). Discussions of 'brain vs mind' likewise pursue—viz., 'the brain bootstraps itself and creates a mind' analogy, and the mind as epiphenomenal.

Poverty of Stimulus.

(19) Environmental factors include (i) The Poverty of Stimulus. (See also G, 4-senteneces). For instance, taking the above pronoun coreference constraint of Lasnik above—namely, again, how/why don't children employ what would be the simpler strategy of adjacency as a driving principle in formulation syntactic structure? As we saw in (15c') above, there's plenty of surface phonology evidence of Pronoun then Name (as counter evidence of Lasnik) but children don't pay attention to such surface phenomena, given that they strictly abide by the principle which states that the Pronoun (He) must be free within its domain, and shall not coreference with a Name (John) it its domain (*He said John was happy), etc.

   a. When [he plays with his children, [John is happy__]].
   b. [John is happy when he plays with his children]
   c. *[He is happy when John plays with his children] (same surface order as (a)).

(The pronoun 'He' in (b) is free to coreference to John (or any other 'John' in the world)).

Children are fine with this sentence despite its surface He + John order, suggesting that children here have access to the movement analogies behind full, formal syntax. How do children 'know' that (a) is fine but (c) is not (given the same surface PF order)? LF must be involved.

Other examples of poverty of stimulus include Peter Gordon's classic 'Rat-eater' experiment (see Gordon 198, cited in G).

Q: 'What do you call a person who eats rats?' (a Move/INFL mother-daughter relation [[Rat]s].
R: a 'Rat-eater'
@R: a 'Rats-eater' (@unattested response) They never say 'Rats-eater'.

In compounding 'rat-eater' the INFL [[ ]s] is deleted forming a sister-siter N+N compound (note how compounds are bricolage in 'internal' nature, build up by [stem+stem] formation). But—in considering the potential environmental, 'external' input—inflections are extremely rare in the input, so how do children gain such knowledge (negative knowledge or otherwise)?. This knowledge must be a byproduct of the [Bricolage…then… [Move]] nature & design of the brain. => Discussion.

Environmental Factors (leading to Faculty of Language/UG)

(20) Input ⟶ LF ⟶ principles: lexical categories/Merge

Parameters: functional categories/Move
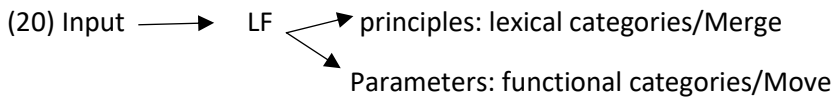
Parameters: Maturational development of LF (UG): phase 1….2…..3….t. (target grammar). E.g.,

+/-Head initial , +/-INFL (Bare Verb Stem), +/- Pro-drop , etc.

*(All parameterized variations the mere result of the binary nature of Internal Merge/Move (IM).

(21)A Note on Merge/Move: (see also left vs right branching languages re. IM).
    (i)   External Merge = combine of two items from the lexical array (a+b)

```
        Y
       / \
     [a    b]
```
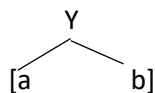
    (ii)   Internal Merge/Move = displacement of an already merged item (a) (up the syntactic tree).

```
          /
    [ a'  / \
       [a     b]]
```
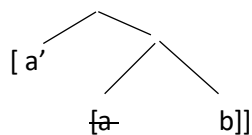
Language is first internal/categorical: cf phonemic perception (e.g., Spanish/English /b/, /p/, /š/, /č/).

Discussion. Environmental gets supplanted by internal factors: How is it possible that two people can hear the same sound differently? Let's assume that environmental is vertical processing while innate categorical is horizontal: this, phonemic representation is horizontal spreading of rule, i.e., horizontal, as opposed to isolated sound chunks on a linear order. Also recall Chomsky's early notion that both (i) creativity, and (ii) errors (omission and commission) are the result not of the environment, but rather of an internal and innate intervening process.

Regarding a strict environmental-based theory of language which suggests that all is to be found on the PF surface level, consider the fact that much of PF becomes vacuous in sound, as in empty categories: When an item moves up as a result of IM, the PF is stripped yet a residual trace of its properties remain at LF.

Empty categories seem to have no PF mechanical/physical reality, yet they are present in the mind.

(Cf. The famous example of the empty category interceding in the 'wanna contraction': (See G, p. 68)).

Note how the PF-form of the sentence: (i) 'He saw the man standing at the bar' has the same LF mental representation as: (ii) 'The man was standing at the bar', such that LF representation = [[He saw] the man standing at the bar].

He saw the man standing at the bar.

The man was standing at the bar.

He saw X, X= the man was standing.

(22) What the above shows is Functional v Formalism'—Language (at PF) seems not to be optimal for functionality:
(i)   Garden-path sentences (G, sentence #3)
(ii)   Ambiguous structures (e.g., a. John saw Mary with a telescope)
          b.   [With a telescope, [John saw Mary_____]](where Move clears-up ambiguity).
               (John saw through a telescope Mary)
          c.   [John saw Mary with a telescope]] (no PP movement):

               (John saw Mary (holding a telescope)

                    i.   Does the PP 'with a telescope' branch off from the DP 'Mary'?
                    ii. Does the PP ;with a telescope' branch off from the Verb Phrase 'saw'?

b.   V'   c. D'

V   PP….with a telescope   D   PP…with a telescope
| |
Saw   Mary

40
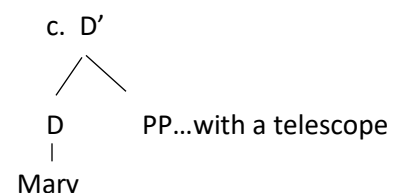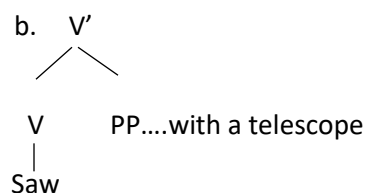
Consider ambiguous compound structures: '

    (i)   'peanut-butter': is butter made from peanuts, but...
    (ii)   'snow-boots': are not boots made from snow...
               (Also see root vs synthetic compounds found in G, p. 20).

    (iii)   Cigarette-smoker (a smoker of cigarettes)
    (iv)   Chain-smoker (Not a smoker of chains).

        Where the former (iii) derives Movement (synthetic) while the latter (iv) employs only merge
        (Root). The same distinction holds of (i, ii). Movement resolves issues of ambiguity.

(23)   **A Note on Displacement/Dislocation.**

Rather, it seems Language is optimal as it serves interface conditions between PF &LF: where (i) PF is mostly an edge-phenomena (where phonological items have moved to Spec-Head of higher projections) and (ii) LF deep structure remains based-generated (such as Theta-marking). The notion that natural Language (L) (and we think all natural languages do this) require displacement suggests that this is the way L can connect items found at the edge (PF) to items not moved (at LF): Then, 'well-designed languages are going to have displacement/dislocation properties' (See C., pp.114ff).

Displacement: Where PF gets displaced-dislocated to edge, while theta-semantics are non-dislocated):

        (24)   Reason for Move:

    PF: edge

                    LF: theta-semantics
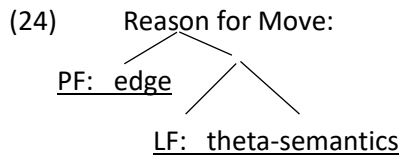
Note that expressions are phonologically interpreted at the edge (of Move) even though they are semantically (thematically) interpreted at the local position (of Merge).

Hence, displacement properties found for L is a way to reconnect Move with Move: 'this is a plausible reason why Languages have this dislocation property'. (C, p. 115). So, there is this disconnect between 'sound and meaning' (far beyond Saussure's classic notion). Hence, Move-based displacement in natural language is just an attempt to remedy the disconnect between edge features related to PF and base-generated features related to theta-roles & semantic mappings.

For example, the displacement of 'bottle' in the phrase 'bottle of wine' vs 'wine bottle' (a merge-based operation) could be seen as the result of bottle' having to take on Genitive INFL properties (up the syntactic tree) which leaves the unmoved bottle' base-generated for other lexical/semantic reasons.

[NP [N wine] [N bottle]] > [GenP Bottle [INFL of [N wine] [~~bottle~~]]] (See G, p. 19).

Since edge-features host displaced items (from lower down in the syntactic tree), functional items/categories become the leading driving of displacement: Move is triggered by Agreement of such features which land via probe-goal search at the edge. (See Miyagawa, 2010, 'Why Agree, Why Move?).

Miyagawa asserts that it is AGR which underwrites language, which drives displacement.

(25)   **A note on Darwin.**

Such displacement in language begs the question of language as 'adaption or exaption' (Steven J. Gould). It seems there is no biological pressure of the Darwinian sort to account for the nature of this displacement as found in natural language (see G., p 28). Chomsky, most recently, has emphasized this matter with his notion of Faculty of Language-broad, vs narrow. (FLb, FLn)—showing that while LFb does indeed have antecedence as found in more general cognitive communication schemes (of the lower primates, etc), FLn on the other hand has no such Darwinist biological account for its onset and development in our species. (See Fitch et al. 2005). Chomsky credits displacement as the exclusive property which shapes and defines Language (all other forms without FLn reduced to communication: iconic, concrete of the sound-meaning association, but if without displacement/Move, the reduced to semantic associations typical of animal communication, etc.).

'The one property which escapes a neo-Darwinian explanation of language is recursive move'. It's good to recall here that [-Interp] features are just that not interpretable (for communication)—i.e., they don't seem to serve communication (Structural Case, Agreement). While the plural on a Noun is surely interpretable (serve communication), the same number feature on its matrix verb seems redundant (the verb agreeing with its subject seems redundant in the case of number). Likewise, if a Determiner is marked +Plural (These), why should we also mark its Noun as plural (These books), surely, as an interpretation, one plural marking across the phase should suffice (These book)…Agreement, in this respect, must be doing another job, unrelated to interpretability.

[-Interp] must be autonomous within its own module as it surely is unrelated to communication—as FLn is unconnected to cognition and other general problem-solving skills.

Speculations abound as to the why/how FLn has arisen uniquely in our species. Perhaps as an artifact of how the brain bootstraps itself and creating a mind….Chomsky asks: No one knows what happens when you take an object the size of a basketball and cram neurons tightly together at a density of about 10^{14}.

Yep! Funny things happen, not the least an array of 'strange looping phenomena….'

This reduces to the Descartes dualism, in a matter of respect: Mind/Brain…

The Nature of Recursive Hierarchy: Topics on Structure

X-bar Theory

Part and parcel of any structural design of language must take into account the following syntactic phenomena. Consider the question/response below):

(26) What are you doing?

    a.   I am going to sleep = Full TP max-projection: TP [TP I [T' am [VP going to sleep]]]
    b.   *_ am going to sleep = Intermediate projection: [T' am [going to sleep]]
    c.   _ _ going to sleep. = Full VP max-projection: [VP going to sleep]
       Let's keep in mind the above discussions about recursive, binary-branching structure here as we address basics of X-bar Theory:

a':   TP   (TP max-projection)

        Spec   T'
         I
             T   VP
            am
                going to sleep

b':   * T' (T-bar intermediate projection: barred from projection)

        T   VP
       am
            going to sleep

c':   VP   (VP max-projection)*

        going to sleep
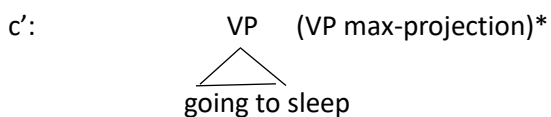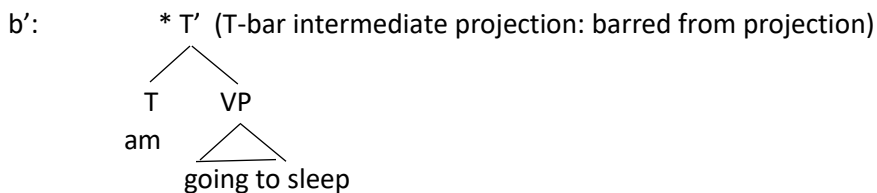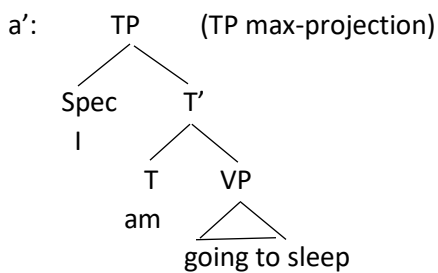
*Fully articulated VP -structure showing infinitive INFL {to} : [VP [going [T' to [V sleep]]]] (where the INFL morpheme {to} being 'featural' inserts as adjunct to the V, keeping a constituent VP.

Clearly if a mere flat merge-based operations on binary branching were operative there would be no way to handle this distinction:

(See G, p. 46-49: for C-command requirements, recursion, binding in local domain).

d'   TP

   I am going to sleep
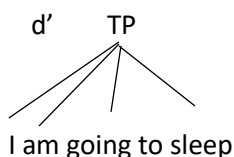
(27)   **Binary Branching: 'sister-sister' 'beads-on-a string' [ ] + [ ] vs. 'mother-daughter' [[ ]] hierarchy**

This same type of binary branching is found in (the syntax of) phonology. Consider the horizontal spreading of Assimilation (a phonological rule): (see G, p. 67). Compare morphosyntax of 'fix' /fɪks/ v speaks /spiks/ as contrasting with phonemes /_rs_/ in 'car' v 'carson'.

Recall, sister relations of are the Merge flavor

[ ]   [ ]
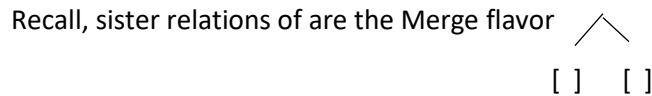
While mother daughter are of the Move flavor:

[

[ ]] ….
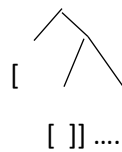
*(Section which follows taken from Preface of 'Reflections on Syntax', Galasso 2021)

Whether it be… constraints placed on *phonological assimilation* which stipulate that in order for the horizontal spreading of voicing to occur between two adjacent consonants, they must first be of a 'sisterhood' relation; whereby, for example, the /r/ in 'cars' provokes assimilation of plural /s/ => /z/, in contrast to the /r/ in 'Carson' which does not. (The former a structural 'sisterhood' relation, the latter a 'mother-daughter' relation in terms of hierarchy):

(28) 'Carson' being broken-up by a *syllabic stress /$/…* /kar $ sƏn/ (CVC $ CVC).

1.   'Ca<u>rs</u>' [K [a [rz]]]   2. 'Ca<u>rs</u>on'

/a/   /r/

/r/   s = /z/ /karz/ (cars)   /s/ = /s/   *(not /z/)
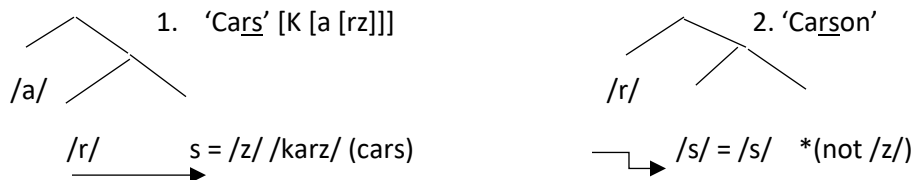
Assimilation of 'voicing 'applies between sisters.   *No assimilation between mother-daughter.
Or whether it be… the naïve view that the two apparently adjacent final-position sounds of /_ks/ as found in the two words 'fix' /fɪks/ versus 'speaks' /spiks/ surely must equally get processed similarly (they do NOT) calls on us to reconsider something much more insidious going on in the underlying structure of morphophonology:

(29) 'Fix' /fɪks/
 1. 'Fix'   2. 'Speaks'

/ɪ/   stem affix

/k/   /s/   [[speak] s]

44
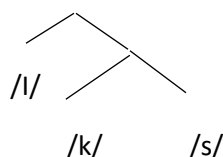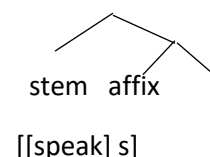
The differences in processing resulting in distinctions held both developmentally (as found in child language) as well as for Second Language (L2) errors of omission: with a lexical stage-1 of child language often deleting the final /_/s/ inflectional affix but never deleting final /_s/ stem-based elements (with similar findings for L2).

And then… to think that such *scaffolding* of hierarchical structure which yields an operative recursive syntax comes to us for 'free'—part-and-parcel of the design of the human brain/mind— is something to wonder. This is what this book is about—'the wonder and unfolding of recursive syntax', and the manner in which it has forced the field of modern-day linguistics to reconsider old assumptions we once held dear—old assumptions which were hard to kill off, but which had to die eventually at the stroke of the *generative grammar enterprise*.

Notice here that simply adjacency (closeness) of phonemes /r_s/ doesn't capture the constraints on assimilation. Take for example the utterances 'John has two cars on the lot', as compared to 'Jonny Carson' Both have the adjacent /rs/ combination at the PF surface level as the two phonemes sit next to each other in the ambient speech stream (i.e., there is no distance, even at the millisecond level, between the distance of 'rs' found in the two environments, yet due to a syllable break in 'Car$son' (but not in Cars), we see assimilation at work: in 'Carson', assimilation is banned, in 'Cars it holds'

    (i)   'There are two [cars] # [on] the lot'. /r s/ (are sister-sister: within same word/syllable)
        ⇨ Assimilation of /rs/ holds /s/= /z/
    (ii)   'Here's Johnny [Carson]' /r s/ are mother-daughter: across word boundary).
        ⇨ Assimilation is barred /s/= /s/

(30)   **Binary branching with Pronouns (Binding). (See C, p. 6)** (Also see sentence in §I (6) above).
        Recall how Pronouns also dependent on structure—viz., how pronouns generate embeddedness in order to create a domain (anything that follows the pronoun (He, She) gets defined as the its ['domain']. (see §1.2 (15)):
         i.   [John$_j$ said that [he$_{j/k}$ was happy]] = [He was happy]: PRN 'He' heads domain.
         ii. *[He $_j$ said that John $_j$ was happy] = PRN 'He' must be free within its domain.

        In (i) above, 'John' can coreference 'He' (as John himself) or any other He (person/John $_k$) since 'He is free'

In (ii) given the Lasnik principle that: a) pronouns create their own domain, and b) that pronouns must be free within their own domain, the only way to handle such binding phenomena is to generate mother-daughter embedded, nested structures [[ ]], ruling out (ii) and explaining (i) above.

(i')   [John..   (ii')   *   [He…   John….]

           [he…]]

As was seen in X-bar theory, flat structures and free binding results are wrong.

*[John wanted [Mary to dress himself]] is wrong since the domain of Mary/Her would be forced to coreference with John/himself outside of the domain. Here, it's the coreference/binding of the reflexive pronouns that captures this domain constraint.

The same mother-daughter relation shows up with the example:

a.   [These stories of you [*'ve/have been going around]

Noting that the clitic 've (have) cannot attach to 'you' (*you've) given it's a mother-daughter nested embedded structure, while a sister-sister structure allows clitic attachment:

b.   [You've/have been going around]

Clitic formations (a PF factor) is also influenced by hierarchical structure: What we can say is that 'have' in (a) doesn't maintain a sister-sister hierarchy to 'You' (but rather to 'These stories') but that in (b) it is. (See G., p 48ff).

Let's consider how a possible extension of sister-sister/mother-daughter relations might be one additional way to explain clitic attachment at PF:

(31)   PF Mother-daughter: Clitic NOT OK: PF Sister-relation: Clitic is OK
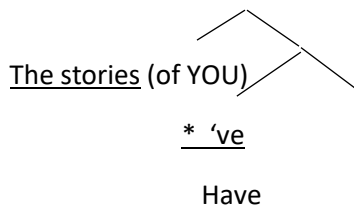
The stories (of YOU)   YOU   've

           * 've   have

          Have

*(Note: The DP 'These stories' doesn't allow Head cliticization)   cliticization (clitics seem to only take pronoun Specifiers: e.g, 'These stories *'ve/have been good').

```
                    TP =S(entence) (showing subject-verb agreement requirement)
          DP                                        T'
      DP      D'                              T          VP
    The stories D    N                    have  circulated around town.
               |     |
               of   you                        *'ve
```

(i)   Clitic {'ve} can't attach to closest stem {you} at PF since LF c-commands doesn't hold. (Syntax at deep level undercuts any surface-level phonological adjacency. (It seems there is tension here between PF and LF— Phonological cliticization is at the mercy of deep level syntax).

(ii)   'These stories' can't attach clitic {'ve} due to the DP 'These stories' not being adjacent at PF (a PF constraint on cliticization). Hence, {have} must be in its full PF form. *(Note: The DP 'These stories' doesn't allow Head cliticization (clitics seem to only take pronoun Specifiers: e.g, 'These stories *'ve/have been good').

(Following examples taken from G, pp47ff)

(32) T**he**se stories of you **have** been going around.

(i)
```
                    TP
          DP
      D        NP                    T' => (T-Head 'have' not c-commanded by 'you')
    These   N     PP          T        PROGP (progressive)
          Stories P    Prn   have
                 of   you      PROG        VP
                              been    going around.
```
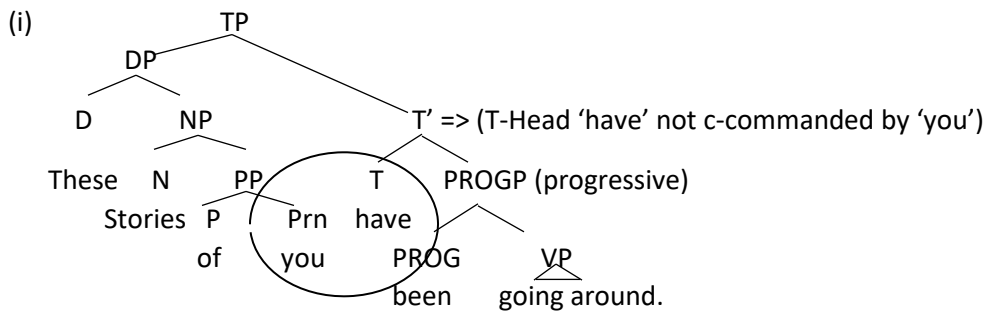
*(Note no clitic usage: 'These stories of *you**'ve** been going around').

(ii)
```
                    TP
          DP
      D        NP                    T'=> (clitic *{'ve} is illicit: non c-command relation)
    These    N     PP         T        PROGP (progressive)
          Stores P   Prn  *'ve
                 of  you      PROG      VP
                             been    going around.
```
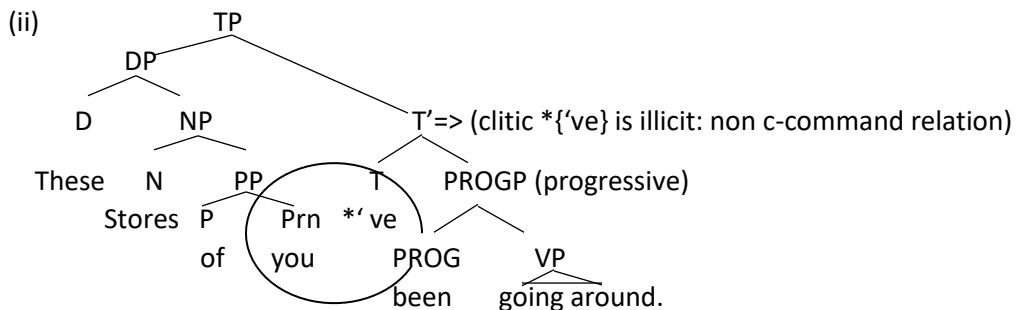
⇨   Hence, while LF maintains {These stories + have} as subject-verb agreement (presumably for a possible clitic {'ve} attachment), non-adjacency at PF bans it ('These stories and 've are mother-daughter relations).

(32a) You**'ve**/have been going around.

(iii)
```
              TP
        Subj      T'  => (clitic {'ve} is allowed: appropriate c-command of Prn 'you)
       You     T     ProgP
              've   Prog    VP
                    been   going around
```
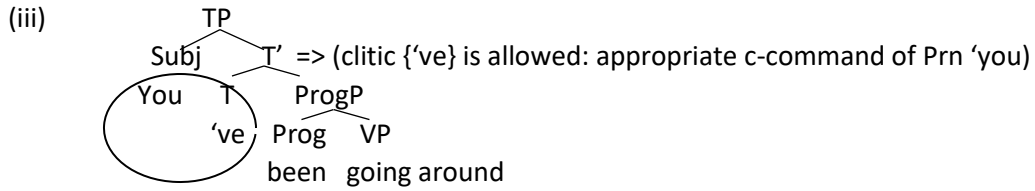
⇨ Here in (iii) we can say that adjacency holds between You and 've both at PF & LF (due to adjacency & c-command respectively).

(32b) This story of you ***have/has** been going around.
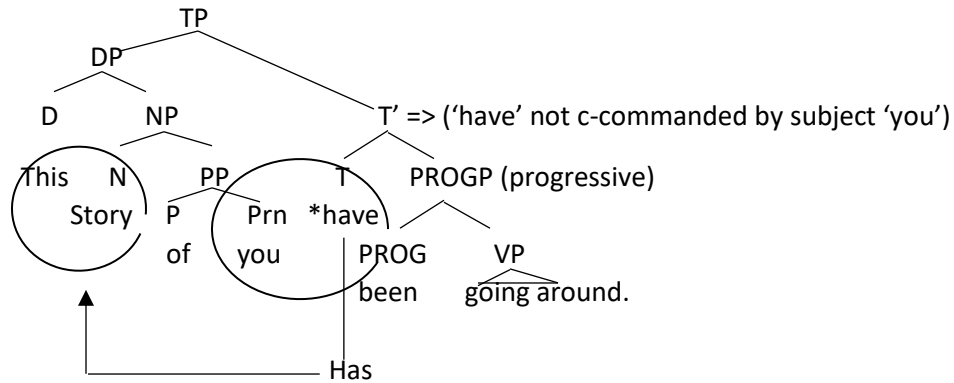
```
                   TP
        DP
      D     NP                 T' => ('have' not c-commanded by subject 'you')
    This    N    PP        T      PROGP (progressive)
          Story  P    Prn  *have
                 of   you       PROG       VP
                                been    going around.

                         Has
```

i.   This story   of [*you have]…
ii. [This story] of   you   [has]…

Note that the verb must be spelled out as 'has' [3 person/singular] since the c-command relation is with the subject 'This story' [3P, sing] and not with the adjacent previous word 'you' (as would be found in the superficial linear word order).

[32c] These friends of the president blame *himself/themselves.

TP   => 'themselves' is c-commanded by 'These friends')
DP

D   NP   T'=> ('himself' not c-commanded by Prn 'the president')

These   N   PP   T   VP
Friends P   Prn
of the president   V   prn/reflex
blame   *himself
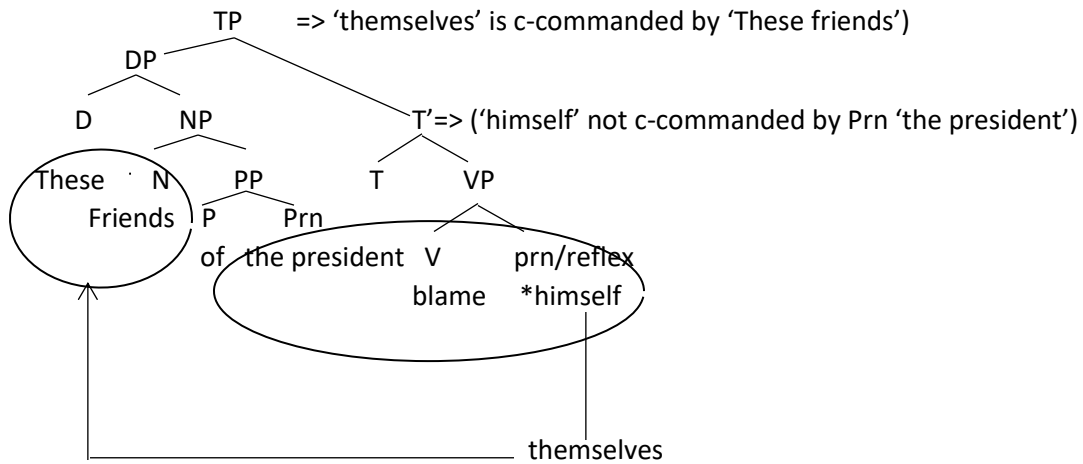
themselves

i. Friends of the [president] blame *[himself].
ii. Friends] of the president blame [themselves].
iii.   The president] blames [himself].
iv.   [These friends] blame [themselves].

In example (i), the pronoun 'the president' is not in a structural relation which would allow c-command to take place (which is required of reflexive pronouns), it being a sister of the preposition 'of' and a daughter of PP.

(32d)   TP

DP   T'   => correct c-command relation: DP c-commands Prn Reflexive
T   VP
a. The president   V   Prn-Reflex
b. These friends   a. blames

b. blame   a. himself
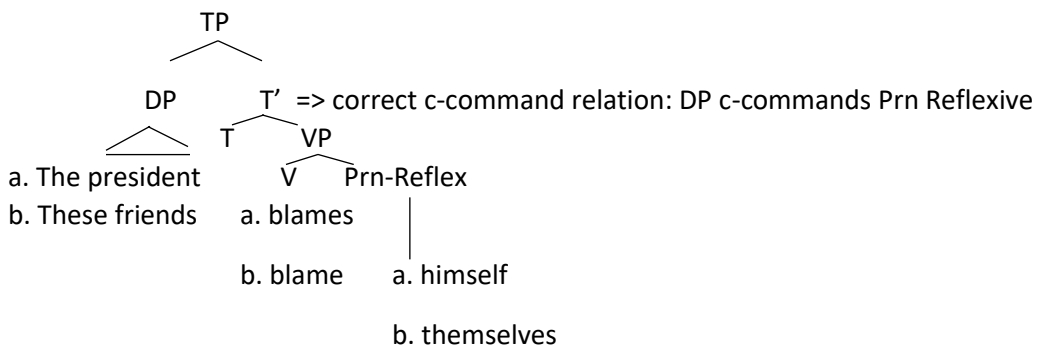
b. themselves

Note how it's the Prn-reflexive '*themselves*' of the VP which must be bound (via **anaphoric binding**) by an appropriate **antecedent**— (the subject (DP/NP) '*The president/These friends*' are the antecedents of the anaphor 'himself/themselves' (respectively). In sum: The two items (the **anaphor** and **antecedent**) must fall within a structural **c-command** relation).

    Below in [33], let's spell out exactly what c-command looks like incorporating terms such as sister, daughter, mother relations found in a family tree.

[33]　　　　　　　**C-Command**　　　　　　　**Specifier>Head>Complement**

```
              A                                    XP
           ┌──┴──┐                              ┌──┴──┐
           B     E                              y      X' (x-bar)
         ┌─┴─┐  ┌─┴─┐                          Spec   ┌──┴──┐
         C   D  F   G                                 X     Z
            ┌┴┐                                      Head   Comp
            H  J
```

                                                      $\langle x, z \rangle$ = sister-sister/flat

                                                      $\langle y < x \rangle\rangle$ = mother-daughter/recursive
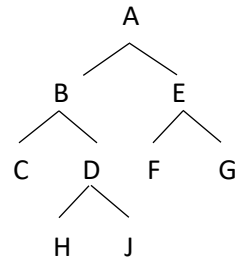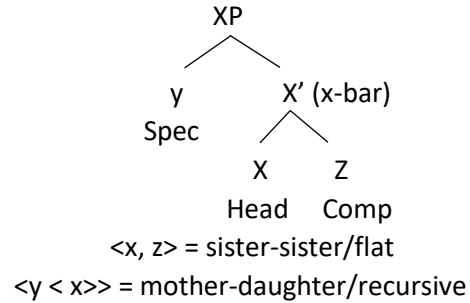
C-command, as a 'linguistic description', is a way of showing just how recursiveness acts upon syntactic constituents, which in turn, brings about hierarchical relations. C-command in this way is described as a syntactic operation found at (Syntax) (prior to PF, LF interfaces). A is the mother of daughters B, E (which themselves are sisters). Sisters can (symmetric) c-command each other (both ways). Mothers can (asymmetric) c-command daughters (one way).

    As observed in the illicit nature of the clitic *[you've] in (32ii) above, the rationale behind the ungrammatical status has to do with the fact that (J) can't c-command (F) (they are not sisters), given the structure in [32] that the pronoun *you* would project from a hierarchical c-commanding position similar to (J), while the clitic ['ve] would be projecting from a similar position found in (F) (as referenced to the scheme in [33] above). While (F) can be *symmetrically* c-commanded by sister (G) {G>F, F>G}, or *asymmetrically* c-commanded by Mother (E) {E>F}, there can be no c-command relation either way between (J) </>(F). And so, what we find in (32ii) is the simple fact that 'you' (J) cannot c-command {'ve} (F), hence the ungrammatical status of (31), while (32a) [you've] is permitted since (B) 'you' asymmetrically c-commands (F) ['ve] as correlated to the scheme referenced in (33).

    So, in sum, what we have here is a two-prong combination of Merge, a combination which functions in one of two ways:
        (i) Merge (local/adjacent) as sisters {x, z} when projected in a flat, non-recursive manner, or as,
        (ii) Merge/Move (distant/non-adjacent) as mother daughter {y {x,z}} when recursive.

    It is this latter 'distant' function of Merge (ii) which MP relabels as **Move**.

1. III Seek out its Functionality

Why Move? Why all this redundant, formalism in natural language? What role if any does formal [-Interpretable] syntax play a role in functional communication? Why Agree? Why Move? (cf. Miyagawa, 2010), Why so much built-in redundancy?

What is language, What is its properties, and How should we go about defining it as a formal computational system? It has been too often considered, I feel mistakenly so, that the study of language and linguistics should be narrowly treated in ways which solely emphasizes the role language plays in providing **communication.** Sure, the role language plays in communication is real enough, and it may do us little good as linguists to talk about language absent its **functional value** as a means of communication (a kind of **functionalism**). But we must be very clear here on this point: while communication certainly involves, indeed requires, at least some fragmentary aspect of language, what we wish to do here is extend the formal definition of language, move it beyond its mere functional value, and attribute to it features which have to do with pure **computational qualities** (a kind of **formalism**). These same formal qualities we attribute to language are thus present even in our silent, inner-most private language, independent of whether or not we are actually communicating to anyone in a public space. True, the counterpart qualities and properties of such a private *language-of-thought* will need to take-on very different shapes when transferring over to a public *language-of-communication*—this is simply due to the fact that there will be differences in demands placed upon the two systems.

For instance, it's plain enough to see that a 'sound-based' (public) language-of-communication would need to satisfy demands placed on it by constraints on the cogno-auditory system. The mere fact that a public linguistic expression has to be uttered →heard →processed presents linguists with some fairly non-trivial demands, both physical as well as mental, requirements of which must be fully met. There certainly would not be the same types of demands placed on, say, a *public* language-of-thought (if one were to exist, other than, say, demands which might creep into processes of telepathy: I don't know what such demands would look like). Of course, what we mean here, and what does exist, is a *private* language-of-thought (not public) which counters our *public* language-of-communication. In any event, we can get by with this definition of a public language-of-communication, in contrast to its counterpart private language-of-thought, by addressing the kinds of demands which must be satisfied in order for either linguistic system as a whole to remain cohesive and coherent.

It may be surprising to most people that it is this latter *private* and more formal aspect of language, labeled herein as the **formal value** of language (**formalism**) that has become the touchstone and leading focus behind the Chomskyan framework of linguistic theory. (Noam Chomsky is an MIT scholar and is considered by many to be the seminal figure in modern-day linguistics). By adhering to these inherent computational features of language—all the while paying our due respect to how such features must manifest and project in communication—we keep to our well-defined notion that language is formal, abstract in nature, and structurally dependent—i.e., that language is always more than 'just the sum of its parts'. So, we must be careful and define language—and the processes which underwrite language (i.e., its computational systems)—in ways which address this higher-level quality.

Holding this more stringent definition of language, we can begin to move beyond language in its communicative capacity and accept, in Chomskyan theoretical terms, that language doesn't (nor should it) simply reduce to a mere functional status. In fact, we may wish to de-emphasize and forego the role of communication all together, and by doing so, strictly turn our attention to evaluating what these more abstract and formal **qualitative values** of language really are. This is a kind of shedding of the outer clothing (language-of-communication), to a reaching-down to the bare skeletal properties (language-of-thought). In so doing, questions regarding the true formal nature of what underwrites language emerge in one fell swoop, forcing us to accept that fact that there could be language without the uttered word, and thought without the spoken sentence.

But does this mean we sell ourselves short of **word, phrase, and sentence?** No! In fact, these very constituents will be rightly called upon as the outer (objective) manifestation of our inner (subjective) language. As mentioned, these constituents serve as satisfying certain demands; they roughly serve as a cognitive bridge which carries the formal inner *impression of thought* to the outer *expression of communication.* So, what we mean is that in addition to the **substantive** *word, phrase, sentence*, there must also be something of a more formal **non-substantive** nature which governs and underwrites the processes of word, phrase, and sentence. The aim of this introductory text is to draw attention to these non-substantive governing processes which allow language to emerge.

Question—What kind of a model of language (*b*), labelled here as the **Language Faculty** (LF), could deliver such anomalous structures as made apparent by contrasting the language (L) input (*a*) *Lxy* with the L outputs (*c, d*), with (*c*) showing **omission** of L input (mistakes of subtraction), and (*d*) showing errors of **commission** (mistakes of addition)?

At the very minimum, LF must serve in some capacity as an intervening computational system, which, as the expression goes, 'has a mind of its own'.

(34)   *a Lxy* input   *b* LF   *c Lx* output (omission)

d Lxyz output (commission)
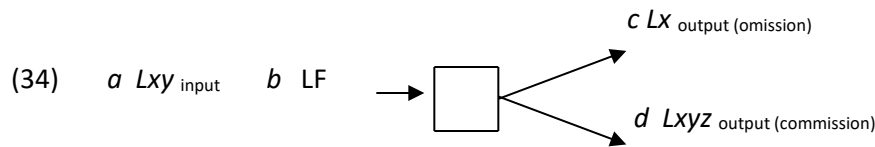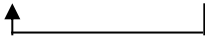
(Note: Given what's been said in our little background found in the previous page, keep in mind what kind of language model would predict such a highly systematic treatment of language regarding what suffers omission or commission in child language development).

Consider (35) below showing one such example of functional omission.

(35) Lx output (c): omission of Determiner, Auxiliary

    a.   *input structure*:

                    What (is) (the) man (is) doing?

                        → What is the man doing?   (target structure)

    b. *output utterance*:

                  What Ø Ø   man Ø doing?

                      → What man doing?   (child utterance)

    c. *Compare*:

                What is the man doing?   (adult/target utterance)

                What   man doing?   (child utterance)

Note that both the Determiner 'the' and Auxiliary 'is' gets deleted in the child language target. These two categories (D & AUX) and considered functional categories.

Next in (36), consider the omission of more subtle functional morpho-syntactic features.

(36) Lx output (c): omission of Case, Tense, Inflection, Number:

    a.   *input structure*:

                  [[3P/Sg/Nom] He] [*IP* [drive] *s*] [[+PL] two] [*Num*[car] *s*]

                → He   drives   two   cars.   (target structure)

    b. *Output utterance:*

                → Him drive-Ø two car-Ø   (child utterance)

    c. *Compare:*

                He drives two cars.   (adult/target utterance)

                Him drive two car.   (child utterance)

(37) Lxyz output (d): commission, copy of Auxiliary

   a. Input structure:

      What is the man (is) doing?

         → What is the man doing?   (target structure)

   b.   output utterance:

      → What is the man is doing?   (child utterance)

   c.   Compare:

      What is the man doing?   (adult/target structure)

      What is the man is doing?   (child utterance)

(38) Lxyz output (d): commission of possessive

   a. input structure:

      This is [[Poss] mine-Ø]

         → This is   mine!   (target structure)

   d.   output utterance:

      This is [[Poss] mine's]

         → This is mine's   (child utterance)

While developmental linguists have long observed that young children systematically omit certain aspects of language over others, and that there is a maturational timetable to these omissions, we are still learning the degree to which some language elements remain more conservative over others, and why there might be cross-linguistic differences found amongst such omissions. Perhaps an even more important question is to ask whether some commonality holds between more vs. less conservative elements found in languages across the world.

Redundancy: Even more problematic than the above developmental cases, consider how *redundancy* works:

Question: Why Agree?

If the determiner is plural (e.g., These), why should the matrix Noun also be marked for plural (These books)? Doesn't the expression 'These book' serve the same communicative aims? The same question can equally apply to subject verb agreement—another redundancy. If the subject is singular (Mary), what necessitates the verb to also be singular (speaks) as in 'Mary speaks' French?

*A Note on Plural as INFL. Plural INFLection on Nouns are extremely natural (e.g., [[book]s]...namely, the fact that the {s} attaches to the Noun/item is a good design.

So, in fact the plurality on nouns is rather like different words: just as you have 'table', 'chair', you have 'plural' (and in fact plural as a new word/item does show up in irregulars (e.g., children, feet, etc.). But herein lies a critical difference: while *table, chair, nightstand* must be *items* (a table is not a chair), plural INFL can't be an item (INFL must only overextend as a categorical/general rule which spreads and operates on separate items).

Category: Everything has to be 'singular or plural'...
Item: But Not Everything has to be a 'chair'. Hence, 'table vs chair' must be a recognizable item-base distinction, while plural INFL only must cover a category. This is a perfection on Number on Noun. But why the same number distinction on the verb (re. verb-subject agreement). This spreading of INFL seems redundant (C., p111).

Chomsky comes to note that while number on Noun is [+Interpretable] (i.e., meaning-based, LF interpretable), the number on subject-verb agreement is [-Interp] (not LF interpretable).

If Determiner is Plural, why also the Noun? 'These books'

If Noun is singular, why also the verb? The boy [[speak]s] French.

(39) **Example of redundancy:**

'I am a friend of John's'

i.	[John's friend]	(step-1 recursion)
ii.	[of [John's friend]]	(intermediate step)
iii.	[a friend [of [John's ~~friend~~]]]	(step-2 recursion)
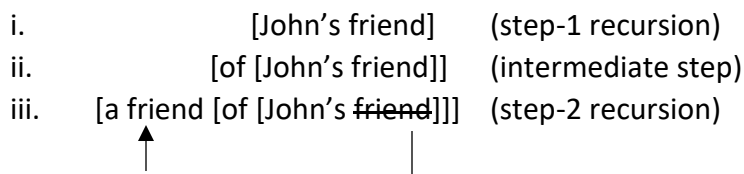
It would seem that the functionality once served and utilized (Possessive), should the get deleted in the deviation. This is not what happens—e.g., the possessive {s} remains at PF (displaying a double possessive structure).

The Pongid/Hominid split, some 6MYA (million years ago), would ultimately be rendezvous-1 with any designs calling for a putative Proto-language which tethers human-like speech to brain/gene evolution (e.g., Broca's, FoxP2). The earliest such attempts at a proto-language/speech would have to fall somewhere between 1-2 million years ago, with the emergence of Homo Erectus. In this respect, scientists look to fossilized hard-bone extractions of early homo skeletal remains, such as cranium imprint of protruding forehead, Broca/front left hemisphere indentation, vocal tract lowering of larynx, jaw-muscle strapping of mouth, etc. The main findings suggest that while some of these critical features which support language do emerge in homo erectus, the question rather becomes one of narrowly defining 'language as formalism' (and not of the mere aforementioned early 'physical' attributes as found in the fossil record—viz., did early homo have a recursive syntax? The literature in this regard pegs the emergence of formal recursive syntax to somewhere between 40, 60 KYA (thousand years ago), a mere drop of water in the evolutionary oceans of time (See Diamond, 1992). The subsequent split between Neanderthal and Cro-magnum highlights this distinction: the two subspecies certainly were very similar to one another, both used tools, cooked with fire, enjoyed social aspects (though Neanderthals may have enjoyed all these to a lesser degree), the main distinction that seems to highlight the separation of the two amounts to the *formalism* of syntax, with Cro-magnum emerging as the champion of symbolic, recursive processing (which may have led to the demise of Neanderthal, notwithstanding some amount of interbreeding between the two sub-species). We, modern man, are Cro-magnum in full glory.

The Minimalist Program advances the argument that there must be critical distinctions upheld between L(anguage) and C(ommunication)—the former (L) being *narrowly* defined and supported by a formalism which takes **recursive syntax** to be of the utmost importance, while the latter (C) as defined in a more *broad* manner, e.g., encompassing a lower-level cognitive motor-control apparatus (bird-calls, animal communication, brute-force memory of rote learning supported by stimulus & response measures, etc.). This distinction makes its way into MP by the dual definitions of Language Faculty-narrow (LFn) vs LF-broad (LFb) (see Fitch et al. 2005). *Symbolism* served as the great leap forward to our species.

LFn = recursive theory : {a, {a, b}} (Perhaps connected to specific cascading genes (FOXP2), Basil Ganglia).

LFb = Beads-on-a-string theory: {a} + {b} + {c}… (cognitive motor-control/general problem-solving skills).

RECURSIVE SYNTAX (taken from Galasso 2021, Preface).

Some 40-60,000 years ago, a monumental clash between two hominids occurred, whereby an older species, Neanderthals, confronted a newly-emergent species called Cro-magnon, whose very existence was drenched in *symbols* (something that had never been seen before on the planet) (*Randall White*, 1989). Both had rather archaic body-types (the former, a slightly larger brain-size, the latter much more boney in frame), but the determining factor as to who would gain the advantage can be readily reduced and expressed in modern-day computer terminology—namely, of how a battle (between the two species) might have been played-out between antiquated hard-ware versus advanced soft-ware. The rather

strange and human-species-specific ability to form *hierarchical recursive syntax*, as exclusively found in human language (sparked by Cro-magnum), would be that unique soft-ware advantage.

At the very conception of Greco-Roman philosophy, we already find Plato convincing us that immaterial essences merely consist of *Forms*, which contain the true and ultimate realities, while the world of *sensible things* is only a vague transitory and untrustworthy copy. (One has an infinitely better grasp of an *abstract & subjective* quality of, say, the color *orange* than any *real* 3-dimensional *object*— 'The former, we have a firm mental grasp of its full essence, its outer covering, its inner trappings, its corners and edges; all is in total mental sight due to its abstract quality. While the latter so-called 'concrete' entity is elusive, an entity upon which one could never fully size-up from mere empirical observation alone, given that human vision cannot process sight in one fell-swoop-panoramic 360° view— there will always be the odd hidden corner or edge which is covered just behind what appears on the surface in front of us, escaping our sight, (mysteriously hidden like the proverbial dark-side of the moon)). (*José Ortega y Gasset*).

A 19th-20th century cognitive perspective would come to show how cued-representations (Icons) could only represent an individual or item, while a detached symbol could stand for an item without the unnecessarily burdening requirement of external stimulation—the former being triggered by direct, environmental stimuli, the latter by a delayed response of its memory—The sign that once expresses an idea will serve to recall it in future. (*George Santayana*).

In the latter part of 20th century linguistics, we would find *Noam Chomsky*, alone, dwelling on these observations, dreaming of a return to a Cartesian Linguistics—particularly thinking, that in order for grammatical syntax to take on its full-operational quality, something must happen whereby a smaller structure gains the ability to lodge itself within a larger structure, all the while preserving its structural integrity with no information lost. Again, the 'item inside a category' analogy would be a critical component to the theory.

Today, on a more personal note—taking-on these concepts as a unifying framework, (even as a 'pedagogical device')—I find that in my own lectures I often resort to extending the metaphor of 'tables, chairs, and night-stands vs. furniture', and show just how, in analogy, the former items (cued-representations) stand in direct opposition to the latter 'furniture' as a category: [furniture [table, chair, night-stand]],

…where we can analyze recurrent items expressed as [α, β], and recursive/embedded structure as [α, [α, β]]—with recurrent flat items forming a lexicon, while hierarchical recursion forms a syntax. (See 'Recursion' definition below).

So, if Neanderthal were indeed stuck in a *flat* world whereby only direct items could be linearly expressed, or possibly stacked on top of each other lexically, (of the 'table-chair-and- night-stand' variety), then, there would be no doubt in my mind how Cro-magnon, drenched in their detached symbolic categories (furniture), would eventually come to outpace their rivals. Such a cognitive thought-process which allows its host to see the world via categorization would undoubtedly lead to the displacement of the earlier species. Modern Homo-sapiens have made the most out of this cognitive niche, up until today (and with artificial intelligence fast approaching—for better or for worse (though I have some thoughts on the topic)).

2. Duality of Semantics

This dual distinction of 'item vs category' (as shown above) parallels what we find in semantics: while some things must signal meaning/semantics [+Interp], other things seem not to denote meaning, and are rather reduced to formal abstraction. This dual nature of semantics (vs syntax) is referred to as the Duality of Semantics'. (Also see (17a) on item-based 'Fascinating' types vs category-based 'Celebrating' types).

(40) **Movement Classification:** (section taken from 'Reflections on syntax', p. 53)

- (i)   Merge
- (ii)   Move
    - a. Local (semantic-thematic)
    - b. Distant (syntactic)[1]

- (iii)   Dual Probe-Goal relations: (cf. Miyagawa)
    - i.   Local VP/Case-Theta
    - ii. Distant CP/AGReement

- (iv)   CP hands-over AGR to T, T hands-over Case to vP: (R, p. 398)
    - a. Features percolate down from C onto T, from v onto V (= a Dual Probe-Goal relation-cf. Miyagawa), 409 (Probe-Goal: CP= AGR, T= Tense), 287 (+/- Interp features):
        - e.g., 'What's the wheels doing?' (p. 398, Ex. 35): where 'is' agrees with wh-pronoun 'what'. (In this case, C keeps AGR-feature rather than handing them over and percolating them down to T.)

Miyagawa assumes that the duality forms a split in the Probe-Goal Union (PGU):

- i.   with *local* PGU relegated to THETA-marking and CASE (two VP-internal configuration, since Case, particularly *inherent* case is tied to the lexical item within the VP, etc.),
- ii.   with *distant* PGU seeking out distant higher Functional (F)-projection relations concerning AGReement.

We view **Merge** as *bricolage* in nature, meaning that it is a (physical) 'lexical builder' at the very local phrase-level. Merge, simply put, creates the **phrase**—a binary operation which is the essential property

---

[1] This semantic/syntactic cut (sometimes referred to as the 'Duality of Semantics' will be further expanded in our discussions culminating with the view that there exist two probe-goal relations which trigger movement: one more 'local' (checking off of Case/semantic, light vP) and the other more 'distant' (checking off of AGReement/syntactic, CP).

of language design. The role of **Move** is then to extend the phrase up the syntactic tree with movement being motivated by the need to check-off and erase formal (abstract) features on the lexicon which have (for some reason) entered into the human computational system (say, at the level of Phonological Form/PF) but which cannot seek semantic interpretation (at the level of Logical Form/LF). So, in a sense, Move creates upward mobility and extends the phrase into higher functional projections. Move seems to be part of a language design which seeks to create recursive embedding and nesting formations—'an exclusive property of human language and that which separates human language from all other communicative systems' (e.g., Hauser, Chomsky, Fitch 2002). Move is a *displacement* property of language which allows an item to be heard in one area of the phonology while being interpreted in another area of syntax.

For example, in passive structures [*Mary was kissed by* [*John* __ __]] the active subject/agent 'John' is dislocated in the surface phonology (PF) from where/how it gets interpreted in the underlying semantic structure (LF) [John <u>kissed</u> <u>Mary</u>]. Other examples of Movement show-up within AGReement mechanisms, say, between number regarding the verb and the matrix subject—e.g., *John speak*-s *French* where the AGR-affix {s} is a move-based inflectional reflex of subject/verb morphology (3P/Singular/Present tense). The fact that both subject and verb must enter into an agreement and correspondently mark for 3 Person (whereas marking on one of the two items should suffice) shows a high level of redundancy built into the design. The same case could be said of number regarding the DP—e.g. *two book-s* where the plural marker on the Determiner *two* alone should have sufficed without need to also mark plural {s} on the Noun *book*. Such redundant aspects of language seem to arise out of an optimality of language design motivated by movement. If there were no optimal need for movement, human language could have evolved as an essentially flat [ ] design.

When we look closely at Move, we find that it too has a binarity of design: it either involves a more local constraint (*closeness*) which we find in examples such as binding, or it is free to displace as (*distant*) movement in cyclic fashion (step by stem) to far-reaches up the syntactic tree (e.g., Wh-movement). Local vs. Distant Move may also capture what we know of the *Semantic vs. Syntactic Cut* (respectively). We'll come to suggest that the properties of language design, as they slowly emerge in child syntax, (viz., Item>Merge>Move-local>Move-distant>…) is, to a certain degree, a creative attempt to rehash the (long discredited) notion of *ontogeny recapitulates phylogeny* (Ernst Haeckel)—being that the progression of language design is pegged to a maturational development of certain cortical parts of the brain. One could extend Haeckel's analogy here by suggesting that any putative 'proto-language' would most certainly have followed and had remained frozen in one of the intermediate steps, along the same progression as taken by the early child—if for no other reason but that it would be stipulated by the progression of design. (See e.g., Bickerton (1990) for thoughts on Proto-language (Also see appendix-). But also see *Punctuated Equilibrium* as argued by Gould & Eldredge (1972) which calls for some aspects of Darwinian evolution to appear in *sudden burst* as opposed to gradual onsets). In any case, the growth of child syntax, if looked at very myopically, shows a fast-closing window of incremental growth over a small period of time. But if we can slow down the progression just a bit, we might have a glance of slow growth. We believe we have done so here with our data—the bilingual child's longitudinal English grammar has been 'slowed' just enough by the two emerging grammars.

Now, while this monograph is not about brain studies *per se*, nor is it about any attempt to tether emergent child language to proto-language, our main objective is rather that of 'language design'—to draw some light on how current linguistic assumptions within the Chomskyan Minimalist Program (MP)

framework might lend an account for what we see emerge in the *slow growth* of early child syntax. As a case in point, let's just turn to one such finding pointing to the aforementioned dichotomy of *closeness vs. distance*, while keeping an eye how the underlying theoretical assumptions of the two processes might prove valuable for our discourse in the pages and chapters which follow. Consider the structures below (Santi & Grodzinsky 2007):

(41)   John knows that [Mary$_i$ pinched herself$_i$]   → Binding: Local

(42) *[John$_i$ knows that [Mary pinched himself$_i$]]   → Binding: *Distant

Here, it seems to be the case that the antecedent *Mary* must have an adjacent closeness to its reflexive *herself* (whereas the ungrammaticality of (42) is due to the distance of the antecedent *John* falling outside the clause and thus breaking adjacent closeness). We can refer to this processing of closeness as similar to what we find regarding retrieval processing of say lexical items—a brain-to-language mapping traditionally assigned to Wernicke's area/Temporal lobe. We will come to call this a **local move** operation which may be one-step removed from a prosaic merge combination (using MP terminology).

In (43) below, we see a very different underlying processing:

(43) John loves the woman$_i$ [that [david pinched___$_i$]] → Movement: Distant.

(where [___$_i$] indicates an index/copy of the moved item—in this case, that *the woman* has been displaced from out of the [__ [david pinched ~~the woman~~]] clause.

When we compare the 'distance travelled' between the two structures, we quickly find that the example in (41) rather mimics what we find of adjacency lexical conditions–viz., that certain words collocate together—e.g., verbs introduce noun phrases forming a Verb Phrase (VP) (*eat the cake*), or that idioms must remain intact (*John kicked/*knocked-over the bucket/*pale = died*) and that very few exceptions allow much space to separate the constituency (and in the former idiom case, no exceptions are allowed at all). It seems under this condition, when too much space separates the constituency there is often a breakdown in processing (e.g., **John ate very quietly in the middle of the room after dinner <u>the cake</u>*). Or when such a breach of closeness is allowed, it is to compensate for higher-order pragmatics: e.g., *John ate __ very quickly <u>the dinner</u> that was prepared for him by Mary* seems to be fine since *the dinner* can be seen as moving rightward (Heavy NP shift) in order to pragmatically coordinate with the lower phrase *prepared for him* (the dinner).

To a large degree, and in overly-simplistic terms, phrase-structure rules as well as lexical retrieval processing come down to this closeness condition e.g., [[*pushed-him*] *down*], but not [*pushed__* [*down - him*]]. Contrastingly, (though still of a systematic nature) the above structure we find in (43), as opposed to (41), pays little heed to closeness and rather shows licit long-distance movement (moving over several

60

phases at a time (though presumably in cyclic stepwise fashion). Let's take a closer look at the two movements, one which is deemed to remain within a single phase (binding, intra-phasal) and one which crosses phasal boundaries (distant/inter-phasal) (Chomsky 2001 defines phases as CP and vP). The notion of pegging movement to phase will become a pivotal aspect of our developing theory on **Merge over Move**:

(44) John knows that [vP Mary pinched herself]

→ Binding: Local/Intra-phasal

= Mary$_j$ pinched Mary$_j$.
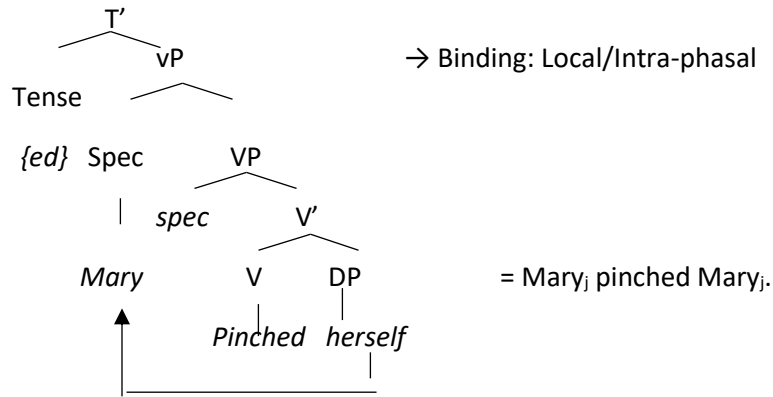
(45) John loves [the woman that [David pinched___]].

→ Movement: Distant/Inter-phasal

= David pinched the woman
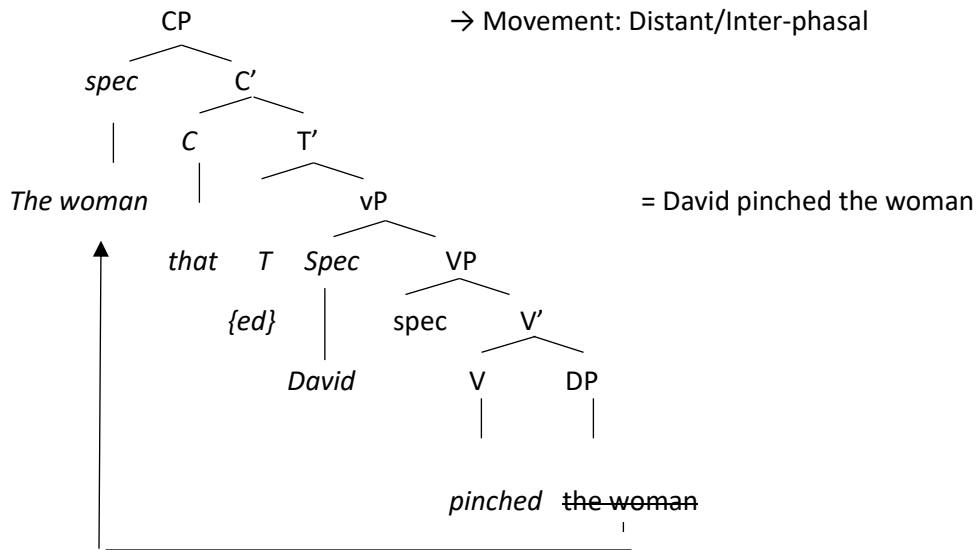
Notice in (44) how the canonical SVO ordering remains intact:

Mary-$_{Subj}$ pinched-$_V$ herself-$_{Obj}$

…but how (45) breaks canonical word order by object raising (OSV_)

The woman-$_{Obj}$ david-$_{Subj}$ pinched-$_V$ __

61

Probe-Goal Union/Relation (PGU)

It seems that **binding** is captured by a local processing operation and that **syntactic movement** is captured by long distance processing. If so, we might expect to draw some parallels with other aspects of syntax which show similar structural distinctions. For instance, we will consider that the **Probe-Goal relation** (using MP terminology) for the checking-off of features will follow a dual-mechanism route:

(i)   Probe-Goal relations of a case/thematic/semantic nature secured by local movement (e.g., vP—handling *case & argument structure*), and

(ii)   Probe-Goal relations of a syntactic nature secured by distant movement (e.g., CP—handling *expression structure*). In addition to these two important phrases (known has *phases* since material from each phase must be independently transferred to PF and LF interpretation), we will consider how the two movements work together informing a cohesive syntactic structure. Any putative lack of distant movement at early stages of child syntax must surely impact the speech of a child.

In sum, the nature of these twin-types of movement represents the classic **Duality of Semantics**—namely, the separation between lexical and functional heads. We assume throughout (following e.g., Miyagawa's (2010) assessment) the following demarcation of 'heads' (or, using more recent MP terminology, 'probes'), that:

a. Lexical heads/probes deal with the *semantic-thematic/argument* structure of language (and is configured via local merge operations), while,

b. Functional heads/probes deal with the syntactic *expressive* structure.

Case & Agreement

We further assume (following much consensus within the MP) that the quasi-lexical projection *light verb* vP assigns a **probe-goal** relation whereby **Case** gets assigned (e.g., as we see inherent Case being lexically assigned), while the quintessential functional projection CP assigns **Agreement.** Hence, as part of our developing story of 'Merge over Move', we find a similar overlap regarding the duality of semantics: with Merge catering more to the lexical/semantic side and Move catering more to the functional/syntactic side of the demarcation. The goal of this monograph is to present a hypothesis regarding the growth of syntax along the lines of this classic split found within the duality of semantics. Having said this, we'll come to consider the following sequence and family of movement (below):

(46) Family of Movement
        a. **basic merge** sequences (pulling two items from out of the lexicon), then,
        b. to **local move** (forming a semantic/argument-structure hierarchical phrase), then finally,
        c. to **distant move** (forming syntactic agreement relations).

The distinction between (b) and (c) is tantamount to what was traditionally called the lexical vs. functional categorical split (respectively) with (a) being the *bricolage* building-blocks of phrase formation. In addition, to these three points, we will also assume that it is **Agreement (CP)**, the quintessential motor driving functional categories, which motivates inter-phasal movement up the syntactic tree (with **Tense (TP)** being adjunct in nature and **Case** being derived via **vP**). What we hope to gain from the general theory as laid out here is that there is indeed reason to assume a ***brain-to-language mapping*** and that such language and movement mapping follows a pegged maturational development of the brain.

If Movement is so critical here, a theory will have to be devised which defines what is the motivation for it. The simple answer will be that abstract *functional features* are what drive movement (up the syntactic tree) where a substantive *lexical* head category is selected by its more formal *functional* head counterpart—viz., where C selects T and where *v* selects V. But the question still remains regarding the nature of the movement: What is the nature of the features and does it drive *local vs. distant*, or *inter vs. intra phasal* movement as discussed above regarding binding vs. movement)? So, a two-prong analysis of movement may be in the offering.

Control vs Raising Predicates and Null Constituencies/Empty Categories (R, 266). As part of this *Family of Move,* two types of **NULLs** emerge as defined by their duality-of-semantics status:

(i)   If the Null Constituency refers back as 'coreferential' to its matrix clause, and both anaphor and antecedent can serve independent theta-roles, then Null is a big PRO (Control).

(ii)   If the Null doesn't refer back and draws on its own separate theta marking, then Null is an Empty Category (EC) (Raising).

   a. John wants [PRO to scare them] (Control)
       i.   John is theta-role EXPERIENCER of the verb 'want'
       ii. PRO is AGENT of verb 'scare'
       Theta-criterion: 'one argument receives only one theta-role' is preserved.
       John and PRO are two different arguments with two different Theta roles.

   b. I will arrange [for him to take the test] : (*him* is subject of infinitive verb '*to take' the test)*

       (*For* triggers accusative case/Control)

       I will arrange [PRO to take the test] : Silent PRO = 'For me'

   c. John seems [ec to scare them] (Raising)
       i.   The verb 'seem' is a raising verb: it doesn't theta mark its subject. Hence, the empty category inside the infinitive VP forces the subject of the infinitive verb 'to scare' to percolate up(raise) to its matrix subject of a finite 'does-phrase' in order to maintain its singular theta role. John and the empty category is connected via trace: John and empty category share the same theta role AGENT.
       ii.   Theta-criterion is preserved: an argument must have a theta role (and since the raising verb 'seems' doesn't assign a theta role to its subject, the percolated subject via TRACE preserves the one lower-level theta role via VP in situ).

3.   Merge: Local Move

While the basic building-blocks of language surely are lexical words (N, V, Adj, Adv, Prep), what glues them together to form their respective phrases is the quite primitive and fundamental notion of Combine, i.e., 'Merge two items together'. Such a primitive 'Combine' of two (or more items) renders e.g., a Verb Phrase [VP [V] [N]] (combine V + N), Adjective Phrase [AdjP [Adj] [N]] (Adj+N), etc. This primitive combine is what constitutes the very earliest stages of child language acquisition, the first stage during which only lexical/thematic items combine to form early two/three-word speech (See Radford 1990, Galasso 2016). The notion of two items sitting next to each other requires an **adjacency condition**, without which no combine can take place.

Simple Examples of Merge:

Compounding:

      i.   N + N : [Adj black] [N bird] => [N blackbird]
     ii.   V + N : [V break] [N fast] => [N breakfast]

*Note how stress [strong ´] [weak `] gets reconfigured after Merge of two items:
    When item is isolated (as one word), it yields a single stress (break) /brék/, (fast) /fǽst/
    When items combine (as two words), it yields 'initial strong stress', 'final weak stress:/brɛkfəst/

(Also see as an example of 'stem+stem' combine Gordon's 'Rat-eager' experiment. (§1, (19) below).

Derivation morphology:

    [V Teach] + {er} => [N teacher]; [V Compute] +{er} => [N computer]; [N girl] + {ish} => [Adj girlish]

First-level Merge

First-level 'Merge operations' over second-level Move shows up quite early in the basic compounding properties of morphology.

Let's flesh out some Functionality distinctions as played out in so-called 'Synthetic vs Root' compounds, utilizing the term 'functionality' to cover subtle aspects of communication.

Question: Does Merge over Move operations show up communicatively in formal aspects of syntax?

(47) **'Root** vs. **Synthetic' compounds**

In the same vein of our two-step derivations found above, consider the 'local-Merge/semantic' vs. 'distant-Move/syntactic' distinction in the examples below:

(i) **Merge (non-Move): Root compound**

'chain smoker' => Not a *'smoker of chains' (so, = root compound)
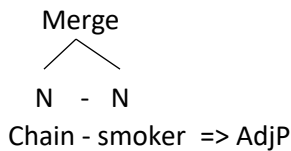
        Merge   Ex. John is a **[N chain-smoker]**

         N - N
      Chain - smoker => AdjP

Note how the above Merge construct would be similarly compared to a simple merge-based Adjectival Phrase 'black bird':
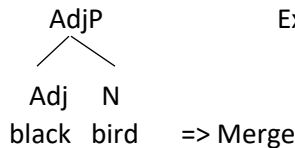
         AdjP   Ex. This is a [**AdjP black bird**]

       Adj   N
     black   bird   => Merge

(Noting that *blackbird,* as a compound in its own right, has a different interpretation from the Adjective Phrase: a *black bird* is a bird colored black, while a *blackbird* is a *type* of bird, presumably black).

So, we note above that there's no necessary movement to derive the double-noun sequence {N+N} as a combined single N (compound), with the appropriate interpretation of the N+N sequence that follows—e.g., *John is a 'chain-smoker'* has the same non-movement quality as found in the structure *John is a 'teach-er'* where the derivational morpheme {er} is of a strict non-move-based procedure. Let's note how 'chain-smoker' doesn't take on a modification-interpretation as *'smoker of chains', which is actually what we do find below in synthetic/move-based compounds.

(ii) **MOVE: Synthetic compound**

Note the stepwise MOVE derivations:
       cigarette smoker => is a 'smoker of cigarettes'

       a) cigarette smoker   (Merge)
       b) smoker of cigarette(s)   (Move-1)
       c) a cigarette smoker   (Move-2)

Stepwise derivations of movement:

(48)   => move-2   (raising of the Noun 'cigarette')
        N   => move-1   (raising of the Noun 'smoker')
   cigarette N   => merge   (the simple merge of two items)
        smoker Poss
              ~~of~~ N   N
                 ~~cigarette smoker~~
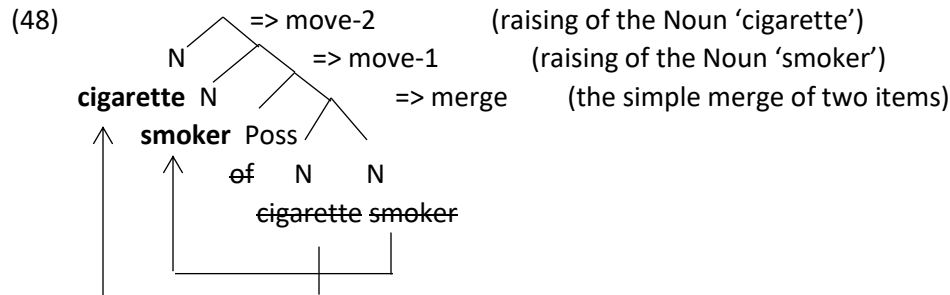
While 'Root Compounds' carry a [-Move] feature, it's clear to see that the 'Synthetic Compound' carries a [+Move] features.

Note: A developmental aspect of root vs synthetic shows-up in the maturation of movement across phases of syntactic development: Let's consider a dual-mode phrase that can be spoken by any English-speaking adult:

(a) 'Wine bottle'   [[N wine] + [bottle]] => N 'wine' becomes adjectival AdjP: 'A wine bottle'
(b) 'Bottle of wine'   [bottle$_j$ of [wine] + [~~bottle~~]$_j$] => Genitive Phrase: 'The wine's bottle'

(where subscript {j} shows MOVE of 'Bottle'. A Move-based operation which displaces the two mere lexical items (bottle, wine) and turns them into a category [__ [of wine ~~bottle~~]].

*A note on child language: In the paper, 'Small children's sentences are dead on arrival' (Galasso 2015), the idea is suggested that young children, owing to their lack of movement, are momentarily (lexical stage-1) stuck at the 'wine bottle' stage, where concrete telegraphic phrases are empty of all movement operations (a non-INFLectional stage). (See (50ff) below).

Theta-marking

Local merge is responsible for theta-markings since such **θ-role** assignment originates within the VP. Let's highlight some basic theta-roles:

(49) Theta-Roles: (See R, p. 246 for further discussion)
   (i)   AGENT (person perpetrating the act: *'John* (AGENT) kissed Mary' <AGENT John>
   (ii)   PATIENT (person under the influence of the act: *'John* (PATENT) fell down'. < PATENT John>
              (PATENT is also referred to as THEME). Note that while 'John' both for (i) and (ii) serves as Subject, the two theta-roles are different.
   (iii)   ACTION (verbal act): John *pushed* <ACTION push>
   (iv)   EXPERIENCER (entity experiencing some psychological state: *'John* likes Mary' <EXP John>
   (v)   INSTRUMENT: 'with a *knife*' <INSTR knife>
   (vi)   GOAL: destination of some entity: John went *home*

Theta roles are generated within lexical categories, e.g., VP:

VP

Agent → Action

Action ←→ Patient
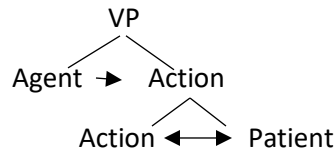
This assumes that all theta-role marking must be done very early in the derivation, and that each argument (e.g., DP) bears one and only one theta-role assignment (Chomsky 1981, 'Theta Criterion').

Consider the double semantic roles of 'which books' below (see ref. in G, p. 35):

'John read which book?'
    (i) [Which book] first starts out as {a}: <object of 'read'> (read *what?*)
    (ii) [*Which book*] then moves to take over a semantic role of {b} <interrogative operator> (which book?)

      Theta role {a}. Object/Theme
      Theta-role {b}: Interrogative

Showing full 'dual semantic role' derivation: [Which books did [John (~~did~~) read ~~which books~~?]]

This shows that an argument can take on two separate θ-roles but only via movement (two different theta-roles are banned from an argument in-situ of a single position. (Chomsky 1981, Each θ-role is assigned to one and only one argument). This provides an additional account for movement—viz., to possibly provide differing θ-roles to a single argument. Chomsky 2013 (p. 4) considers that movement allows for the thematic-semantics of SCOPE. (See G, p. 8).

Move-based Theta role assigner:

 i. [TP-Dec [VP read which books?]] (TP-Declarative): [which book] = i. OBJECT of 'read'
 ii. [CP -Int Which book (did) [TP [VP]]]: [which book] = ii. Interrogative theta feature

  ⇨ For which books X, John read books X.
  ⇨ [CP-Int   [TP-Dec   ]]

Young children at the so-called lexical-thematic stage can only produce utterances such as 'Him do it', 'daddy drive car', etc. where we find INFLectional morphology completely missing: A Non-INFL stage-1 (See Galasso 2016. For the classic first seminal study, see Radford 1990).

Note: Also see **Ergative** structures in the same light (Also see (52) Split VPs).

4. Distant Move

All items must move at least once—all right branching structures end in a trace (Kayne). Phase-based theory (Marantz).

The 'Phase' (Section taken from G, 2016, Ch. 6)

Chomsky (2000) argues that derivations (which get transferred [/$/] to full interpretation) proceed in chunks called 'Phases'. There are working-memory limitations as well as syntactic motivations behind the theorizing of phases. Regarding a 'working-memory' limitation, the mere fact of trying to hold a very long syntactic string in one's mind and then to perform some syntactic operation on it would surely tax working memory to the brink of catastrophic breakdown.

In order to lessen the burden of working memory, Chomsky suggests that phases must get transferred [/$/] via spell-out as soon as possible in the derivation, based on these phase boundaries which minimally include the argument structure of the vP light verb and the expressive structure of a CP: VP and TP are not considered phases (and so much continue build-up projection until they reach vP and CP status respectively— viz., VPs with build-up projection to the light verb vP must then get sent to LF-transfer for theta-argument structure interpretation, and TPs with build-up to CP (T being adjunct in nature hosting Tense lowering from CP (CP houses tense) must then be sent via CP. Hence, a defining aspect of what constitutes as a phase is this notion *of movement and transfer*—namely, if movement is allowed (i.e., build-up delaying transfer), then it is not a phase. VP allows build-up/movement (to vP), TP allows build-up/movement to CP, hence, the two are not phases. (In the section below, we'll come to reanalyze such a move-defined phase-based theory in the way of child language acquisition: that young children's sentences don't contain movement hence all projections get immediately sent to transfer. (See Galasso 2015, link to discussion presented below. See 'Wine bottle' vs 'Bottle of wine' analogy in Marantz section below, whereby a lack of movement at stage-1 of child syntax forces a phrase to immediately get sent to transfer/spell-out, thus freezing the child's utterance at the semantic [+Interp] phase of its derivation—hence, 'small children's sentences are dead on arrival'.

https://www.academia.edu/15155921/Small_Children_s_Sentences_are_Dead_on_Arrival_Remarks_on_a_Minim alist_Approach_to_Early_Child_Syntax_Journal_of_Child_Language_Acquisition_and_Development_vol_3_no_4_D ec_2015

Abstract pulled from paper:

As the title suggests,
Small children's sentences are 'Dead on Arrival'—if by that we mean that the young child's syntactic parser is unable to advance (MOVE) a morpho-syntactic utterance, both at PF (phonology form) and at LF (logical form)up the syntactic tree (whereby **MOVE**ment would thus save the derivation from being sent off immediately to early semantic transfer). The deficient for a lack of movement is not just a surface-level PF deficit, but is also pervasive at interpretation. Hence, as a metaphor for the lack of movement (both at PF and LF), children's early utterances are indeed semantically frozen deep within the prosaic trappings of the bottom portion of the tree (namely, within the VP phrase) and are thus sent immediately to spell-out. In this paper I propose an initial 'merge-only' stage of child syntax which can account for a rather wide spectrum of implications leading to the impoverished state of early child syntax. Using Chomsky's current Minimalist Program (MP) framework, I adopt a 'Merge-over-Move'

Marantz's 'little heads'

(For 'wine-bottle' analogy on movement/phases, see G, p. 19).

If for whatever reason movement is blocked and transfer ensures, then the project acts as a phase. Let's look at the sentences below and see how this notion of phase would apply, focusing here on the two distinct operations performed on 'wine bottle'. As we will see, the notion of an AdjP being defined as a 'phase' has been cited in the literature (as well as notions of DP, PP arguments for and against their status as phases). For instance, regarding the AdjP, Marantz (2001) proposes that all heads which form word-identities based upon a simpleton merge operation (such as all lexical categories noun, verb, adjective) contain a *little x-head*, to generalize from the *little v-head* that creates the extended verb and phrase vP and that these little heads may uniformly correspond to phase-heads. As a consequence, syntactic computation could be unified above and below the word level.

Category changing morphology could yield multiple phases within a single word and or phase depending on the nature of the underlying processing of word/phase, and that cyclic phonological effects within words could be related to – hopefully reduced to – the cyclic operation of phase-based syntax (Marvin 2002). Consider such a dual status of AdjP below:

(50) 'John brought <u>a wine bottle</u> for the party'. [TP John brought [vP John bring /$/ [AdjP a wine bottle] for the party]]. => vP…[AdjP [wine bottle]] No MOVE/ => /$/ AdjP Transfer The AdjP lexical category 'wine bottle' must proceed to transfer since no further upward mobility (up the tree) would enhance its interpretation. In this sense, since it must transfer here, we can take the AdjP as having the qualities of a phase (AdjP with small-head, in Marantz terms, an s-AdjP (s-AdjP). Hence, the s-AdjP is forced to proceed as a phase to transfer /$/ due to the fact that there is no forward-looking movement (or escape hatch in a higher edge position) which would attract it up the tree. So, [s-AdjP] is a phase].

However, consider a non-phase big-AdjP below:

(50) i. (John brought <u>a bottle of wine</u> for the party'. [TP John brought [vP John bring [DP a bottle of wine] for the party]]. => vP…[[DP a bottle [+Gen] of wine] => [AdjP wine bottle]] MOVE/AdjP => Non-transfer (Where [+Gen] marks for the Genitive/possessive feature) The 'escape hatch' attracting such movement up the tree might be the edge position—a specifier/multi-specifier position of DP (2-prime).

(50) ii. DP Spec D' edge AdjP… [wine bottle] Let's take the derivation step by step (bottom-up).

In the PP 'for the party'—it not being a phase—can survive as a derivation up a notch until it gets tagged onto the next highest phrase, the DP 'a bottle of wine-(for the party)'. But DP—it too not being a Phase—also survives until it tags to the light verb vP John/He bring…It is at this point where both phrases (DP>PP) merge in the syntax and may get transferred as a single derivational

DP/PP chunk. In fact, the PP (and perhaps DP) here is traditionally analyzed as an adjunct providing adverbial information and therefore can front to the beginning of the sentence (e.g., For the party, John brought a bottle of wine (for the party), or even consider e.g., The bottle of wine John brought [the bottle of wine]…was great!). The fact that PPs (and potential DPs) can front suggests that they may survive transfer (i.e, the DP/pp didn't transfer as a chunk). Otherwise, if they/it did transfer—since post-transfers can't remain as a surface derivation— certainly in the case of the PP—it would not be able to stick around to front. So, the first possible transfer is at vP John bring a bottle of wine.

Transfer phase #1. (sentence 1) 'John brings a bottle of wine'. It is here at phase-1 transfer where any thematic/argument structure related to Semantics and Case receives full interpretation. (Recall we suggest in our present working theory that Case straddles the thematic/syntactic divide and is thus located within vP—it being a hybrid lexical/functional category—and not TP which is usually assumed).

Transfer phase #2. 'John brought [+past]…' It is here at phase-2 transfer where we pick-up formal Tense and Agreement relations— Agreement being the feature par excellence which drives movement up the syntactic tree. Case is not seen as a motivator for movement). So, according to the above two-step derivational process, we first proceed to interpret (at the semantic/Case level) the Agent/Subject and its course of Action/Verb, (where x = bottle of wine).

The second step then incorporates what we have already interpreted from phase-1 transfer and now tags that chunk onto phase-2 where we deliver Tense/Agr. Now, for sentence (1), (restated below in (3)) although similar sounding, things proceed a bit more early-on in the derivation.

(3) 'John brought a wine bottle for the party'. [TP John brought [vP John bring [AdjP a wine bottle] for the party]]. => vP…[AdjP [wine bottle]] => No MOVE When there is no Movement, this traps the derivation at the level of phase within a position where it is forced to proceed to transfer (and get interpreted). In other words, Move keeps a string viable as it moves upward for future transfer. The same PP step would apply as in (1) with 'for the party' delaying its transfer and tagging to the AdjP. The AdjP, not being a phase, then tags onto the vP John bring…and at this point all transfers.

But crucially note that the two sentences differ not only regarding their syntax but regarding their semantics as well—viz., 'a bottle of wine' is interpreted as 'bottle that has wine in it [+Possessive/Belonging to wine], whereas a 'wine bottle' denotes a mere adjectival interpretation and doesn't carry a [+Possessive/Belonging to] feature. What these distinctions tell us (at the semantic level) is that the notion of phase has to be extended to covering DPs and AdjP constituencies—hence, we must now add DP, AdjP to vP and CP as potential phases. TP and VP remain as non-phase boundaries. (See e.g., Chomsky's 2006 'Approaching UG from below' discussion for DP as potential phase). Why? Well, in order to maintain two different readings between (1) and (2), the relevant phrases in question would have to proceed to transfer at crucial constituencies—e.g., [wine bottle] would have to transfer as soon as possible in (1), but MOVE in (2) delays transfer.

Consider the steps of movement for 'wine bottle' > 'bottle of wine'

(51)   => move-2   (raising of the Noun 'wine': yields Adj 'wine bottle'
            N   => move-1   (raising of the Noun 'bottle': yields genitive 'bottle of wine'
      w**ine**   N   => merge   (the simple merge of two items)
          bottle Poss
              ~~of~~   N   N
                  ~~wine~~   ~~bottle~~   => no phrasal constituency: 'mere sister-sister combine'
                                      (e.g., wine bottle, bottle wine: non-AdjP)
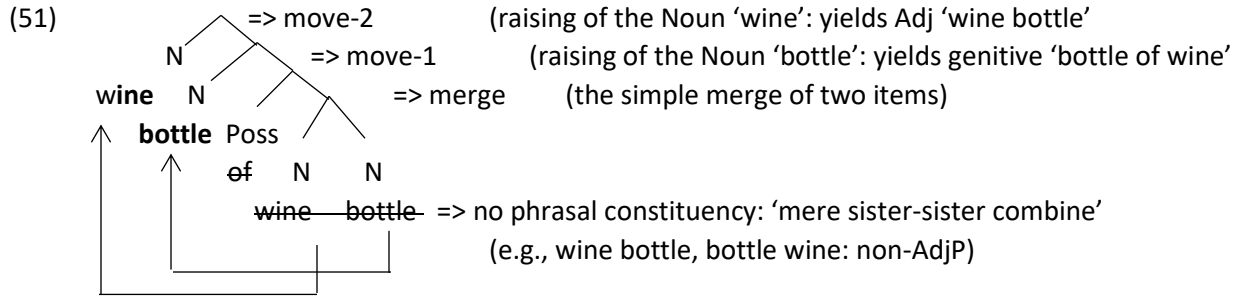
*Note and caveat: What we can say here regarding child language, is that in lieu of no movement all structures proceed to transfer—hence, in the above example, young children prior to their INFL-functional stage-2, don't have access to genitive structures. The fact that they seemingly projective lexical phrases such as AdjP at the lexical stage-1 comes with the caveat that such prosaic lexical projections might simply be merge-based combined structures carrying a default value of AdjP, if we assume that AdjP must first proceed from out of Genitive Phrase. This might be a mere theory-internal consideration, but considering that the merging of two nouns [N+N] seems to provide an adjectival interpretation (glass/plastic/wine bottle) => AdjP there might be some merit to this theory-internal rationale.

New definition of 'Phase'

Recent theoretical accounts have sought to promote movement as that definitive operation which solely triggers syntax. Chomsky (2000) argues that it is in fact movement which drives syntax alongside and the unique language property of displacement, with movement being forced to act as a condition to check-off formal features of a functional Head within some strict adjacency position/agreement relation (once considered to be a Spec-Head relation, though this stipulation has eased). Hence, movement was feature-driven and would arise as a need to mollify some feature checking condition. Recent MP attempts at defining exactly 'what' moves 'when' has brought on a new account which defines a phase as that constituency which must check-off as early as possible for full interpretation (Logical Form (LF)) to proceed.

In-note on Kayne: 'All lexical items must move at least once: no right branching nodes can end in a trace'

That 'all right-branching structures must end in a trace'—viz., that everything must move at least once (Kayne).

Due in part to the 'labeling algorithm' (of PAIR not SET) as well as the Linear Correspondence Axiom (LCA) of Kayne (1994) (see 'The antisymmetry of syntax. MIT Press.), in order for a lexical item to gain visibility (word order), the item must be displaced from its base-generated position, (i.e., mustn't be in situ). Let's consider two cases below dealing with the items of a category DP (e.g., a proper Name) and NP (an expletive Pronoun 'it'). (Taken from Uriagereka, 2000).

Proper Name: Consider the sentence 'I like Ike' (focusing on the [VP like Ike])…

(i)   VP
          /   \
        V   N
       like   Ike
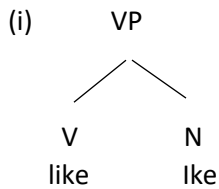
Here, the VP as it stands is a mere SET {a, b} product of merge/Combine with two sisters {like, Ike}. Recall, that in order for a Labeling Algorithm (LA) of a Phrase to ensue, at least one member/item must move creating 'Dynamic Asymmetry' (DA). If the phrase stands as is, without a member moving out to create a PAIR, there can be no visibility per word order. There are three ways, theoretically, which can motivate movement: (1) one is by upward movement of the verb (to check some functional [-Interp] feature: this would be the case of Head V of VP moving up to insert into the Head of a light verb vP. A second way (2) is to always create more structure. So, in the case of the Noun 'Ike', it being a name, one way to create DA is to postulate that the Noun must move upward in forming a DP (names have referential qualities, a DP property). The added structure always entails, by definition, movement. Hence, one of two ways motivates movement in taking us from an unorder SET {a, b}, to an ordered PAIR <a, b>, either via VP>vP, or via NP>DP (noting that both projections percolate upward from lexical to functional category—upward movement being instigated by some checking-off of a formal functional feature). (3) When we reach rock bottom, as a last resort, we have to cliticize.

Let's flesh out these two ways below:

(ii)   VP>vP => SET {like {like Ike}} => PAIR <like, Ike>

                    vP
                  /   \
                v   VP
             { like   /   \
                     V   N
                   { like   Ike } }
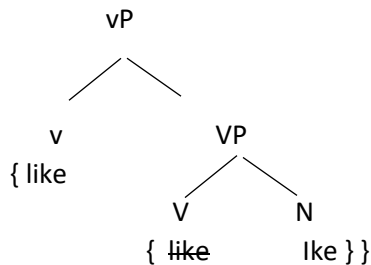
N=DP => MEMBER {Ike} => PAIR < F, Ike> where F stands for a Functional (silent) category/word such as a determiner ['The' + N], DP {e.g., **The Ike** that I know will do his best to win!)

(iii)　　　　　　DP　　　'We like Ike (for president)'

　　　　　　D　　　　D'
　　　　　(The)

　　　　　　　　　D　　　　NP

　　　　　　　　{ Ike　　N　　　　PP…..

　　　　　　　　　　　{ Ike, }}
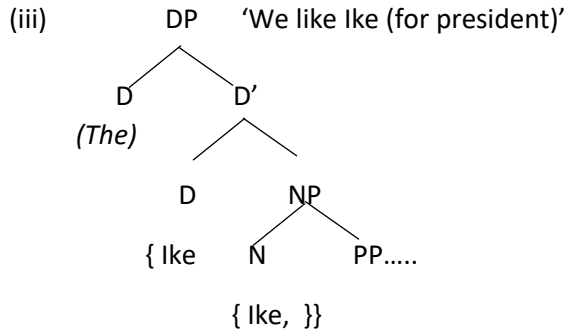
　　　'I Like it' (cliticization of expletive pronoun)

(iv)　　　　　　VP

　　　　　　V　　　　N
　　　　　like　　　it

(v)　　　　　　VP　　　(noting that clitics form as part of the host word)

　　　　　　V　　　　N
　　　　Like+'t　　(it)
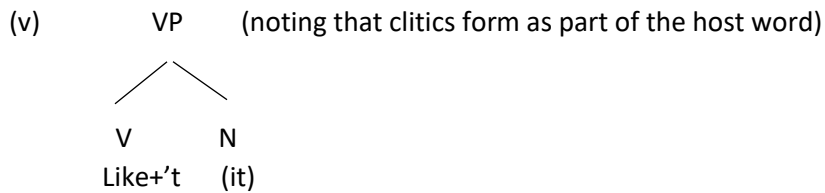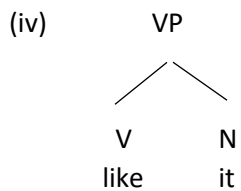
'That all right-branching structures must end in a trace': (Kayne).

Given we already have a step-1 VP>vP movement, a step-2 N>DP movement secures this LCA notion that all items must move at least once in a derivation. Clitic formation of weak expletives as last resort of LA leads to DA forming step-3.

(vi)　　　　　　XP-2

　　　　　Spec　　　X'

　　　　　　　　Head　　　XP-1

　　　　　　　　　　Spec　　　X'

　　　　　　　　　　　　Head　　Comp

　　　　　　　　　　　　　　　　trace
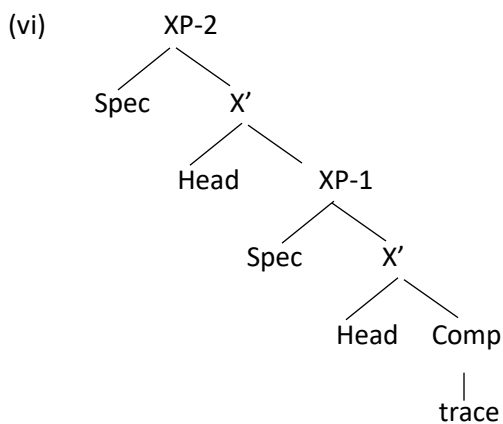
'So, the general rule is always the same: linearization problems in right branches disappear at the moment we posit more structure. Of course, at the very end of the right branch we hit rock bottom, and we have to cliticize' Uriagereka, p 221)

A Final look at Distant-Move via Probe-Goal Union (PGU)/Relation (Miyagawa).

Assuming a duality of semantics treatment of PGU, we find that while some properties of **External Merge** (= Merge) might be the result of a PGU in relation to CASE/SEMANTICS, **Internal Merge** (= Move) is motivated by formal AGREEMENT/SYNTAX. Let's flesh this out (following Miyagawa's analysis):

Probe-Goal/Local: vP>VP…searches for Thematic properties (also related to CASE). Case partially seen as a semantic property, particularly Inherent Case, Oblique Case, whereby the marking is assigned via the merging of a specific lexical item (the preposition in these cases…e.g., [PP with [DP [-Nom] him]], not *[PP with [DP [+Nom] he]]. Such local PGU has the flavor of Merge properties insofar that the assignment (Case) is lexically driven, and not structural driven (as in Structural CASE, where the subject of a finite verb is marked as [+Nom] (Nominative).

Probe-Goal/Distant: CP>TP…searches for the higher-order relations as driven by functional-projections (CP, TP). These higher projects have to do, some say (also following Miyagawa here) exclusively with AGREEMENT (a CP>TP hand-off relation of PGU) as well as FORCE (CP).

Let's go over an example of how PGU might work within a given sentence (see below). (See G, p. 9).

*Notice how structure is a restatement from the same structure found in (8) of the prologue ex.
'John broke the window' , 'The window broke').

Light verb **vP-shell:** *Probe-Goal relation and Feature Valuation*

TP>vP>VP: Instances of Head Movement, wh-movement

Split VPs: Ergative Predicates
e.g., 'He rolls the ball down the hill'   **PGU targets:**
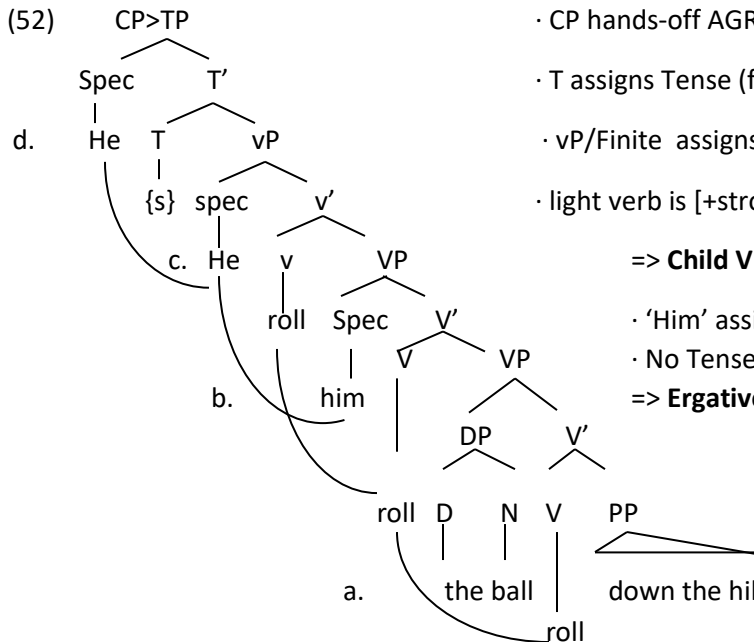
(52)   CP>TP   · CP hands-off AGR features down to TP*

          Spec   T'   · T assigns Tense (featural or affixal)/AGReement via CP

d.   He   T   vP   · vP/Finite assigns case…T/AGR yet to be assigned

              {s} spec   v'   · light verb is [+strong affixal]: attracts Verb raising

                    c. He   v   VP   => **Child VP structure: 'Him roll (the) ball'**

                        roll Spec   V'   · 'Him' assigned [-nom] case via default.
                                                  · No Tense/AGR on verb
                              V   VP   => **Ergative Predicate: 'The ball rolled'**
                    b.   him

                                  DP   V'

                            roll D   N   V   PP

                    a.   the ball   down the hill

                                            roll

(1)   vP Case / VP—Thematic-Lexical Structure: Ergative Prediates (see. R, 347) (See G, p. 9)
   a.   The ball roll down the hill.   ('roll' assigns **Theme argument** to 'the ball'). **(=> Ergative)**
   b.   Him roll the ball ~~roll~~ down the hill. ('roll' raises to v to assign **Agent argument** to 'Him')
   c.   He roll the ball ~~roll~~ down the hill. ('He' received [+Nom] case in vP).

(2)   TP—Functional Structure: Tense/Agreement features.

   a.   He rolls the ball ~~roll~~ down the hill. (EPP forces 'He' to raise. Tense/AGR project)

   The above structure shows both **Head-to-Head movement** (of the verb: V>V>v) and **Spec-to-Spec movement** (of the subject: VP>VP>vP>TP)

(3) Percolation of Features to lower Head: C hand-off AGR to T, T hands-off Case/Tense to vP*
   e.g., What's the wheels doing? (Note: H of CP 'is' agrees with Spec of CP wh-pronoun 'What')

*See Radford 2009 (p. 398, Ex. (35)) for analysis which states that CP hands-off AGR to TP.

Let's consider as our final tree some possible PGU configurations for the 'breakfast' sentence below:

'We are breakfasting…'

(53)

Move / AGR

CP   CP hands-off AGR features onto TP

{AGR}   TP   Merge / Tense, EPP-features

Adjunct   T'

vii. We   T   AspP   Move / Aspect / Case

vi. are   Asp   vP

v. {ing}

Spec   v'

         Merge/Θ-marking

we   v

iv. breakfast-ing   set-merge / copy

              pair-merge / pull

α   *i.* α   β

iii. [breakfast]   *[break]*   *[fast]*

ii. [[break] [fast]]
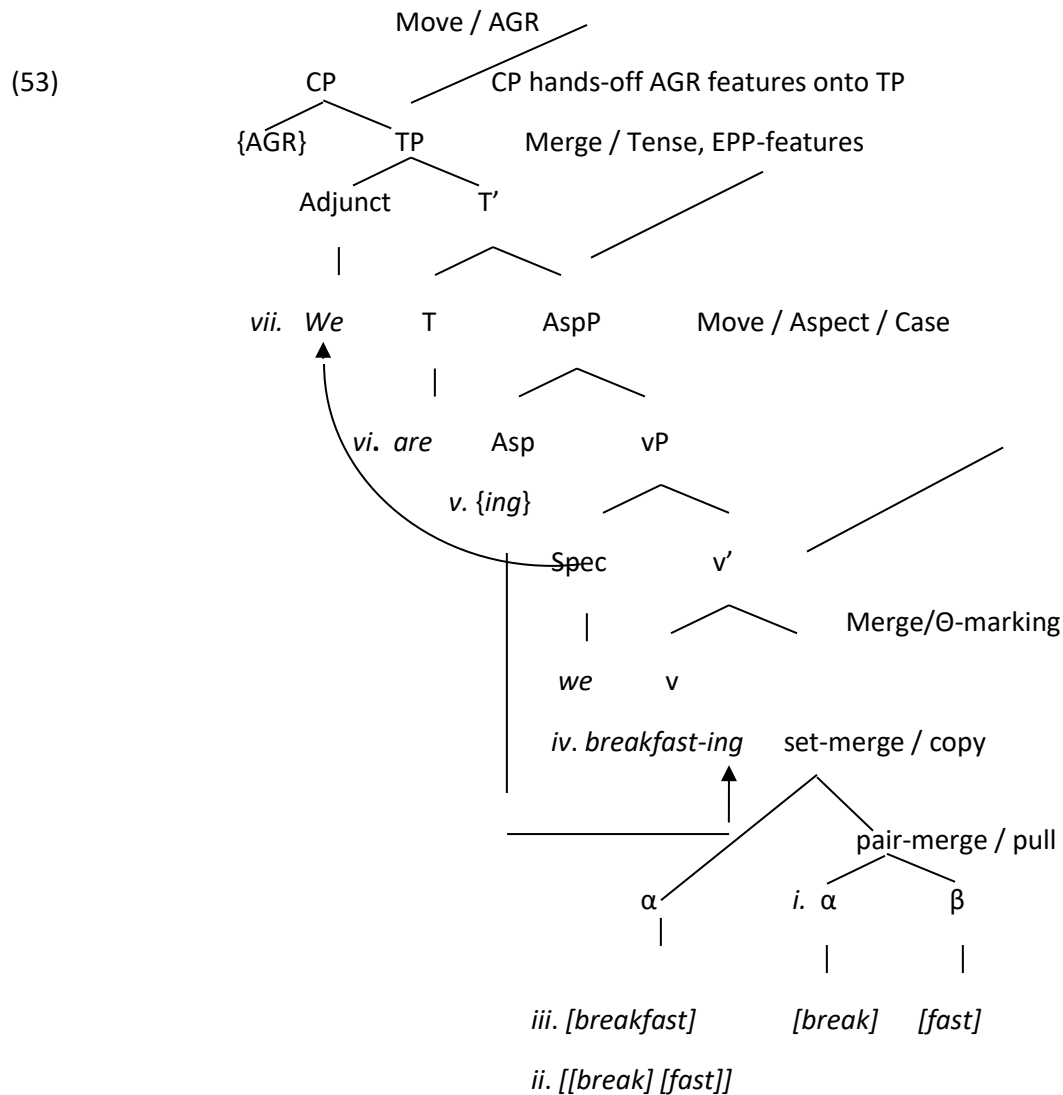
Sequence of Merge to Move: <u>Two probe-goal operations:</u>

         a. *Set-Merge-based* 'Probe-goal' to establish identity among set-membership and thus create labeling of phrase.
         b. *Move-based* 'Probe-goal' motivating raising and stripping-off of formal features.

(i)   **Pair-merge**: Items {α [break]}, {β [fast]} are 'pulled' directly from the lexicon (no linear order): the two items under pure pair-merge are sisters and could either come to be interpreted as *break-fast* (V: 'to eat in the morning', N: 'what you eat in the morning') or *fast-break* (a 'basketball term').

(*Note that it is not until 'Set-merge' {α {α, β}} that we obtain word order directionality).

(54)   set-merge / copy

a = 'breakfast'

                                                                                     pair-merge /pull

b = 'fast-break'

                                                        α   α   β
                                                        |
                                                                         | |

                                    a.   *[breakfast]*   *[break]*   *[fast]*

                                                                         | |

                                                                         α   β

                                    b.   *[fast-break]*   *[break]*   *[fast]*

What the Set-merge allows us to do is create an overlapping template upon which copy can play-out. If we have as a membership which includes $\{\alpha, \beta\}_2$, $\{\alpha, \beta\}_1$, then we have a parallel structure in place for an item to break sisterhood and be promoted into a dominance position. Of course, what that item is will be determined by C-I (the semantic/conceptual-intentional interface), with some residual factors bleeding in from phonology (since, as Chomsky (op. cit.) claims, syntactic determinants of order fall within the phonological component where prosodic stress may select and label and head).

The operation 'pair-merge' will be later viewed as a possible account for variable word order (SV, VS, OV, VO) found amongst Single Argument Structures (SASs) in early child speech (Galasso 2016). On the other hand, Double Argument Structures (DASs) seem to fix word order in child speech. The SAS vs. DAS templates suggest an overlap with pair-merge vs. set-merge respectively, whereby DAS requires asymmetrical c-command to be spread across the DAS [Spec/adjunct [Head-Comp]] (with Head placing either initial or final based on head directionality parameter). Hence, it seems DASs necessarily trigger hierarchy out of design. 'Set-merge' (and not pair-merge) is defined as a structure which allows the performance of some minimal 'probe-goal' operation to act upon set-membership for the sole purposes of identity. This probe-goal relation is different than what we find regarding MOVE since this Merge-probe-goal seeks out function-arguments for identity purposes—i.e., determining which element will become the head. 'Move-based' Probe-goal rather seeks out formal features for reasoning having to do with check-off.

(55) **Two probe-goal operations:**

      a. *Set-Merge-based* 'Probe-goal' to establish identity among set-membership and thus create labeling of phrase.

      b. *Move-based* 'Probe-goal' motivating raising and stripping-off of formal features.

(ii)   **Set-merge**: 'anti-sister/c-command' provokes a 'copy' of one of the 'pulled' item. The operation 'copy' forces a selection of one of the two pulled items to serve as 'Head'—copy thus creates function-argument status. With newly created c-command, coupled with a 'head status' (and head directionality parameter), we now can derive what was once originally a compound *[[to break] [a fast]]* to lexicalized *[[break] fast],* with *break* serving as verbal head. We now derive *[breakfast]* (to eat in the morning) which becomes an unproductive and undecomposed new lexical item, used here as a verb. Note that it is here at Set-merge that we lose independent productivity of isolated lexical items—hence, set-merge is a quasi-grammatical operation (since it is move/copy related) and may be *semi-productive* at best (as seen with derivational morphology). Set-merge yields function-argument status whereby word-order is imposed once head directionality parameter is set [+/- Head initial]. We take it that word order naturally falls out of the computational design of set-merge in conjunction with the head directionality parameter, given that all trees, by definition, are binary branching. (Note that a tertiary tree could not establish such word-order in this manner).

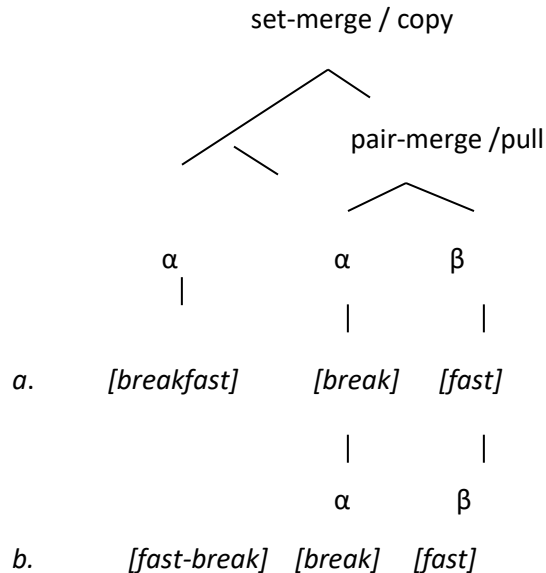*(iii)*   The verb *[breakfast]* raises to *(iv)* light verb vP (via MOVE) and gets into a position whereby the verb-stem can receive *(v)* an Aspect affix {ing} from the Aspect Phrase (AspP), a process known as 'affix-lowering'. Since T in English is not strong enough to force verb raising, the verb must stop at Asp.

(iv)   The tensed verb [+Present] *are* gets pulled directly from the lexicon in its bare form and inserts directly into head of T. (There is no agreement yet at this point of derivation between the verb and subject. But once CP projects, phi-feature AGReement is established between verb and subject).

    *(Note: in this account CASE (structural) is established at this point (as a result of *light verb* vP projection. Other accounts peg CASE specifically to TP. Since TP is an adjunct position in this account, as well as being a non-phase [+Interp] projection (cf. Bare Phase Theory), we suggest that it's the Phase vP which projects CASE).

(v)   The subject inserts via adjunction into Spec/adjunct of T (due to an EPP feature on T which stipulates that all clauses must have a syntactic subject).

For further discussions into Probe-Goal relations, see the following: (Exercises §§7-10, and the 'Summary' section found at the end of 'Exercise' section.

Final Notes on Phases (Chomsky 2001)

Some initial assumption:

(1) **Phases** are *vP** (*transitive *light verb*) and CP, their domains of which **transfer** to spell-out (PF & LF). A **Domain** is the Complement of a phrase/phase, with Head and Multi-Specs serving as an 'escape hatch' to host moved elements (up the tree).

(2) **Probe-Goal (PG)** searching must be local (economy considerations).

(3) **Case** is assigned according (via PG): (i) Accusative via *light verb v* to V, (ii) Nominative via T to v.

(4) **Movement** is driven (i) to check [-Interp] features via PG (**Case, AGR**), (ii) to attach verbal **affixal feature** (V raising to v), (iii) to satisfy **Edge-feature,** (iv) to satisfy **Extended Projection Principled (EPP).**

(5) **Transfer:**
   (i) At the end of each Phase, the domain (Complement of the phase Head) undergoes Transfer.
   (ii) At the end of overall derivation (CP), all remaining constituents must transfer.

Let's flesh this out with the sentence below (See R, 393).

What have they done?

```
                        CP = transfer-3 (end of derivation gets transferred)
                       /\
             PRN          C'
             What        /\
             (ix)   C         TP = transfer-2 (They raises, Probe-Goal via have)
                  have-Ø     /\
                    (viii)  PRN        T' (tense) = (Adjunct in nature/non-phase)
                           they*      /\
                           (vii)  T       vP = Spec-2 (outer-spec)
                                 have    /\
                                  (vi)   T      v' = Spec-1 (inner-spec)
                                        /\
                                     Prn    Prn    v'
                                    what   they  v    /\
                                                    VP = transfer-1
                                      (iv-v)  done+Ø   done what
                                                        /\
                                              (iii)      (i, ii)
```
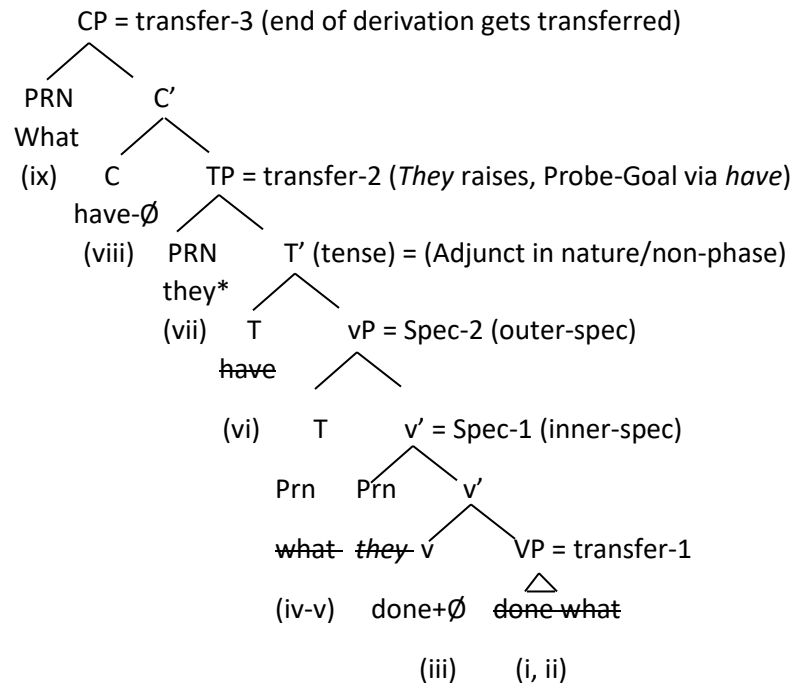
transfer-1: [done what] (domain of vP)

transfer-2: [what they done ~~done what~~]] (domain of CP) *(Prior to Case, *they* (Spec-1)= default THEM)

transfer-3: [What have [they ~~have~~ done [~~done what~~]]] (end of derivation transfers)

Where strikethroughs ×̶×̶×̶ indicate transfer and so get NULL spell-outs.

Let's take each derivation stepwise bottom-up:

(i)   [VP done what]: the verb DO (spelled-out as *done*) **merges** with its complement *what* and assigns its **thematic role** (Theme-Object). [VP [DO] [what]]. Case not yet assigned since VP can only assign theta-roles via semantics.   Merge

                                   [DO   What]
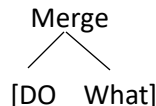
(ii)   *Transfer-1* of VP (since it is in the **domain**/COMP of vP). [VP what done] immediately transfers with **NULL** spell-out.

(iii)   *Affixal {Ø} in light verb v* attracts V raising from V to v. [vP {v-Ø} [VP V̶]]. Affix feature of light verb being semantic equivalent (light verb 'to make': 'make-DO'). Persian has a particular verb here: *Kardan (to make) (e.g., Persian translation: 'to make-play the piano' (= play the piano).*

         At this point in the derivation, we have X-bar Spec-Head-Comp (They = external argument); [They done-Ø DO what] (note VPISH would spell-out as [VP Them do what] prior to NOM case).

(iv)   The light verb v assigns **ACCusative** Case to *what.*   vP

                              PRN   v' (*They* = external argument)

                           what   they   v'

                                        done-Ø….[VP d̶o̶n̶e̶ ̶w̶h̶a̶t̶]
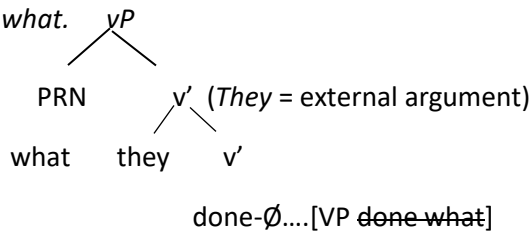
                                            ⇨ ACC case assigned

(v)   *What* raises due to **Edge feature** on vP (hence **multi-spec** position 'What they').

(vi)   *Have* is directly inserted via the lexicon (T-adjunct).[1]

(vii)   *They* raises to check-off features (Case/AGR) via Probe-Goal Union. **Probe T** (features) *have* —which can see through *what* (since *what* has already valued its ACC CASE feature)— searches for its nearest/closest **Goal** (the PRN *They*). (Also EPP can be evoked). *Have* assigns NOM case to PRN *They.* (NOM case assigned from T to Spec-1 of light verb, with subsequent PRN raising)[2].

         (*What* can't raise to edge of TP due to **Mixed Chain Constraint[3].**

(viii)   *Have* raises from T to C due to the null Comp having an affixal tense-feature (cf. (iv)).

(ix)   *What* raises due to edge feature on Spec-CP. (Force-feature or EPP).

1. Whether T is adjunct or not is only a theory internal consideration, nothing hinges on this.
2. Theory internal considerations can both evoke the twin scenarios whereby the(Case-assigned) subject may (i) remain within vP, or (ii) raise to Spec of TP (the latter satisfying EPP).
3. Mixed Chain Constraint bans a constituent from moving into phase, then subsequent movement into a non-phase projection.

Notes

[1] For 'Single pathway mode', otherwise known as the 'Full-listing Hypothesis, as it relates to morphological processing (Storage & Retrieval), see Seidenberg & McClelland (1989), McClelland & Rumelhart (1981, 1988) which represents such full-listing models due to connectionists stance on language (e.g., Seidenberg & Elman, 1999) Networks are not "hidden rules". *Trends in Cognitive Sciences*, 3, 288-289.

(For arguments for a dual-pathway, see Bertram, Schreuder & Baayen, 2000. Gary Marcus along with colleagues Steven Pinker, Harald Clahsen, Peter Gordon et al. also argue extensively for a Dual Mechanism Model. (Also see Galasso 2016).

> • Bertram et als. (2000) The balance of storage and computation in morphological processing. *J of experimental psychology: learning, memory and cognition. 26, 489-511.*

> • Marcus, G. (2001) The algebraic mind. MIT.

> • McClelland & Rumelhart (1988) Explorations in parallel distributed processing. MIT.

> • Pinker, S. (1999) Rules and Rules. Basic Books.

> • Seidenberg & McClelland (1989) A distributed developmental model of word recognition. *P sychological Review, 96, 523-568.*

[2] in Galasso (2016, p. 31) an analogy is given which pins items as vertical processing and category as horizontal—the former, an encyclopedic list of entries [ ], [ ], [ ]…, the latter a spreading (horizontally) of rules [[ ]]…Recursive syntax manifests only via a kind of 'horizontalness' as in uniquely found in language.

[3] For a full description of the 'Four sentences', see Galasso 2019).

Exercise Section: Syntax

⇨ **Ex. 1: Structure**

Introduction: What is Structure and Structure Dependency?

⇨ **See PDF no. 1 'a brief note on structure dependency, binary branching and recursive structure'**

(0) **c-command**

a.
```
            A
         /     \
        B       E
       / \     / \
      C   D   F   G
                 / \
                H   J
```

b.
```
            XP
          /    \
         W      X'
         |     /  \
        Spec  X    Y
              |    |
             Head Comp
```
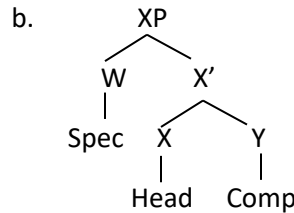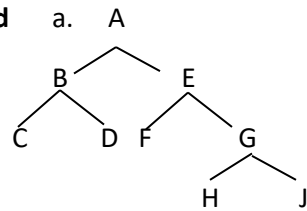
C-Command: A constituent X c-commands its sister constituent Y and any constituents Z which is contained therein. (See examples in notes 0.1, 0.2 below).

Think of a family tree: In the above structure (a), A has no sister constituent (it is a mother and has only daughters, grand-daughters, etc.) and so A doesn't c-command any of the other nodes. (A is equivalent to the (mother) declarative sentence TP). B has one sister E, so B c-commands E and since E contains F, G, H, J, B c-commands those nodes as well. Vice versa: E c-commands its sister B, and contained C, D. B c-commands (daughters) C and D, and (inverse) D and C c-command each other (they are both sisters). (Note that within sister relations there can be no 'word order' hierarchy).

In structure (b), we can say that Spec (W) c-commands X-bar (X') which contains the Head (X) and its sister Complement (Y). The sisters Head (X) and Comp (Y) only c-command each other. Note that for English, structure (b) yields an SVO word order given that English is a [+Head Initial] Language (moving left to right), thus VP=> [V' [V, Obj] in that order. The Spec subject position (on either side of X') is due to either a left or right branching parameter (English being right-branching, Japanese being left-branching).

C-command will become a crucial syntactic configuration as we examine movement.

⇨ Phonology (syllabic structure) abides by the same hierarchical binary branching:

```
            syllable
           /        \
        onset       rime
                   /    \
               nucleus  coda
```
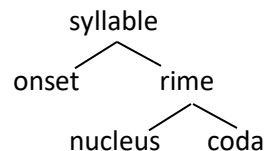
phonological tapping experiments with young children either show phonemic awareness (so the word 'cat' /kæt/ [CVC] taps out three phonemes /k/, /æ/, /t/, or children who are not phonemic tap out twice (at the syllable units) /k/ (for onset) and /æt/ (for rime).

*Note on C-command via X-bar:

The DP 'These stories' doesn't allow Head cliticization)   cliticization (clitics seem to only take pronoun Specifiers: e.g, 'These stories *'ve/have been good').

(0.1)   TP =S(entence) (showing subject-verb agreement requirement)
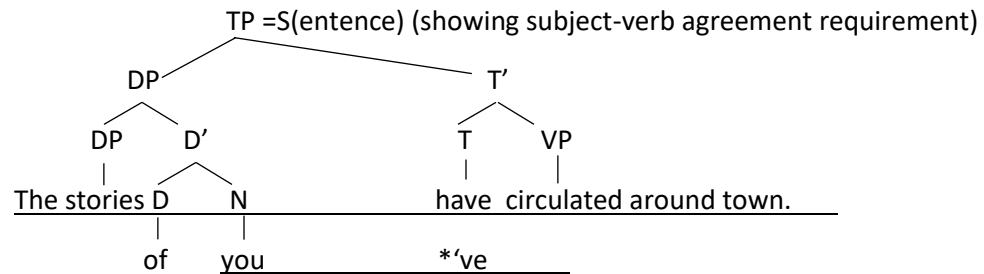
```
              DP                              T'
           DP     D'                       T      VP
        The stories D      N           have  circulated around town.
                    |      |
                    of    you                   *'ve
```

(i)   Clitic {'ve} can't attach to closest stem {you} at PF since LF c-commands doesn't hold. (Syntax at deep level undercuts any surface-level phonological adjacency. (It seems there is tension here between PF and LF— Phonological cliticization is at the mercy of deep level syntax).

(ii)   'These stories' can't attach clitic {'ve} due to the DP 'These stories' not being adjacent at PF (a PF constraint on cliticization). Hence, {have} must be in its full PF form. *(Note: The DP 'These stories' doesn't allow Head cliticization (clitics seem to only take pronoun Specifiers: e.g, 'These stories *'ve/have been good').

(iii)   These friends of the president blame *himself/themselves.

(0.2)   TP   => 'themselves' is c-commanded by 'These friends')

```
            DP
        D       NP                    T'=> ('himself' not c-commanded by Prn 'the president')
     These    N    PP            T      VP
           Friends  P     Prn
                    of  the president  V      prn/reflex
                                      blame   *himself

                                      themselves
```

83

(1) **The Tree Diagram:**

```
                          TP
                     ╱         ╲
                  Spec          T'
                   |         ╱      ╲
                   It      T         VP
                        {affix}   ╱     ╲
                         {s}   Spec      V'
                                     ╱        ╲
                                   V            TP
                                   |         ╱     ╲
                               takes    spec        T'
                                          |       ╱     ╲
                                       money    T         VP
                                                |      ╱      ╲
                                               to   Spec       V'
                                                     |       ╱    ╲
                                                   money    V      DP
                                                            |      △
                                                          raise  a family
```
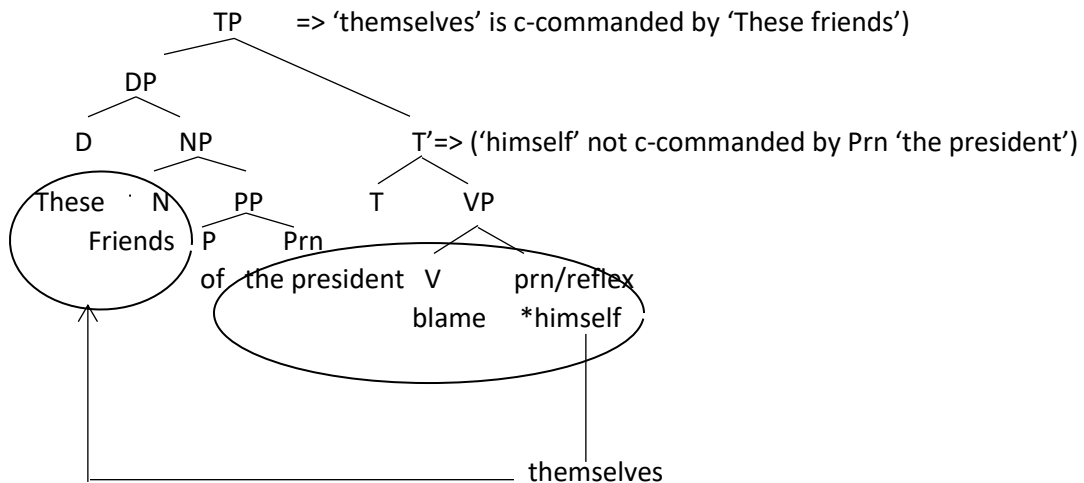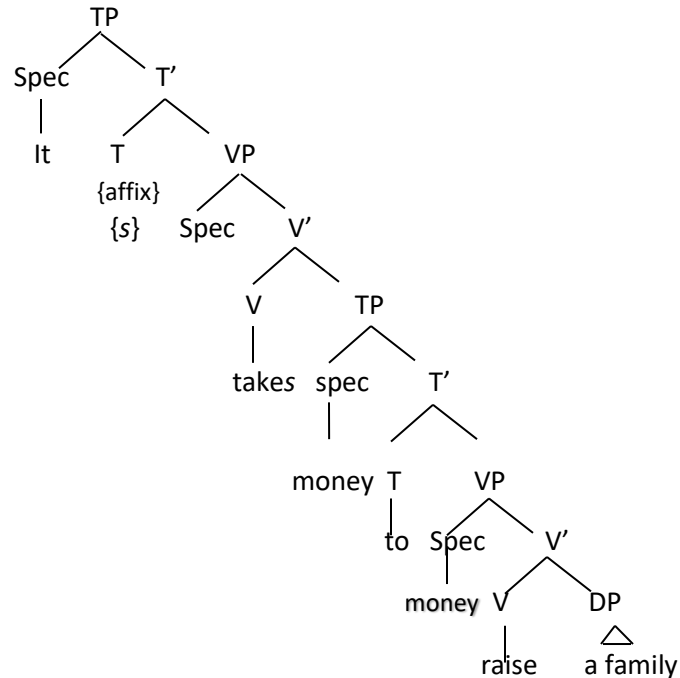
(2)   Subject of VP becomes surface object of TP (by underlying subject of TP): A parallel structure to 'existential expletives' e.g.,

  a.   [TP   There is money [VP money going in the bank]]
  b.   [TP Money is   t   [VP   t   going in the bank]]
  c.   [TP There [T' [T have] [VP arisen several problems]]]
  d.   [TP several problems [T' [T have] [VP arisen ~~several problems~~]]]

One aspect of the seminar will be to drop all orthodoxy regarding traditionally assumed structure, parts-of-speech, etc., and rather to ask the more unassuming yet elegant question of how bare elements of language enter into specific relationships.

For instance, consider the *bare* particles 'to' and 'of' (*bare*—just phonological items without assuming their traditional roles as infinitival and prepositional (respectively). Let's just consider them as neutral particles for now.

First, let's consider 'to' below:

(3) <u>'To-infinitive' vs 'ing-Infinitives'</u> ([__] indicates the position from which an expression has moved):

  a.   ***To*** *raise* a family, it takes money___.   (movement/verbal)
  b.   ***To*** *raise* a family   takes money.   (no movement/nominal)
  c.   *?Raising* a family, it takes money___.   ?(movement/verbal)
  d.   *Raising* a family   takes money.   (no movement/nominal)

84

(We take it that an expression must be nominal if it becomes an argument (without movement)—e.g., b, d).

As is, all four structures are fine (with (?) being somewhat degraded for some). So, 'to' and 'ing' seem optional (both expressions can be verbal/nominal, with/without movement). However, it seems there might be a gradient cline regarding preference when we adjoin 'in order' to the two verbal phrases.

Let's try to flesh it out. Now consider below:
e.   In order *to raise* a family, it takes money__.   (movement/verbal)
f.   ? In order *to raise* a family   takes money.   ?(no movement/nominal)
g.   *In order *raising* a family,   it takes money__.   *(movement/verbal)
h.   *In order *raising* a family   takes money.   *(no movement/nominal)

It seems that the (+adverbial) phrase *in order* prefers to adjoin to 'to-infinitives' over 'ing-infinitives' (e.g., (g) and (h)).

Let's now suppose that 'to-verbs' carry an **adverbial verbal** [+Adv] feature—hence, the acceptability of movement (adverbs being relatively free to move) (cf. examples (a) vs. (c)). Let's also consider that 'ing-verbs' carry a Nominal [+N] feature—hence, their reluctance to move (noting that the 'ing' particle is also gerund forming: Verb to Noun). So, in order to flesh out the marked utterances, consider the following syntax:

(4) **Particle Feature specificity**
    a.   'To' particles (to-infinitives) are +Verbal, plus carry a [+Adv/-N] feature.
    b.   'Ing' particles (ing-infinitives) are +Verbal, plus carry a [+N/-Adv] feature.

If so, let's consider these hidden features a bit more closely:

If 'ing-verbs' carry a nominal feature [+N/-Adv], we might find two consequences to their syntax:
        (i)   They prefer to be nominal arguments over verbal expressions whenever possible (whenever movement operations allow them to slot within an argument c-commanding position. This is what happens in examples (c) and (d)—the expression 'raising a family' prefers to be an argument over remaining a verbal structure, thus the expletive 'it' deletes and we get –e.g.,

                [DP Raising a family [TP takes money]] vs. *[VP Raising a family, [TP It takes money__]]
                (It = 'the raising of a family').

(ii)   They prefer the 'of' particle (which too carries a [+N] feature):
    a.   It's a matter [of [DP raising children]].
          [TP [DP **(The)** Raising **(of)** children] [T' [T is] the matter]]].
          I have a question [PP about [DP the raising of children]]
    b.   * It's a matter   [of [+N] [VP raise children]].
    c.   It's important [to [+V] [VP raise children]].
    d.   It's a matter   [of [+N] [DP raising children]].
    e.   *It's important [to [+V] [DP raising children]].

(Note that (e) works only when the feature specification of 'to' turns from being [+V] to [+N]—
hence, the (phonological) ambiguous nature of 'to' :

    (i)   It is important [**to** raise children] with love.   ('to' is verbal)
    (ii)   Love goes far [**to** the raising of children] (= towards)   ('to' is nominal)

So here, we can tease apart in an unorthodox way how the Prep 'of' is [+N] and enters into a structural relationship with DPs, or even verbal expressions such as 'ing-infinitives' since they too carry an [+N] feature.

(5) **Particle summary:**

    a.   'to' carries a [+V] feature, otherwise known as 'to'-infinitive.
    b.   'for' carries a [+N] feature, otherwise known as a Preposition.

(Note: whenever 'to' serves as a preposition, it carries an [+N] feature, as found in ex. (e) above).

Hence, we can arrive at traditional parts-of-speech via syntactic substitution tests on structure.

of (the) raising (of) children => 'of raising children'   **to** raise children.

a.   PP   b.   T'
         /   \   /   \
        P   DP   T   VP
        | /   \ | / \
     [+N] D   NP   [+V] V   N
        | | / \ | |
       Of | N   PP   **to** raise children
        | the | / \
        | N   P   DP
     about raising | / \
        | of D   N
       to |
                   children
    (towards)
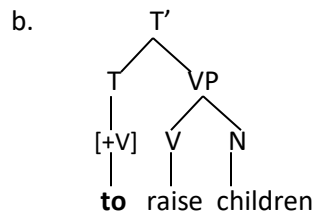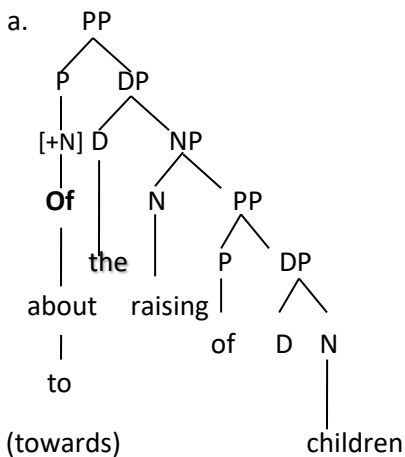
⇨ **Ex. 2 What is Morphosyntax? What is Movement?**

Understanding the internal structure of the word (heyday of Skinner vs Chomsky).
· The Dual Mechanism Model.
· 'Fascinating' vs. 'celebrating' (vertical vs. horizontal processing) language types/processes.
· decomposed vs. undecomposed morphologies
e.g. 'shopping' (N, Adj, V processing): Derivational vs. Inflectional Morphologies e.g., 'How do you do?' (where the two 'do' verbs have different lexical entries: aux vs. main verb—Substitution Test).

Phrase structure rules.

Evolution of tree diagrams:
1. **From NP to DP:**

NP   =>   DP (Abney 1987)[2]
/\   /\
D   N   D   N
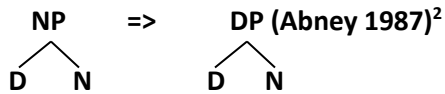
2. **From DP to Bare Phrase (Chomsky 1995)[3]**

DP   =>   the
/\   /\
D   N   the   book...*is on the table.*
| | | |
The   book   [+D>N]   [+D<N]]
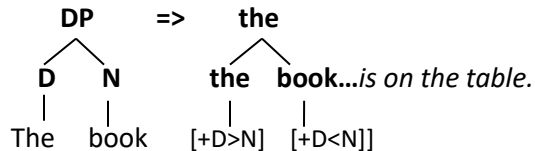
1. Abney reminds us that not all NPs nicely package under a uniform x-bar theory. (The class will stay with Abney's analysis of labeling).
2. Chomsky states that 'language design' is perfect: all features (e.g., X-bar, c/s-selection, features) come out from the lexical element itself. No new objects are added in the course of a computation. For example, the DP [The book] is derived by the very nature of the lexical item 'the' itself (with its internal bundle of features already in place). No need for indices, x-bar theory, etc.

e.g., [NP [John's [VP quietly having read the book]] has all the trapping os a verb phrase with no apparent Noun head. Hence, Abney analysis such NPs as DPs which are headed not by nouns but rather by determiners, this expanding the tree (to allow for the co-occurrence of a possessor and a determiner within one NP (DP): e.g. 'La mia mamma ' (It. The my mother). In English possessives and determiners are in complementary distribution—e.g., * 'John's the book'.

---

[2] Abney (1987). *The English Noun Phrase and its Sentential Aspect.* Ph.D. dissertation. MIT.
[3] Chomsky (1995). *The Minimalist Program*. (Ch. 4). MIT Press.
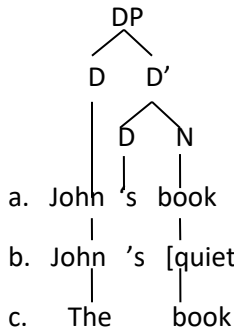
DP   both possessive ['s] and determiners [the] occupy the head of DP

D   D'   with both taking a noun (phrase) complement [book].

D   N

a. John 's book

b. John 's [quietly having read the book] = > gerund NP.

c. The   book

Also, a DP analysis allows us to maintain that functional features (AGR/Case) remain under a functional node (D).

3. AUXP => TP

VP *(Syntactic Structures 1957)*   T'

AuxP   T   VP

Aux   V   {s}   spec   V'

[3/pres/sg]   V

{s}   [[walk] s]   [[walk] s]
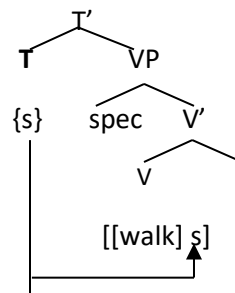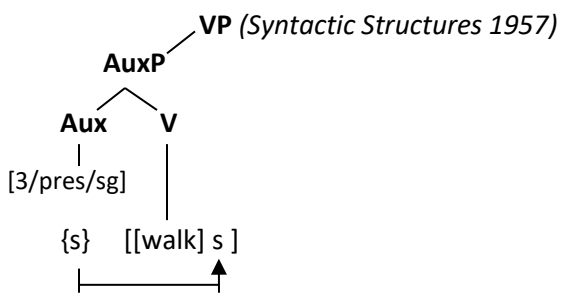
What was nice about the out-dated AuxP analysis is that it nicely captured the functional to lexical relation between D to N and Aux to V. We know that D always introduces an N, like how Aux introduced a V: Aux to V relation serves both to host modal/Auxiliaries as well as verbal inflectional material: 'will walk', 'to walk'

e.g.,   a. John [AuxP [Aux *will/does*] [VP walk home]].

b. John likes [AuxP *to* [VP walk home]].

The class adheres to Abney's DP analysis while maintaining the CP>TP>VP tree.

Comparing French vs. English: Verbal Morphology (taken from Howard Lasnik's 'Verbal Morphology: Syntactic Structures meets The Minimalist Program' (p. 263, 'Minimalist Syntax: essential readings)'

⇨ **Ex. 3 Parsing Morphosyntax/Structure Dependency: Some thought-experiments**

· Processing mistakes and slips are systematic, not random?

Target: what about tacos night?

Slip: What about taco_ tonight_s?

Processing/parsing: [[What]] [about] [[taco] s] [[tonight]].

Wh-word 'What' allows for a morpho-bracket frame/parse as [[ ]], as in e.g., [[What]'s up?,

Q-1: Try to tease out the other morpho-brackets per words: about, taco, tonight…

R-1 (from above): [[what]_] (as in [[what]'s], [about], [[taco]_], [[tonigh]_] (as in [[tonight]'s] party)…

*NB. Historical, wh-words might have been broken accordingly:

[Wh[at]] > [th[at]], [[wh]en]>[the[en]], [wh[ere]]>[the[ere]], etc.

The word [about] is not processed as 'where-abouts' *[[about]s] since 'where-abouts' is a formulaic idiom.

Adverb (19th century) 'Tonight' was once composed as preposition 'to' + Noun 'night: [to [night]].

Q: Do trees come from words, or do words come from trees? => Structure Dependency (so words come from trees (?).

Chomsky (1995).   Arguments for 'bare structure' stems from the notion that words get pulled from out of the lexicon with all the syntactic properties via design (the structure surrounding a lexical item is made precise by the syntactic structure it occupies and therefore it doesn't need to be restated as part of the items selection properties. For instance, the AUX 'will' gets pulled from the lexicon stating that it occupies a T label (T-bar/Tense) , an intermediate projection.

Movement: French v English Overt v Covert movement

Difference between French vs. English verb morphosyntax is a matter of *overt vs. covert* movement.

However, there are two theories for French verb movement:

(i)   French main verbs come fully inflected as a result of there being no bare stems.
     So French main verbs come fully inflected (undecomposed) and directly insert under T. (Lasnik),
 or,
(ii) French overt V-to-T movement (decomposed) (Chomsky).

The structure in (1) below shows French V-to-T movement across NegP.

(1)   [aime]   TP   Jean (n') *aime pas* Marie. (John doesn't like Mary).

```
              TP
        Jean      T'
                T     NegP         (see French vs. English movement in (4) below).
              ne- aime   pas   Neg'
                          ne+v        VP  => (VPISH: Verb Phrase Internal Subject Hypothesis)
                                  spec        V' => [-BVS]. There are no [+BVS] verbs in French.
                                                            *aim
                                  Jean    V      DP
                                         aimer   Maria          (NB. BVS = Bare Verb Stem)
```
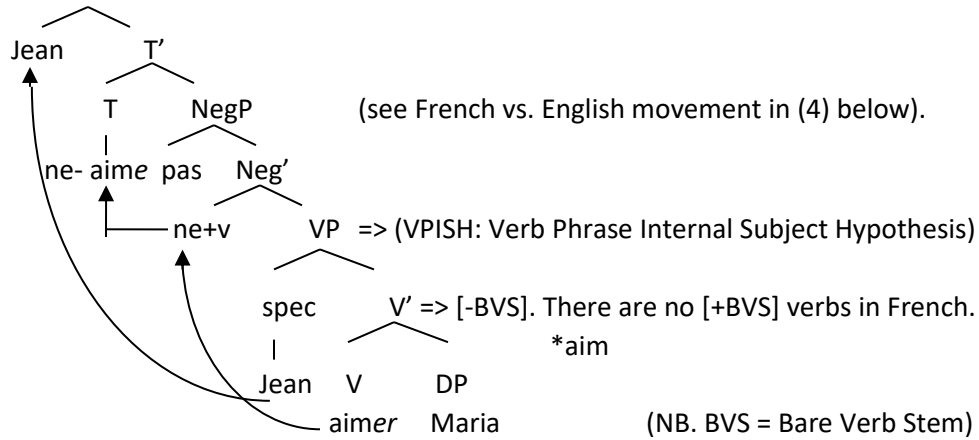
Note English main verbs don't situate above NegP (since they pull out of the lexicon without inflection and must process via decomposition of affix hopping (covert affix lowering from T to V). (See (2) below).

*John[T' [T likes]] not Mary. *John *walks not.*

  John [T' [T {s}...(NegP) [VP [V like-s]]]] Mary

  John [T' [T does]] n't *like Mary. John hasn't walked.*

English Auxiliary verbs 'Do, Be, Have' (and Modals: *will/would, shall/should, can/could,* etc.) like their French main verb counterparts also directly insert under T (without raising), hence, the grammaticality of 'John *does/is/has* not walked' (similar to what we find in French main verbs e.g., *Jean (ne) aime pas*. ('John likes not').

Max projections come from design forcing a Spec of TP. Hence, top-down syntax (design) contributes to the lexical items' properties. So, words come from trees.

Other evidence for this is the well-known stated axiom of structure dependency. (cite examples).

As an example, when one says 'walks' there is a decomposed nature of the word which generates the tree below:

Using the tree to derive movement applications.

English 'walk*s*' / French 'marche' decomposed   vs.   French [march*e*] undecomposed (Lasnik).

(2) [[walk]s]   TP   {2') [marche]   TP

                         T'   T' => direct insertion

                 T   VP   T

                 [*s*   V'   march*e*

                 | [walk]*s*]   => covert: affix lowering
          [marche |
                 [march*er*]]   => overt: verb raising
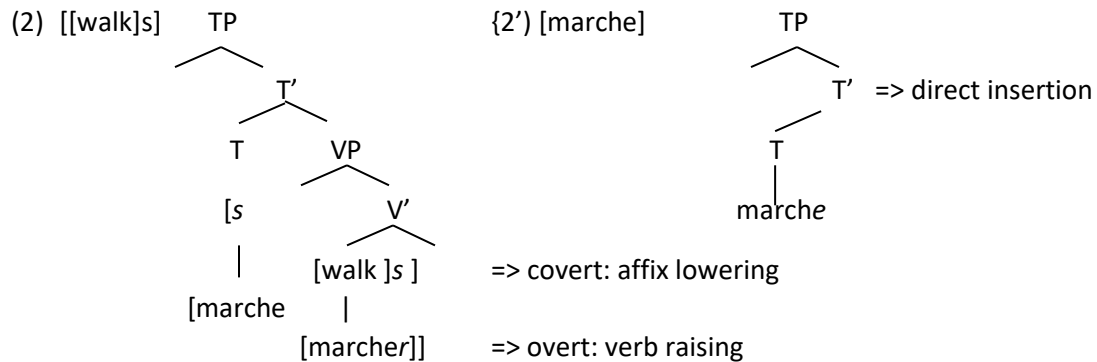
              [[celebrat]ing] : decomposed   [fascinating] : undecomposed

(3)   T'   (3')

            T   AspP (she is 'celebrating' her BD)   AdjP   (this is a 'fascinating' class)

         is   asp   VP   adj   N

               [ing   V'   [fascinating]   class

                         V
                    [celebrate]]
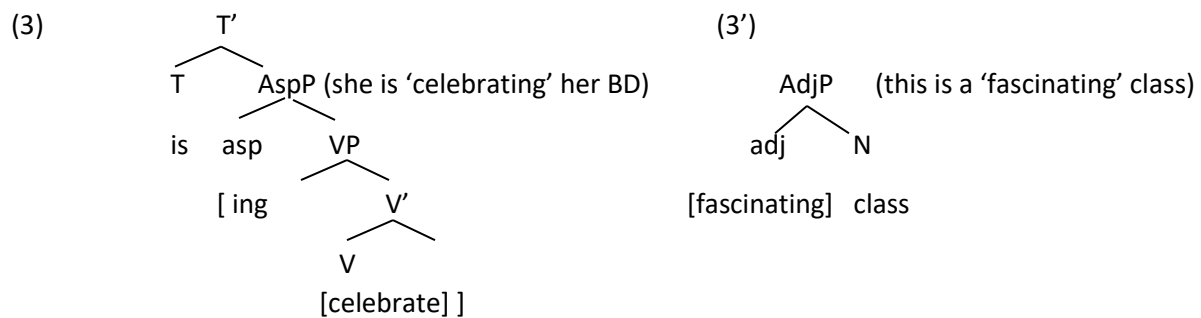
    See affix hopping ex. (2): INFL on English Verb lowers to main stem.
    Classic example of affix hoping:
    Phrase-Structure (PS) rules: (Chomsky: *Aspects of Theory of Syntax*)
    Aux→ Tense (have + en) (be+ing)
    AuxP begins with obligatory Tense, (then optional perfect/progressive aspect markers).
    e.g., 'The man s have + en be+ ing read the book'

    INFL/Move; PS affix hopping places affixes in correct position:
    e.g., The man [_ [have] s] [_ [be] en] [_ [read] ing]] the book
        (The man   has   been   reading   the book)

    Non-INFL/Non-Move:
    The man is in the [[reading] room].
    Note distinction between derivational adjectival [reading] vs. inflectional [[read]ing]. The tree
    structure can capture a decomposed vs. an undecomposed analysis.

Ex. 4 Testing Structure

⇨ **French** vs **English main verbs** (also see Italian below).

(1) [aime]   TP   Jean (n') *aime pas* Marie. (John doesn't like Mary).

                  T'

            T   NegP => (note word order: 'pas' is in spec of NegP, 'ne' is Head of NegP)

    Fr.   ne-aime pas   Neg'   'ne' must raise with verb to T.

                   ne-   VP   => nonfinite verb 'aimer' must raise to T to get tense.

                          V   DP

                        Aimer Maria

    *Eng. likes   **not**

           does   not   like   Mary
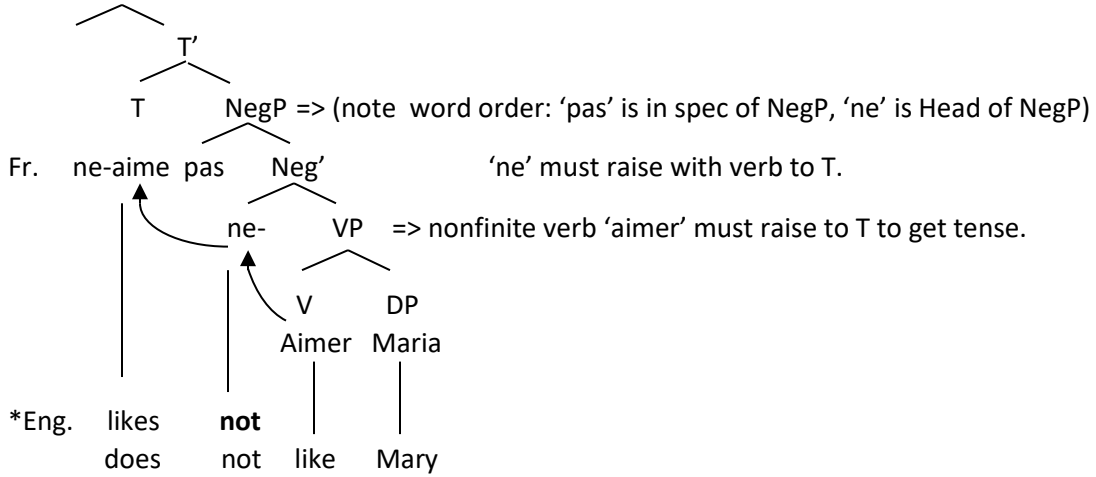
⇨ **A note on Italian Clitic Climbing: The NegP relation to clitics.**

      An argument for treating clitics as a TopicPhrase TopP (for Focus).
      An [f] Focus feature motivates checking of host Clitic . Clitics behave as adjuncts hence their climbing /various adjoin. TP = Tense Phrase, CleftP (Cleft Phrase (clitic)).

(2)   TP   ({e̲} is [-BVS] morpheme. {r} is infinitive marker).

  Spec   CLeftP

g. lo   spec   CL' =>   [f] focus feature which ensures movement: S-O-clt-V word order

    f.   me   clitic   T'

           [f]   T   VP => [+BVS]   Note: {e̲} is [-BVS] morpheme. {r} is [-Fin] marker).

        e.   lo   spec   V'

             devo   V   CLeftP   (= TopicP) => 'clitic climbing' focus and verb selection.

                b. dover spec   CL'   (=FP) => *lo* as strong pro-clitic

                        clitic   VP [-BVS]

                        [f]   spec   V'

                      d. lo   V   clitic

                         a.   dover̲e   ~~lo~~

                    e. f. g.   fare
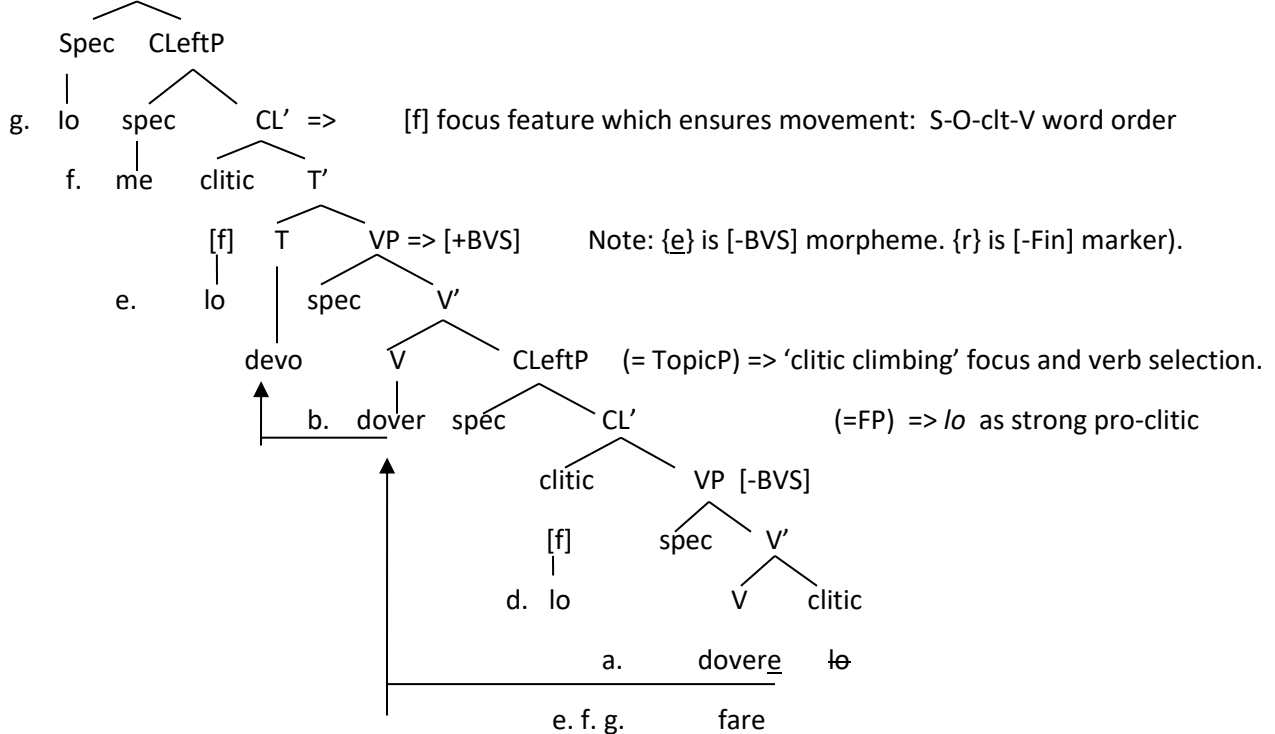
Io me lo devo fare   Note Italian overt verb raising: (like French)

'I to-me it have to do '   {finite}   /   { non-finite }

[T *devo*, / [+ bvs *dover*, [-bvs *dovere*]]]

Infinitive/[-BVS] can't host clitics. Infinitive/[+BVS] stems can host clitics

a. *dover<u>e</u>lo (clitic can't attach to [-BVS/VP] stem).   d. [lo [VP dovere]] (clitic can raise above [-BVS]

b. [doverlo]. 'dover(e)' must raise ({e} deletes).   e. [Lo [TP devo [VP fare]]] (It I have to do).

c. *lo dover (clitic can't raise above [+BVS/VP] stem.   f. [Me [lo [devo [fare]]]] ((I) to-me it have to do)

                                                          g. Io me lo devo fare (Io ('I')).

Why pro clitic movement?
1. Let's first assume that all phrases must be of a full Spec-Head-Comp configuration (they must fully expand).

(3)   XP
                          /   \
                      Spec   X'
                               /   \
                            Head   Comp

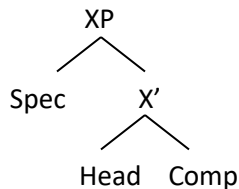Chomsky proposes that this is what is behind the notion for both an **EPP feature***, the **(Extended Project Principle)** and the **Edge Feature**, the former extending T' into a full TP and the latter expanding a C' into a full CP. Yes-No questions too require full CP (albeit via a phonological null Spec with abstract [Q] feature).
(***EPP** that all declarative phrases/TP must have a subject-filled Spec position).

Declarative sentence: ** Does John go home.* This is an improper construct since the Spec of TP must be filled by closest argument (**A-movement**). Note VPISH (verb phrase internal subject hypothesis). Recall, one argument for VPISH is theta markings must be done within VP.

(4) Under-expanded TP/declarative   Fully-expanded/declarative
            * T'   TP   EPP motivates subject raising out of VP.
         /   \   /   \
        T   VP => VPISH   spec   T'
      does   / \   [EPP]   /   \
           Spec   V' | T   VP
            | / \   John   does   / \
          John   V   N   spec   V'
                go   home | / \
                                                              John V   N
                                                                   go   home
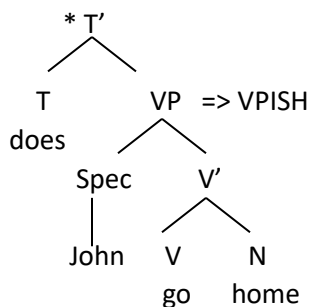
93
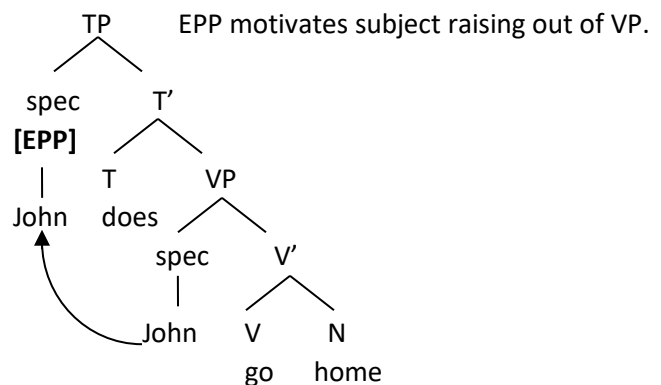
The same logic applies for CP.

Wh-movement: *Main Clause* Questions contain a force **[Q] feature** in Spec of CP, turning TP declarative into CP interrogative. The Head of CP houses two features **[Tense]** (making it a main clause) and an **Edge Feature [EF]**. Tense in C requires that the verb raises out of T. The [EF] feature must be checked by an argument in spec of C (thus ensuring A-movement of wh-word). The feature [T] gets checked-off via Aux-inversion. However, the [EF] feature is still not checked. The wh-word 'where' must raise to check [EF]. **Wh-words** carry an embedded **[Q] feature** (question) motivating movement.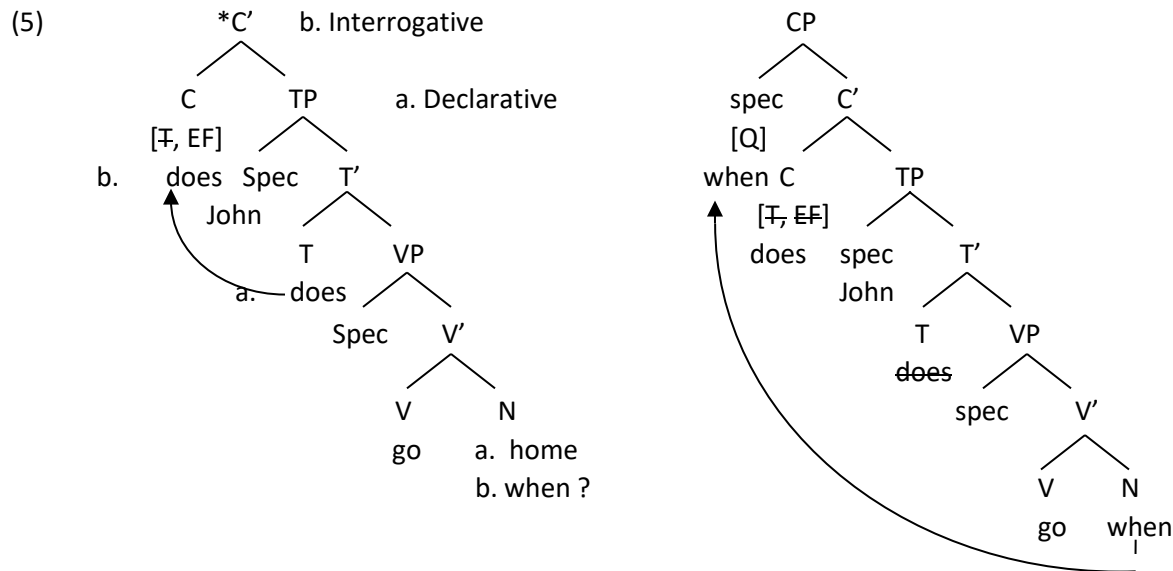 The noun 'home' doesn't carry a [Q] feature thus ensuring no Complement/Object A-movement of the noun in this environment.

(5)

2. **Why MOVE?** If all phrases must be full expansion XP structures, then one consequence of movement is that it ensures full phrase expansion (EPP, EF). A second motivation for movement is to check-off formal functional features. Typically, movement into higher (functional) projects provides lower (lexical) elements with formal features regarding

    (i)   Discourse,
    (ii)   Syntactic

These two properties motivate movement. In addition, we can assume that **Focus** is one such discourse property, labeled herein as an **[F] feature**.

Back to clitics.   Let's assume that there is a **[f] focus feature** found within clitics (of CLeftP) which is (partially) responsible for clitics to move up the tree In accordance with full XP expansion). Hence **clitic climbing** could be viewed as something similar to what we find with **Wh-movement**, where the Wh-constituent must move for reasons of focus. Recall, in Chinese wh-pronouns do not have such a [f] features, hence the wh-pronoun is allowed to remain base-generated within VP—e.g, 'John (does) go when'?

So, the typical analysis that CP houses a [Q] feature for **Interrogative Force** can be extended to include an [EF] **Edge feature** (similar to EPP for TP) coupled with a [f] **focus feature.** Wh-words must move to CP in order to check both EF/F features. We can treat (*syntactic*) clitic movement much in the same manner (though there are *morphophonemic* considerations: weak stress, bound morpheme).

Note how in the wh-pronoun 'who' in the sentence 'Who called the police' must raise out of spec of TP and situate within spec of CP due to such feature checking. No verb raises to [Tense] of C since Tense here is affixal in nature. The wh-pronoun 'Who' is already tensed (virtue of TP) and so satisfies the checking off of the Tense–feature in C.

(6)

```
                        CP
                      /    \
                  Spec      C'
                  [Q]      /   \
                  Who     C     TP
                        [T̶, E̶F̶]  /  \
                          Ø    Spec   T'
                               who   /  \
                                    T    VP
                                 {affix} /  \
                                  {ed}  Spec  V'
                                             /  \
                                            V    DP
                                        call-ed  the police
```
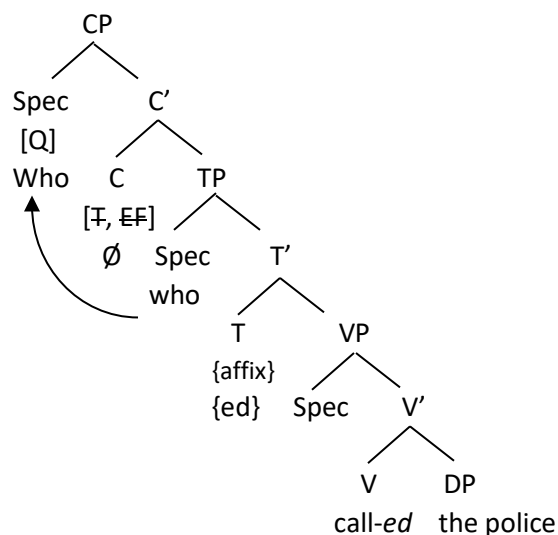
In summary, one possible account of clitic climbing is to check-off a focus feature. But there are constraints on clitic movement. Once such constraint is that clitic movement is blocked by Negation Phrase (NegP).

Some Data: Negation Phrase 'non…mai' is used in movement analysis.
a. Vorrei   [non   doverlo   mai fare]   => infinitive / [+BVS]
    (I) would-want   not (to)have-to it ever to do.
    'I would like to not have to ever do it'

b. Vorrei   [non dover   mai   farlo]   => {e} deletes with clitic 'lo'
    (I) would-want   not (to) have ever (to) do it.

c. *[Lo vorrei   [non dover   mai   fare]]
    (I) It would-want   not (to) have never (to) do

d.   [Lo vorrei   vedere]
     (I) It would-want   (to) see.

95

Phases.   Chomsky has recently proposed that chunks of certain constituents (what are called phases) get sent to **transfer** at certain points in the **derivation** along the syntactic tree projection. Movement is said to delay transfer (e.g., *bottle of wine vs. wine bottle*). The Phases are vP (*light verb* just below NegP) and CP. VP, NegP and TP are not phases and thus must either allow movement up the tree to keep the derivation alive or get immediately transferred to spell-out .

While (d) allows TopP fronting of clitic 'Lo', in (c) above NegP is a **merge constituent** which blocks clitic raising to a TopP since transfer to spell-out has already occurred. NegP blocks movement out of phase and secures transfer. Move/Phases are vP and CP and may secure transfer to spell-out. Merge/non-phases are VP and NegP-TP which may allow for further movement up the tree. **Merge>Move>Merge>Move (MeMo)** must follow—a sequence roughly mapping what Chomsky proposes of 'non-phase-to-phase' sequences, e.g., VP/merge>vP/move>NegP-TP/merge>CP/move. Two Negations are proposed for Italian: one lower lexical Neg 'mai' with scope over the lexical verb, and a higher functional Neg 'non' with scope over Tense. 'doverlo' must raise to situate in a proper T node of a FP/TP max-projection—e.g., 'Non doverlo'.
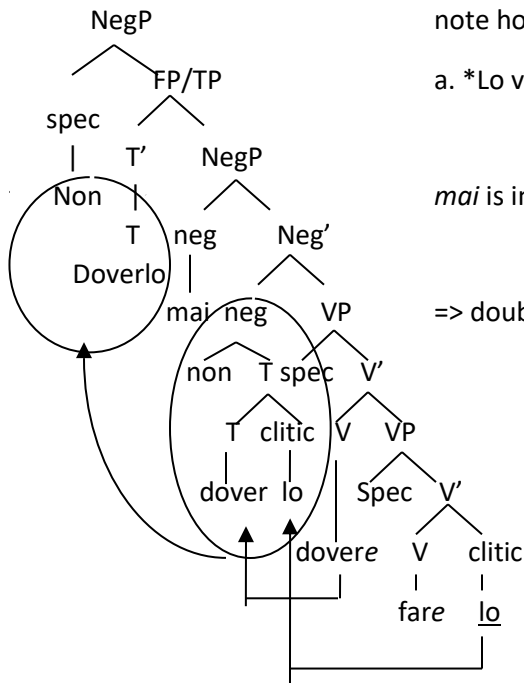
(7)   NegP   note how NegP blocks pro-clitic climbing:

             FP/TP   a. *Lo vorrei [non dover mai fare].

spec
   | T'   NegP
  Non   *mai* is in Spec of NegP, *non* is in Head of NegP
      T   \neg   Neg'
  Doverlo |
            mai neg   VP   => double infinitives: 'dovere', 'fare'. Dover raises to T/NegP
            non T spec   V'
                 T clitic V   VP   {e} [-BVS] morpheme is deleted with clitic. 'Dover/far'
                 dover lo   Spec   V'   with {r} as marking infinitive 'non dover'
                          dover*e*   V   clitic   ('not to ever have to')
                                      far*e*   lo

Summary: Understanding Focus. Topicaliation, Cleft sentences. Topicalization (TopP).
     a.   I have been having trouble with the syntax seminar.

          a' 'That kind of class', everyone seems to have trouble with [that kind of class].

     b.   John has fallen in love with Mary.

          b' 'That girl', everyone falls in love with [that girl].

Topicalization allows **fronting** of old, previously stated material for **focus**.

Cleft Sentences.

   a.   It was [syntax$_i$ that [he hated most syntax$_i$]]   ('It was syntax that he hated most')

   b.   It's [freedom $_i$ that [we want to have freedom $_i$]]   ('It's freedom that we want to have')

Note: movement of cleft constructs allows Objects to raise up above matrix subject for focus.

Also note movement within relative clauses:
   a.   **The cat** that [__chased the dog] was very big. Subject movement.
   b.   The dog that [the cat chased__] was very big.   Object movement

*As a footnote: **Broca aphasics** have a harder time dealing with object movement relatives and their interpretations whereas subject movement relatives get interpreted at chance. Broca aphasics have difficulty dealing with long distance movement over short distance movement (Grodzinsky).

Note the subject 'the cat' raising through the various projections, ending in CP for Focus.

(8) 'The cat that chased the dog'   CP   'The cat' (Focus Phrase /CP)

97

(9) The dog that the cat chased.　　CP

```
                          CP
                  Spec        C'
              The dog      C          TP
                        that   TP    spec      T'
                            Spec   T' the dog  T      VP
                            the cat  T   VP   [+past]  Spec    V'
Obj move                      {ed}            the dog   V    Adj
                              V   DP           V      big
                           chased  the dog   was
```
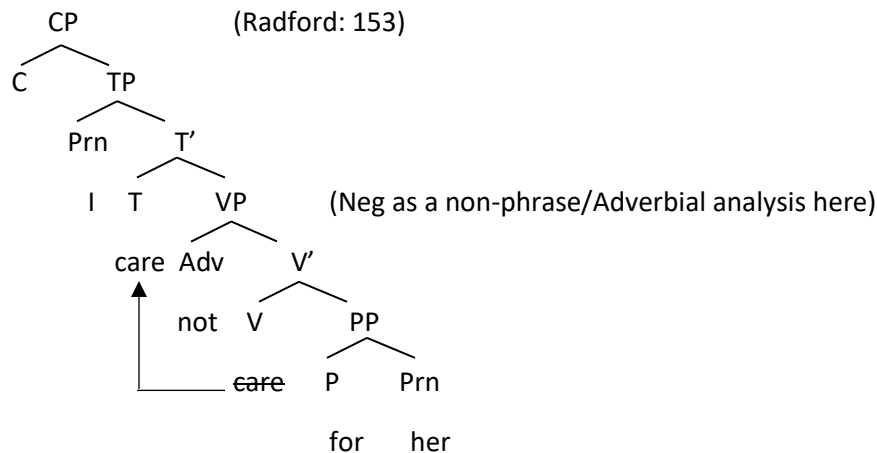
Obj move
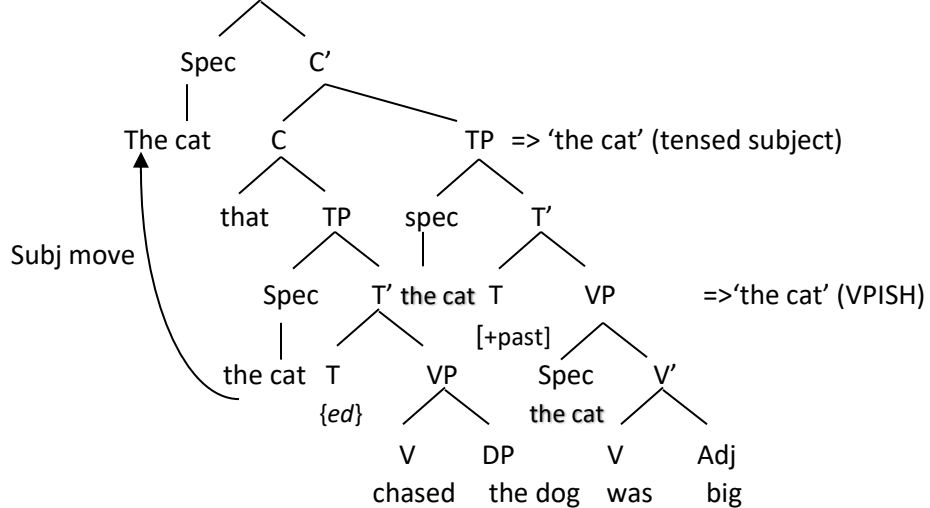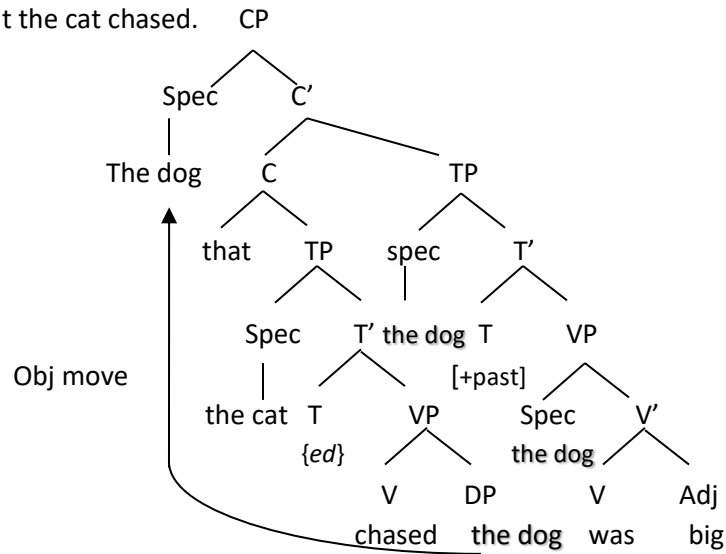
⇨ **Ex. 5 Head-to-Head Movement**:

(1) **Head Movement Constraint/HMC:**

Head movement is only possible between a given head and the head of its complement.
In other words, locality is a condition—namely, a head can't skip an intervening head on its way to a higher-up head of a functional phrase (FP).

(2) V –to – T (Elizabethan English, Early Modern English (EME))

'I care not for her'

```
        CP              (Radford: 153)
     C      TP
        Prn     T'
         I   T     VP         (Neg as a non-phrase/Adverbial analysis here)
           care Adv      V'
               not   V      PP
                    care   P     Prn
                          for    her
```

In EME grammar, English main verbs were 'strong' (a specification based on an [+/-] inflectional parameter. For instance, in 17th century Shakespearean English, we find the following rich inflectional paradigm: second person/singular {-st}, third person/singular {-th}, {-s}, and third person/plural {-n}:

(3)   a. Thou *sayst* (you say)

          b. love *feedeth* love (love feeds love)

          c. It *looks* ill (preserved in Current English (CE))

          d. Their lips and laugh *waxen* their mirth.

Hence, from V-to-T, EME allowed continuation of Head Movement from T-to-C (maintaining the Head Movement Constraint (HMC)):

'Care I not for her?' (EME)

(4)   CP   (Radford: 156)

('not' as Adverb analysis: see (6-7) below for NegP analysis)
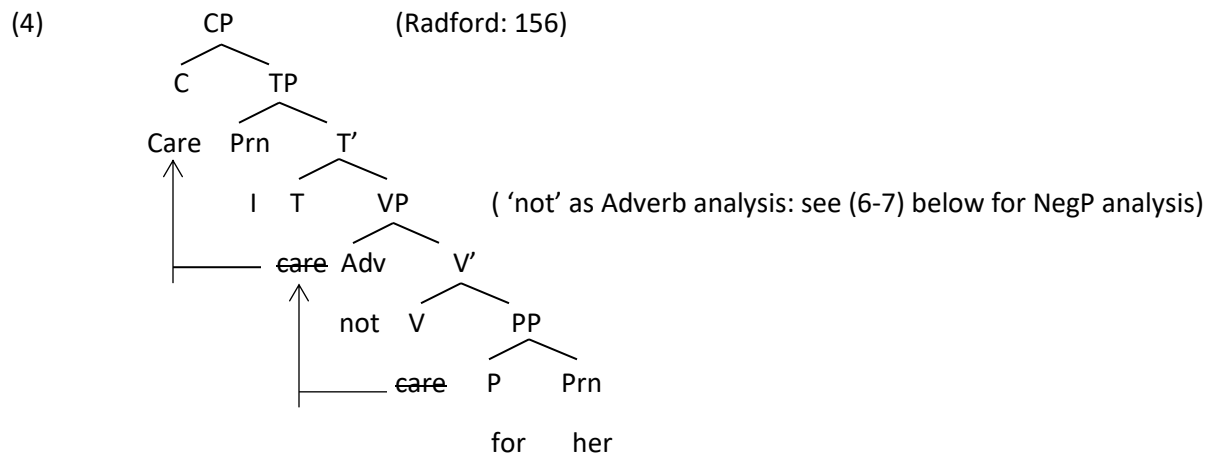
(5) **EME Negation: an NegP analysis**
     A lord *ne* hath *nat* al of gold (AR p. 164)
     (A lord has not all of gold)

     Negation: ne+verb+nat (not)

(6) **Shouldn't Negation project its own phrase, a NegP?**

Now dispensing with the *Neg-Adverbial analysis*, we can project negation as its own phrase, with a proper Spec-Head-Comp configuration. Counter to what was assumed in EME, where 'not' (*nat*) was placed within Spec of NegP, current analyses today place 'not' in Head of NegP (with a null spec):

(7)   NegP

               spec   Neg'

                Ø   Neg   VP …
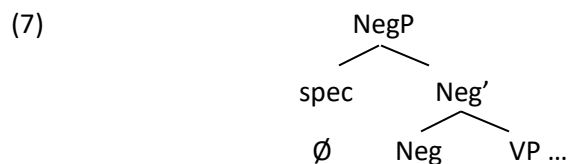
Reduced to: NegP
                                  ╱   ╲
                               Neg   VP…
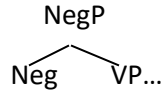
(8) {ne+V} could be seen as a complex-head drawing the verb upward and attaching onto {ne}.

    By Shakespeare's time, {ne} had dropped out of use, leaving the Head of NegP null (available for
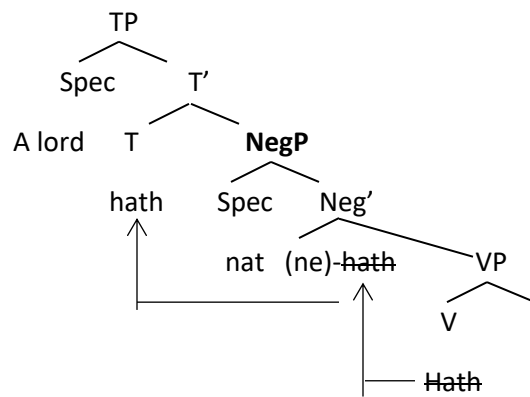   Head movement, as shown below).

    A lord hath (ne) nat every vessel of gold

    Similar to French {ne…verb…pas} whereas the French {ne} has dropped out of use.

(9)   TP
                            ╱   ╲
                        Spec   T'
                                ╱   ╲
                   A lord   T   **NegP**
                            ╱   ╱   ╲
                        hath   Spec   Neg'
                          ↑   ╱   ╲
                        nat   (ne)-~~hath~~   VP
                                       ↑   ╱ ╲
                                      V
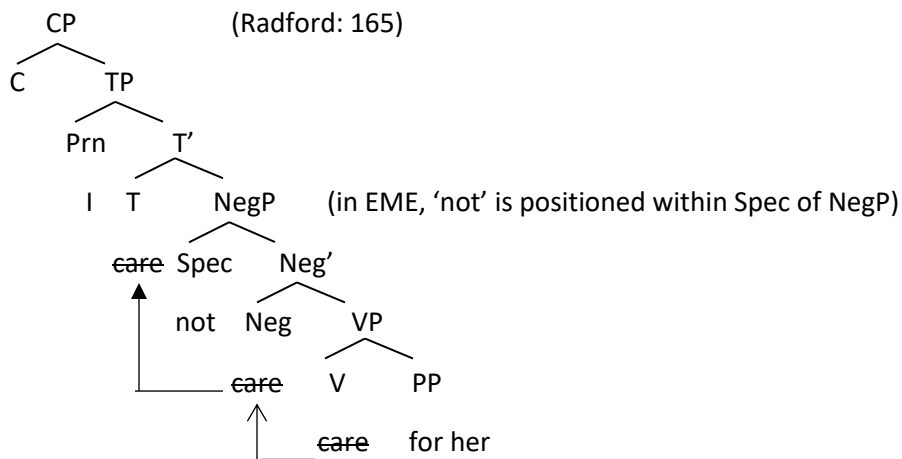   ~~Hath~~

(10)   **Arguments for NegP**

Moving away from an Neg-Adverbial analysis and projecting a proper NegP ('not' is positioned within
spec of NegP for EME) made available a Head slot within NegP for Head to Head movement, maintaining
the Head Movement Constraint: Head of V, to Head of NegP, to Head of TP(and all the way to Head of
CP for interrogatives):

(11)   CP   (Radford: 165)
              ╱   ╲
           C   TP
                  ╱   ╲
               Prn   T'
                      ╱   ╲
                  I   T   NegP   (in EME, 'not' is positioned within Spec of NegP)
                     ~~care~~ Spec   Neg'
                       ↑   ╱   ╲
                     not   Neg   VP
                          ~~care~~   V   PP
                             ↑   ~~care~~   for her

                              100
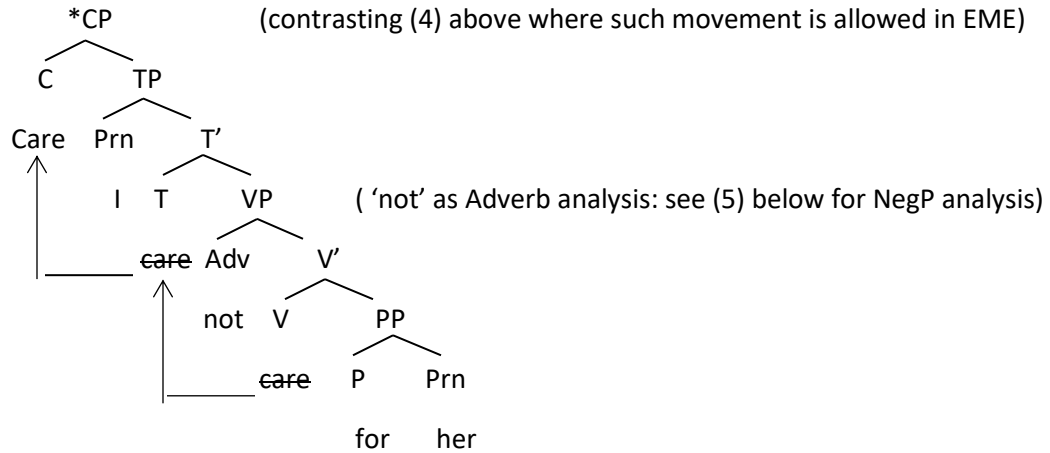
In Current English (CE), note how the verbal morphological paradigm ({-s} only shows on 3P/singular) has become reduced (shifting the parameter of [+INFL] for EME to [-INFL] for CE, and therefore disallowing V-to-T and as a consequence V-to-C movement (HMC).
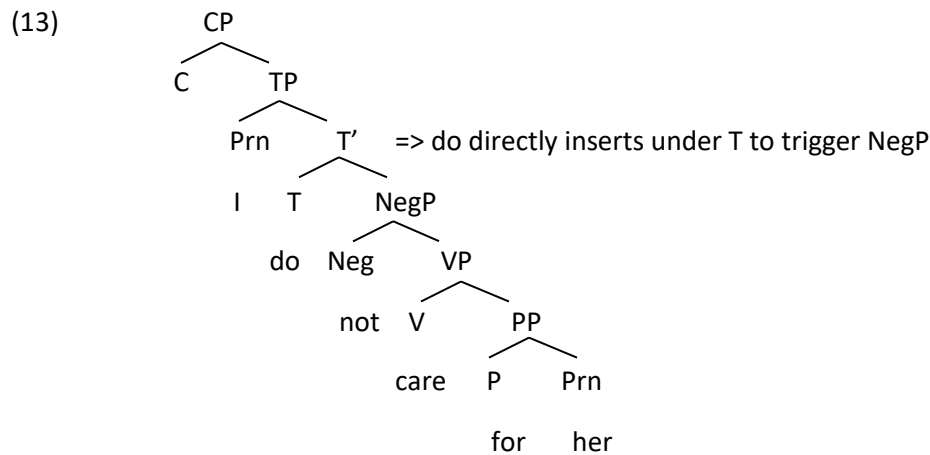

* 'Care I not for her?' (CE)

(12)   *CP   (contrasting (4) above where such movement is allowed in EME)
          ╱   ╲
         C   TP
               ╱   ╲
       Care   Prn   T'
                    ╱   ╲
                   I   T   VP   ('not' as Adverb analysis: see (5) below for NegP analysis)
                      ~~care~~ Adv   V'
                            ╱   ╲
                          not   V   PP
                             ~~care~~   P   Prn

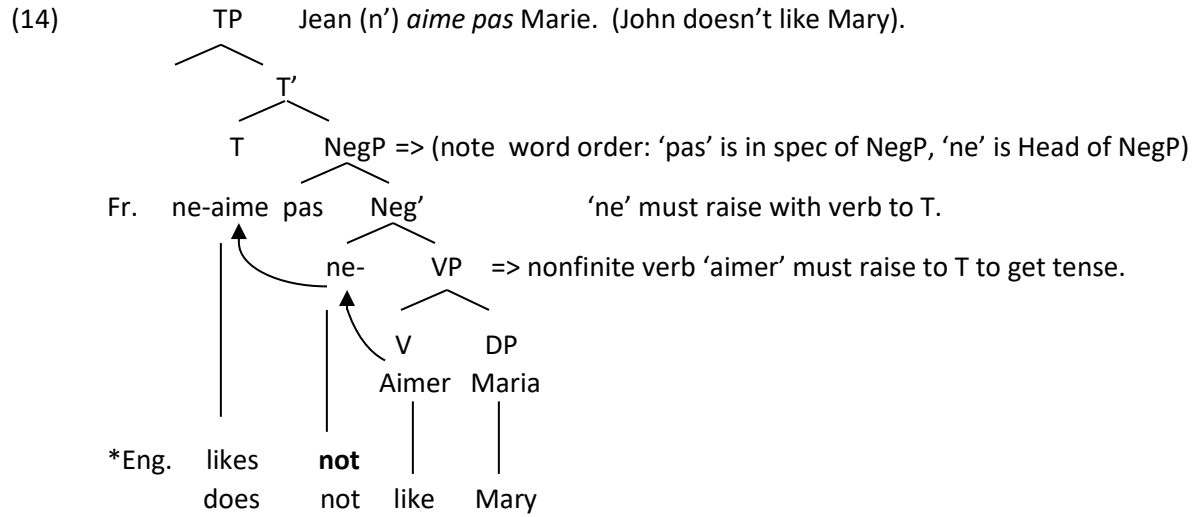   for   her

In (12) above, **Do-support** is required to trigger negation (shown in (13) below):
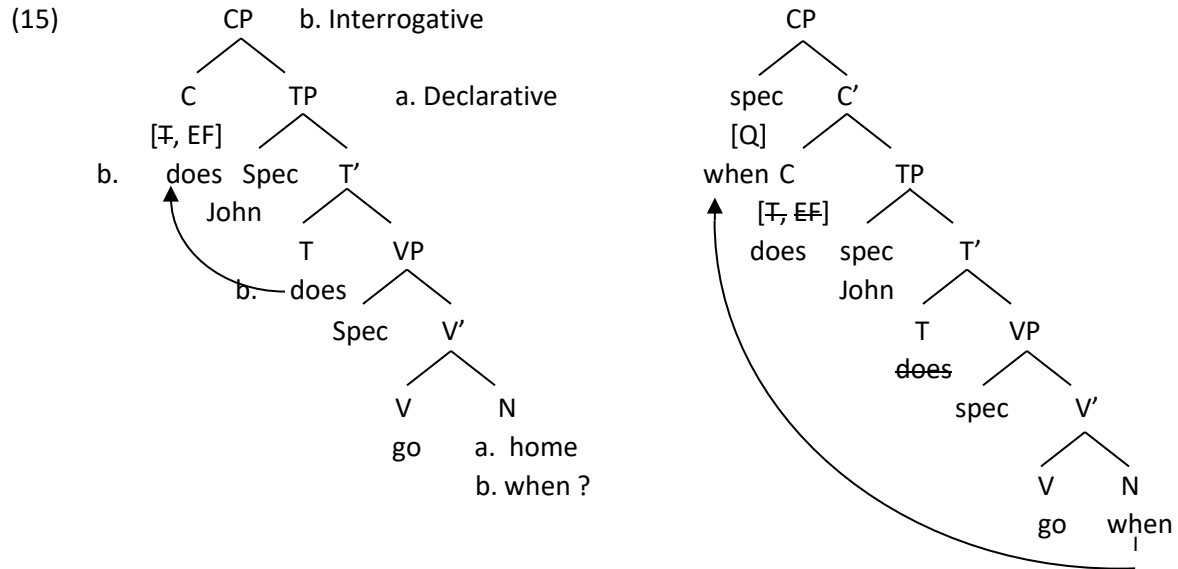

Negation trigger: [TP [T Aux] [NegP [not] [VP verb]]]. Most analyses today place 'not' within the Head of NegP.

(13)   CP
          ╱   ╲
         C   TP
               ╱   ╲
             Prn   T'   => do directly inserts under T to trigger NegP
                   ╱   ╲
                  I   T   NegP
                      do   Neg   VP
                          ╱   ╲
                        not   V   PP
                             care   P   Prn

                                    for   her

French vs **English main verbs**

(14)   TP   Jean (n') *aime pas* Marie. (John doesn't like Mary).

                      T'

            T   NegP => (note word order: 'pas' is in spec of NegP, 'ne' is Head of NegP)

Fr.   ne-aime pas   Neg'   'ne' must raise with verb to T.

                  ne-   VP   => nonfinite verb 'aimer' must raise to T to get tense.

                         V   DP

                     Aimer Maria

     *Eng.   likes   **not**

            does   not   like   Mary

Current English (CE) Head to Head movement: T-to-C (Aux inversion)

(15)   CP   b. Interrogative   CP

         C   TP   a. Declarative   spec   C'

       [T̵, EF]   [Q]

   b.   does Spec   T'   when C   TP

          John   [T̵, EF]

               T   VP   does spec   T'

           b. does   John

              Spec   V'   T   VP

                               ~~does~~

                V   N   spec   V'

                go   a. home

                    b. when ?   V   N

                                      go   when

But again, notice how in CE there is no V-to-C movement due to HMC (in addition to CE verbs not being 'strong')

*Care I not for her?

(16)   *CP   (contrasting (4) above where such movement is allowed in EME)

        C   TP

    Care Prn   T'

            I T   VP

               Adv   V'

               not V   PP

                    ~~care~~ P   Prn

                          for   her?
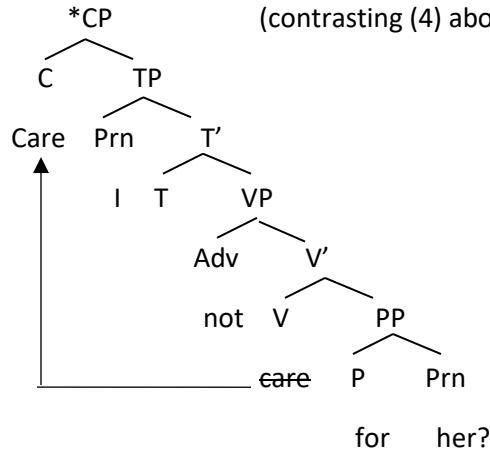
⇨ **Ex. 6 Null Constituents**

Null Subjects

The Extended Projection Principle (EPP) basically is a reflex of the simple fact that all declarative sentences must have a subject (in forming a TP max-projection). However, in some languages, the stipulated subject can be phonologically Null (not pronounced)—e.g., Spanish, Italian.

    (1) Italian: Maria é arrivato?   The null subject is labeled as *pro.*
          Si, _____ é arrivato.
          Si, *pro*   é arrivato.
        (yes, she has arrived).

    (2) English allows Imperative null subjects;
       a. *pro* Push the car!
          (You push the car)

    (3) A third type of Null constituent is found in non-finite structures:
       a. I would like [you to stay the night]
       b. I would like [Pro to stay the night]
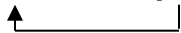
       We say that PRO is **controlled** by the pronoun I.

Let's consider the syntactic tree for the bracket clause in (3) above:

(4)

```
              TP
      You/PRO     T'
              T      VP
              |    V    DP
             to   |   D    N
                stay  |    |
                    the  night
```

Strong evidence of Pro is found amongst **reflexive anaphors**.

Coreferential within same Clause.

The first point is that in order for features (such as person/number/gender) to work in a coreferential manner, the two items (**pronoun/antecedent** *he* and the **reflexive/anaphor** *himself)* **must remain within the same clause**. 'Movement and Distance traveled' now becomes a defining aspect—'closeness is preferred over distance'.

(5)   a. John wants [Jim$_j$ to prove himself$_j$]

         b. *John$_j$ wants [Jim to prove himself$_j$]

 Hence, a PRO must be inserted within the bracketed 'prove-clause' in order to maintain feature coreferential control.

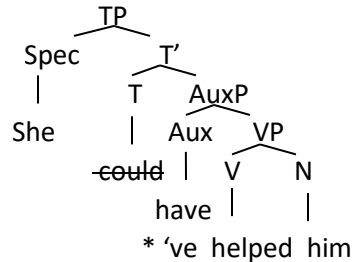(6)   John wants [TP PRO$_j$ to prove himself$_j$]

```
        ┌──────────┐
        ↑          │
```
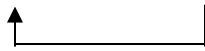
         Pro is controlled by John.

⇨ **Null Auxiliaries/ have-clitics**

(7)   a. He has helped her.

         b. He's helped her.

         c. *He could have helped her and *she've* helped him.

         d. He could have helped her and *she ~~could~~ 've* helped him.

         e. He could*'ve* helped her. ('ve pronounce like 'of')

Why can't *She've* (d) behave as a clitic? If the modal *could* is null (c̶o̶u̶l̶d̶), the null 'empty category' still fills the constituent slot thus blocking clitic formation—since clitics have to be **adjacent** to the host—both in overt phonology as well as in covert syntax—then we have a syntactic account for the blocking of clitics in such formations. (Note that the modal 'could' occupies a Tense and not the auxiliary verb 'have').

(8)

```
              TP
         Spec      T'
          |      T    AuxP
         She     |   Aux   VP
               c̶o̶u̶l̶d̶  |   V    N
                    have  |    |
                     * 've helped him
```

Since the Aux 'have' is not adjacent to the pronoun, clitic formation is blocked.

This same kind of covert empty category/blocking was what was behind our 'wanna contraction' example:

(9) **Empty Categories / wanna contraction**

a)   Who do you __   want to help ___? => (Who do you *wanna* help?)

             You (do) **want to** help who? => (base order)

b)   Who do you __   want__   to help you? => *(Who do you *wanna* help you?)
             You (do) **want who to** help you?   (want & to are separated by an empty category)

(10)

```
            TP
       Spec      T'
        |      T      VP
       You    T         VP
              |
             do   V        TP
                 want   spec    T'
                  a.    ___    T     VP
                  b.    who    |    V    N
                              to  help    |
                               a.  Who
                               b.  You
```
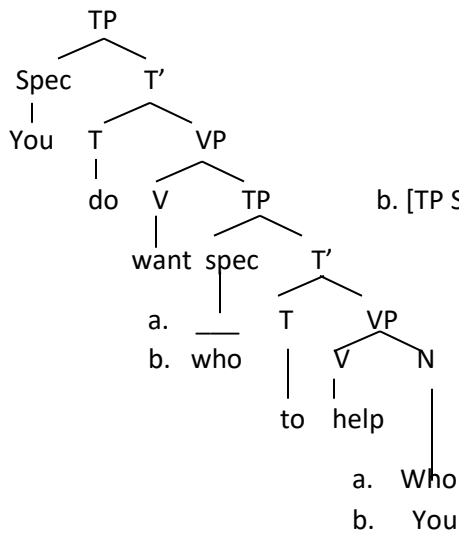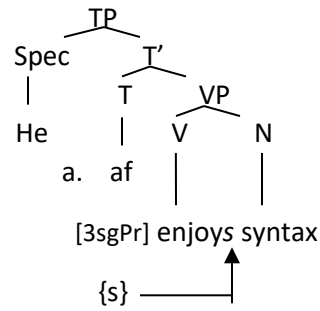
a. Who do you want [T' [T *to* [VP help you]]]

b. [TP She [T *can* [VP help you]]]

Null T   (Tense affix (af) inserted under null T)

(11)
```
                TP
          Spec      T'                    => Null T hosts af/features
           |      T    VP
          He      |   V    N              => 'affix hopping'
              a.  af  |    |
                 [3sgPr] enjoys syntax
                 {s} ————┘
```

        b.   doe*s* enjoy syntax   => T is filled with verb/inflection

Null T in Subjunctive Clauses

(12)   a. He suggests [that I have a physical exam]
          b. He suggests [that I ~~should~~ have a physical exam]
          c I've an exam (=> clitic 've')
          d. *He suggests [that I've an exam] => clitic blocked due to null T ~~should~~

A theoretical note:

I'd like to advance the notion that TP is affixal in nature (it is not a true phrase/phase)…but due to EPP and the fact that affixal properties must be housed under T, a TP is projected. Hence, the vP is the first real phrase/phase projection…on top of which an extended spec is adjoined onto a T-bar (intermediate projection). Recall, that T-bar is not a max-projection—hence, the EPP condition.

Null T in Infinitive/Small Clauses
  a.   I have never seen   [Tom __ speak to anyone].
  b.   I have never seen *[Tom to speak to anyone].
  c.   I have never seen *[Tom   speaks to anyone].

(13)
```
                TP
          Spec     T'
           |     T    VP
          Tom    |   V    PP
                to   |    △
                    speak  to anyone
```
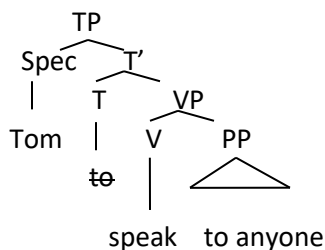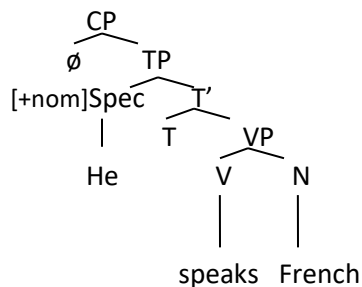
106

Null complementizer and Case marking

It can be argued that Case (nominative, accusative, genitive) is a result of a case-assigner feature which is associated with C (of CP). (An alternative approach, one which I pursue, is that the light verb vP assigns nominative case). For instance, complementizers such as '*that/for*' assign specific case:

(i)   I think [that he speaks French]
(ii)   *I think [that *him* speaks French]   =>   'that' assigns [+Nom] case.
(iii)   I want [for *him* to speak French]
(iv)   *I want [for *he* speaks French]   =>   'for' assigns [-Nom] case.

If case is assigned by C, then all declarative sentences must be CP projections by default due to case.

(14)
```
              CP
            /    \
          ø       TP
    [+nom]Spec    T'
            |    /  \
           He   T    VP
                |   /  \
                   V    N
                   |    |
                speaks  French
```

Now note both null C (for case)as well as a null subject (PRO) in spec of TP:

(v)   I will arrange   [PRO to see a specialist]
(vi)   (I will arrange [for me to see a specialist])

(15)
```
                 CP
               /    \
             ø       TP
                   /    \
    v. [-nom]Spec   T'
             |     /  \
            PRO   T    VP
             |    |   /  \
             |   to  V    DP
             |    |  |    /\
             |    | see  a specialist
    vi. for  me  to see  a specialist
```

107
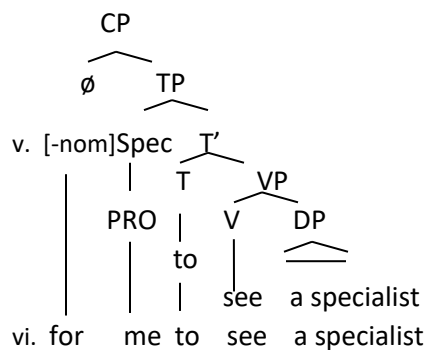
Defective Clauses (TP-embedded declaratives)

One account for [-Nom/-Fin] small clauses is that they are in fact defective TP clauses (and not full CP clauses). Consider the analysis:
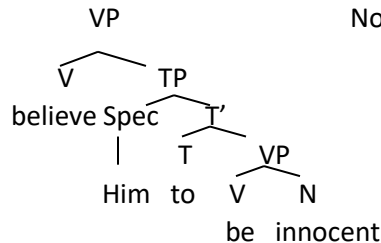
> **CP assigns case:**
> (i)   A finite complementizer such as 'that' assigns [+Nom] case,
> (ii)   A nonfinite complementizer such as 'for' assigns [-Nom] case,
> (iii)   Or a defective TP by default assigns [-nom] case.

We can assume that clauses which lack a CP (such as small clauses) are exceptional in that case gets assigned via default (otherwise known as **Exceptional Case Marking (ECM)**). Verbs like *believe, intend* seem to be ECM verbs. (What is exceptional is that the verb is in a different clause from the subject which it assigns accusative case to).

> **Defective clauses**
> a.   They <u>believe [him</u> to be innocent]   = >(ECM)

(16)   VP   No CP case assigner.

```
            V       TP
        believe Spec      T'
               |      T      VP
            Him   to    V      N
                       be   innocent
```

(An alternative approach is to have vP assign Nom case, TP assign Tense/AGReement and have all otherwise defective clauses assign [-Nom] by default. In this sense, TP is only an adjunct phrase and is affixal in nature. See attached paper 'Dual Probe-Goal Relation').

> **Null Determiners**
> a.   *[DP [D ø] [N John]]* admires Mary
> b.   *[DP [D The] [N John]] that I know is very bright*

Functional Features must be specified under a Functional Head, hence a null D must be projected within a DP:

(17)   DP
         ╱‾‾‾╲
        D   N
        ∅   John
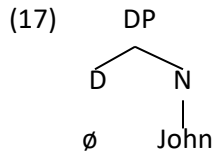
      [+Def]

     [+Nom]

      [3p]

      [-Pl]

    [+masc]

_____

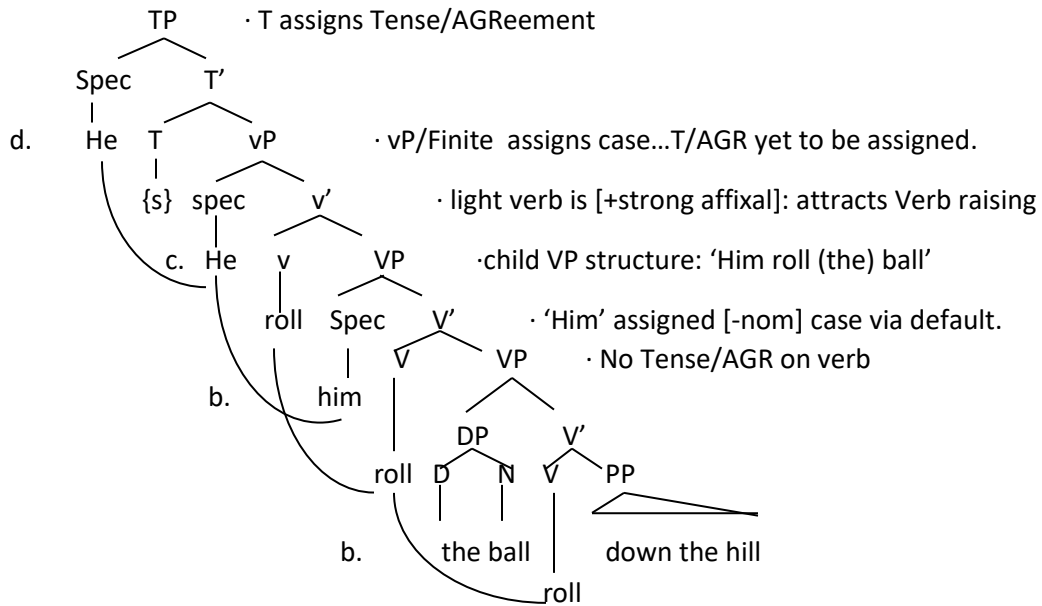⇨ **Ex. 7 A-Movement: VPISH (verb phrase internal subject hypothesis).**
   Light verb vP-shell
   Probe-Goal relation and Feature Valuation

   TP>vP>VP: Instances of Head Movement, wh-movement
   e.g., 'He rolls the ball down the hill'

(1)   TP   · T assigns Tense/AGReement
           Spec   T'
    d.   He   T   vP   · vP/Finite assigns case…T/AGR yet to be assigned.
                {s} spec   v'   · light verb is [+strong affixal]: attracts Verb raising
                    c. He   v   VP   ·child VP structure: 'Him roll (the) ball'
                             roll   Spec   V'   · 'Him' assigned [-nom] case via default.
                                           V   VP   · No Tense/AGR on verb
                        b.   him
                                         DP   V'
                                    roll D   N   V   PP
                        b.   the ball   down the hill
                                                  roll

109

(2) **vP Case / VP—Thematic-Lexical_Structure**

        d.   The ball roll down the hill.   ('roll' assigns **Theme argument** to 'the ball').

        e.   Him roll the ball ~~roll~~ down the hill. ('roll' raises to v to assign **Agent argument** to 'Him')

        f.   He roll the ball ~~roll~~ down the hill. ('He' received [+Nom] case in vP).

(3)   **TP—Functional Structure: Tense/Agreement features.**

        b.   He rolls the ball ~~roll~~ down the hill. (EPP forces 'He' to raise. Tense/AGR project)

        The above structure shows both **Head-to-Head movement** (of the verb: V>V>v) and **Spec-to-Spec movement** (of the subject: VP>VP>vP>TP)

Nominal Movement: Movement violations:

   Head Movement: Movement of a Head from a lower phrase into a Head of a higher phrase.

(4)   **The Head Movement Constraint (HMC):** A Principle of UG which specifies that movement between one head and another is only possible between the head of a given structure and the Head of its complement. Head movement target must be closest possible position: Head movement cannot pass over the closest c-commanding Head (cf. example c. below).

        a.   John will read the book

        b.   Will John _t_ read the book

        c.   *Read John will _t_ thebook

    (b) satisfies the HMC , (C) breaks HMC since the verb *read* crosses the nearest possible target occupied by *will.*

(5)   **Anaphoric Expression**

        a.   John criticized himself   (*John* c-commands *himself*)

        b.   *Himself criticized John (*himself* not c-commanded by John)

        c.   Himself, John criticized___.

In (5b), **the anaphoric/reflexive pronoun *himself* requires an antecedent that c-commands it**. Here, c-command is broken. But anaphoric connection is interpreted before movement (5c):

(6)   **Assignment Condition** (also see probe-goal relation).

      (i)   A. **Nominative case** if c-commanded by Intransitive finite complementizer (that, ~~that~~, if or null complementizer).
               B. (Alternative) Nominative if assigned by a Spec-Head relation of light verb vP.
      (ii)   **Accusative case** if c-commanded by Transitive head (transitive verbs like *meet, hit,* and transitive prepositions like *with or* transitive complementizer like *for).*
      (iii)   **Null case** if c-commanded by null Intransitive nonfinite complementizer ø (as with PRO to insure no phonological spell-out).

    a.   John/*he* is certain to win the race   => [+Nom] case via (i).
    b.   It is certain that John/he will win the race.
    c.   I'd like for John/*him* to win the race.   => accusative case via (ii)
    d.   I want [John/him to win]
    e.   I will arrange [ø [PRO to see the coach]]   => accusative case via (iii)

Note how (f) below doesn't enter into a case assigning relationship: *John* is not in a position to receive case. In addition, the tensed verb lacks the probe-goal relation to assign case.

    f.   * [__reads [**John** to prepare for exams]]

But if we assume (as we will below) that the light verb vP assigns case, and that an affixal Tense node could sit on top of a vP (as we will assume), then, in a sense, the only thing that makes the structure in (f) illicit is the condition of the EPP (that all clauses must have a subject) and the fact that English main verbs don't rise above vP to T (which affects word order here).

    g.   * [John reads [~~John~~ to prepare for exams]]

(7)   TP   · EPP shows subject raising into spec of TP.

          Spec   T'   · T affix for tense inflection.

       [EPP] T   vP   · light verb vP

      John {s} Spec   v'   · agentive null feature {ø}

            John v   VP   · Subject Verb Internal Hypothesis (VISH)

             read-ø Spec   V'   · V to v verb raising (agentive verb-ø)

                John V   T'   · Here we speculate on simply an affixal T-bar

                   read T   VP

                       to V   PP

                      prepare P   DP

                           for D   N
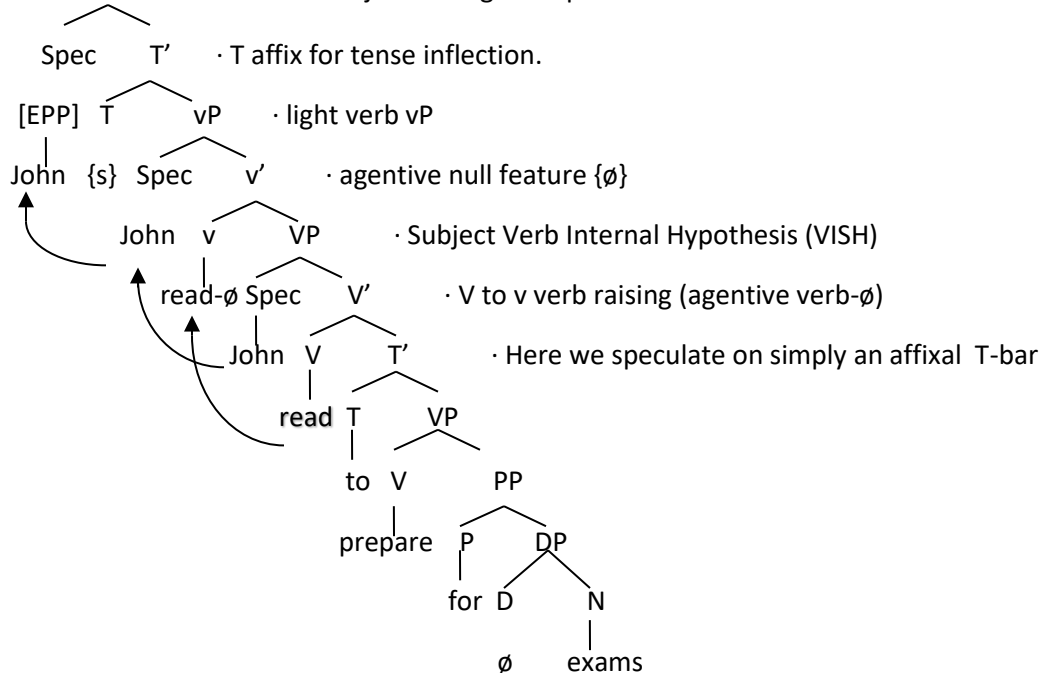
                              ø   exams

Chomsky (1995) assumes that verbs raise from V to (light verb) v in order to acquire a null agentive {ø} **light verb agentive** feature. So the structure above would be paraphrased as *'John performs the action of reading in order to prepare for exams'* where 'performs' serves as a kind of agentive *light verb.*

⇨ **Ex. 8 Case/Agreement: Probe-Goal relations (Miyagawa)**

(1) **T- probe / case: Nominative for T-probe [+Fin], accusative for v-Probe [-Fin].**

Probe T : [+Fin] (raising verb)

  a. [TP [probe] _____ are likely [vP ___ to be awarded *several prizes*]]
  b. [TP [probe] *There* are likely [vP ___ to be awarded *several prizes*]]
  c. [TP *Several prizes* are likely [vP ___ to be awarded]].

Probe v: [-Fin] (ECM transitive verb)

  d. [TP We expect [vP [probe] _____   to be awarded *several prizes*]].
  e. [TP We expect [vP [probe] *several prizes* to be awarded]].

The example of (1b) above now largely discredits the once held view that agreement involves a Spec-Head relation given that, here, while the Spec-subject 'there' has no inherent AGR features of its own (e.g., there is/are…) the verb 'are' must agree with the nominal 'several prizes' lower down in the tree. Such AGR within passive structures are problematic for a local top-down 'Spec-Head Agreement'-based relation. Hence, the less constrained **c-command** relation whereby Agreement enters a **probe-goal** relation can account for such passive constructs. Consider the sentence below:

  (f) There *are* awarded *several prizes*.

In (f), the verb 'Be'/(are) serves as a **probe** which searches for its nearest **c-commanded** nominal **goal** to agree with. The nearest and only nominal goal c-commanded by the probe 'Be' is the nominal QP 'several prizes (lower down in the syntactic tree).

(1f)   a.   CP
                    ╱   ╲
                  ø   TP
                        Spec   T'
                              T   VP
                           [probe]
                         There │   V   QP
                              are │   [goal]
                                 awarded   │
                                        several prizes.
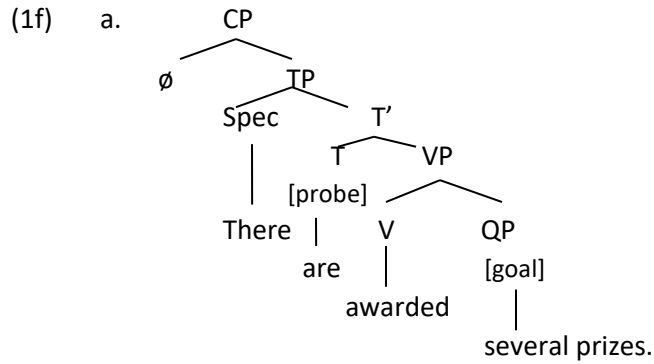
Therefore the probe 'be' agrees in **person** and **number** with the goal 'several prizes' to the extent that 'several prizes' becomes the active subject of TP—e.g., 'Several prizes are awarded' as found in (b) below:

(2)   b.   CP   c.   T'   = 'They were arrested' (b, ii)
               ╱   ╲   ╱   ╲
             ø   TP   T   VP
       [+nom]   Spec   T'   Be
                 [+Nom]   **[past]**   V   Prn
                         T   VP   [*u-Pers*] arrested   THEY
                      [probe]   [*u-Num*]   **[3-Pers]**
                              V   QP   **[Pl-Num]**
(i)   Several prizes   [*u-case*]
                           are │   [goal]
                          awarded   │
                                 ~~several prizes.~~

(ii)   They   were arrested   ~~they~~
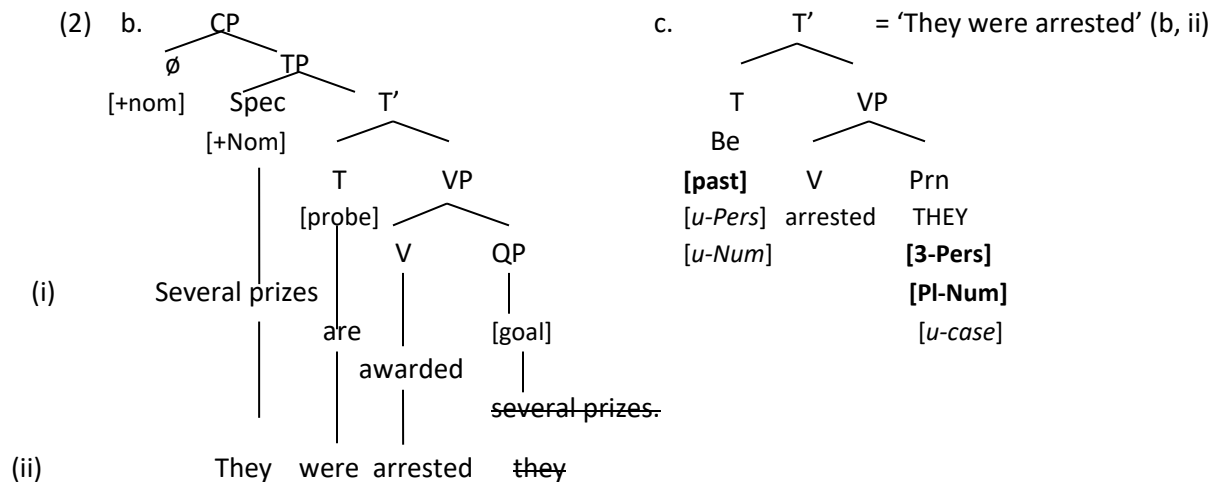
For (b), the raised (goal) 'several prizes' now the subject spec of TP (to satisfy EPP) is assigned **structural** nominative case via the (structural) agreement relation with [T are].

The passive structure in (c) follows Chomsky's recent terminology calling for **unvalued features** (u-Pers(son), u-Num(ber), u-Case). Firstly, consider the feature spell-out of **Be.** Before 'Be' enters into its probe-goal relation (getting its Pers/Num values based upon the goal **they**), it is said that **Be** enters the derivation with its Pers/Num features still unvalued. Secondly, consider an item, say, a pronoun, that may be pulled from the lexicon with person/number features already valued but with its case features unvalued. This could happen due to the fact that it's the Agreement mechanism which determines the value of the case feature , and prior to AGR it is said that the case feature is yet to be valued. So case is unvalued. In the sense, the derivation doesn't determine *a value of Case* but rather agreement does.

The above examples deem the probe-goal relation (and not the Spec-Head relation) as the only credible account for the dual nature of passive/active agreement.

Case and Agreement properties within the specific location (probe T, probe v) are determined by the properties of the matrix probe. Case gets deleted by the probe when it enters into the AGR relation. Thus agreement is what triggers deletion of the [-interpretable] case feature.

In (1f, a) the expletive 'There' in Spec TP contains a probe the searches for the nearest argument, noting the plural/number agreement relation between (true subject lower down in the tree) '*several prizes*' and the verb '*are*'.

Note that TP has no intrinsic property of Case/Agr, rather their feature-spell-outs are the sole property of the probe of the head—and not of the head of T itself since T is a **defective head** and is unable to determine Case/AGReement features on its own, but has only an EPP feature. But between the two, Case/Agr, AGR pertains to a real mechanism with an index related to a real argument. Case on the oter hand has no such concrete property. Case is a problem— regarding transfer properties, it has no semantic (LF) properties of its own, while containing only phonological (PF) properties. This mismatch is a non-trivial problem. Case seems to be the one [-interpretable] feature which defies any notion of a optimally-designed language system.

It can be argued that the T-head/probe is the result of 'merge' operation rather than 'move' (**merge over move**) since TP/T-head in not the kind of a phase which has **semantic (LF)/phonological(PF) transfer properties**. (Only CP and transitive vP are phases). As claimed, Case/AGR bears no relation to TP (which is only responsible for affixal/inflection tense. In the examples, above, it seems to be the case that it is the independent (spec-head) probe relation of [+/-Fin] within T –head which establishes case.

Case assignment via C-command (restated from Ex. 7 (§6) above):

(3)   **Assignment Condition.**

      (i)   A. **Nominative case** if c-commanded by Intransitive finite complementizer (that, ~~that~~, if or null complementizer).
              B. (Alternative) Nominative if assigned by a Spec-Head relation of light verb vP.

      (ii)   **Accusative case** if c-commanded by Transitive head (transitive verbs like *meet, hit,* and transitive prepositions like *with or* transitive complementizer like *for).*

      (iii)   **Null case** if c-commanded by null Intransitive nonfinite complementizer ø (as with PRO to insure no phonological spell-out).

Nominative Case (cf. 3i)

(4)　　　　　　　　CP　　　　=> Intransitive ø Null Complementizer Head /Nominative Case
　　　　　　　ø　　　TP
　　　　　[+nom]Spec　　T'
　　　　　[probe]　│　　T　　VP
　　　　　　　　│　[goal] {s}　V　　N
　　　　　　　That　│　　│　　│
　　　　　　　　│
　　　　　　　　if　　he　　speaks　French
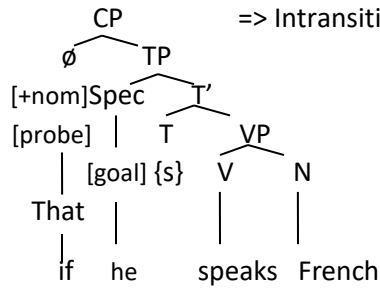
Alternative: Nom case via Spec-Head of [+Fin] Head/ functional vP.
(by defining *light verb* vP as [+finite], and VP as [-finite] v ia Case)

　　　　　He speaks French.

(5)　　　　　　　T'　　　(T-bar) (since TP is not a phase—a consequence of merge, not move)
　　　　　Adjoin　　T"　　　　　(merge T' to vP).

　　　　　　EPP　T　　　vP [+Fin]

　　　　　He　{s}　Spec　　v'
　　　　　　　　　[+nom]　v　　VP [-Fin]
　　　　　　　　　　│　　　spec　　V'
　　　　　　　　He　　[-nom]　V　　N
　　　　　　　　　　　│　　│　　│
　　　　　　　　　　　him　speak　French
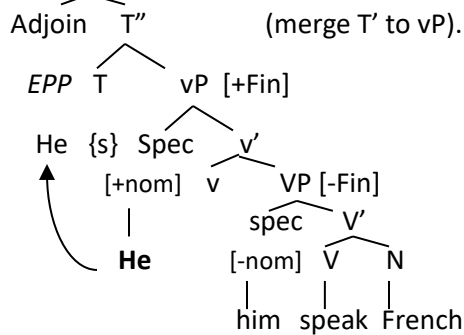
(Note: Adjunction (of T-bar) is generated for no other reason other than for EPP. Otherwise, the top is a 'merge-based' T-bar –adjunct: Case is assigned via vP and Tense inserts under a merged 'affixal T' (of T-bar). Using a 'merge over move' analysis, we can argue that phases vP and CP are 'move-based' due to their having a complete spec-head-comp structure. T-bar, not being a complete phrase/phase is merge-based affixal only).

(Alternative) Accusative case via Spec-Head of [–Fin] Head.

e.g., He wants [*him* to take the class].

(6)   vP   (not showing [+Fin] TP)
                    Spec   v'   [+Fin]
                  [+nom]   v   VP
                     | | spec   V'
              He want-ø [-Nom] V   T'   Accusative case via Spec-Head of [-Fin] head 'to'
                                          Adjunct   T" [-Fin]
                              him   [-nom]   T   VP
                                  want | V   DP
                                                        him   to   take the class
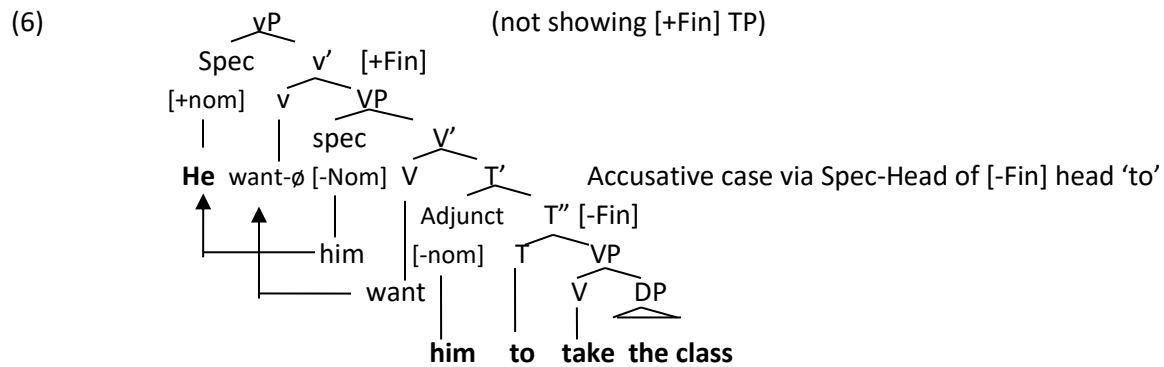
Accusative case via c-command by transitive head

(7)   TP
                      Spec   T'
                        | T   VP   => Transitive Head
                       He   {s}   V   N
                                      | |
                                   meets   **him**
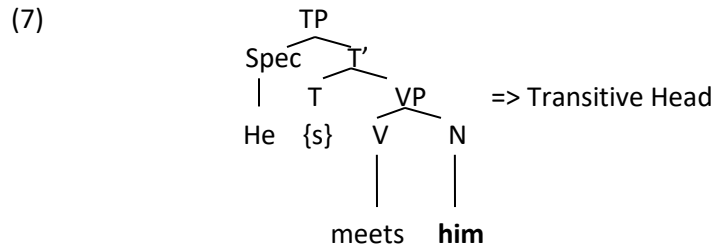
Accusative Case via c-commanded by transitive Complementizer 'for'

(8)   CP
                            TP
                 C   Spec   T'
              [-nom] | T   VP
                for   **Him** | V   N
                              to | |
                                   meet   **them**
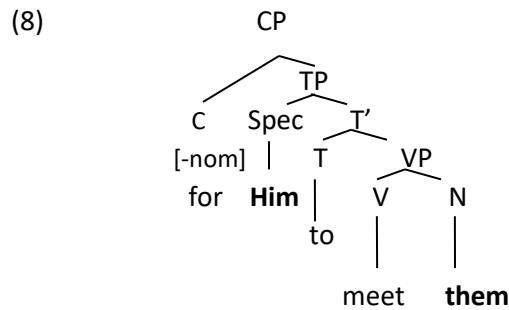
Exceptional Case marking (ECM)

ECM subjects are said to be 'exceptional'   in the sense that they are assigned accusative case via a verb not in the same clause (they receive accusative case via the verb in a higher clause).

Recall from (3ii) above that normal **Non-ECM Accusative case** is c-commanded by a transitive head of the same matrix clause (transitive verbs like *meet, hit,* etc.—e.g., We'll *meet* **them**, John *hit* **him,** etc.)

(9)

```
                    CP
              ⌐‾‾‾‾‾‾|
             ∅        TP
                  ⌐‾‾‾‾‾|
        [+nom] Spec      T'
                      ⌐‾‾‾|
                     T      vP
                         ⌐‾‾‾‾|
          We    'll   spec    v'              => c-command by transitive head /accusative case
                         ⌐‾‾‾‾‾|
                        we  v     VP
                            |   ⌐‾‾‾|
                           meet Spec  V'
                                    ⌐‾‾|
                                   V    N
                                   |    |
                                  meet  him
```
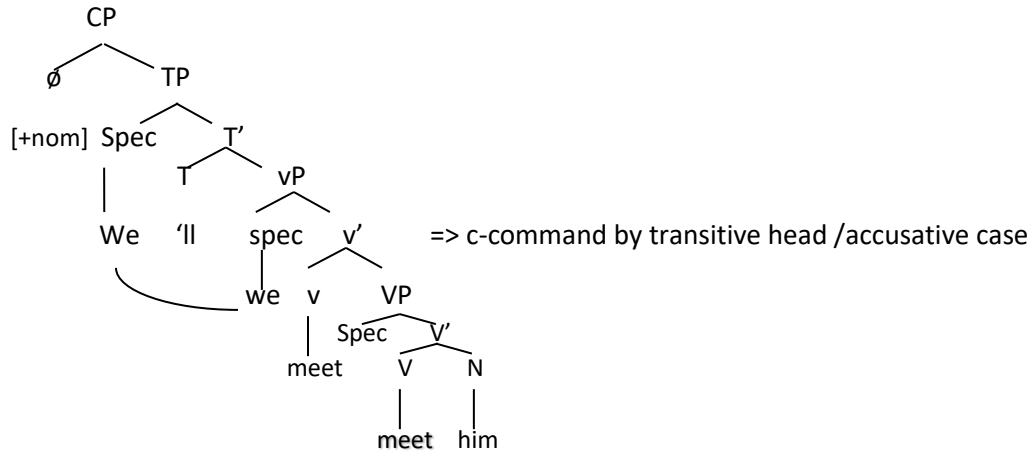
But notice how ECM subjects are not in the same clause:

(10) ECM verbs are verbs like *believe/intend…*

      a.   *[They <u>believe [TP him</u> to be innocent]].*
      b.   *[They believe [CP for [him to be innocent]]].*

It's a hallmark of ECM verbs that their case marking subjects seem to appear within TP clauses—sometimes considered as **Defective TP clauses** since they are not contained within a full CP (similar to our T-bar alternative account above). Note how the complementizer '**for** him <u>to be</u>' (known as 'for-infinitives') cannot be projected showing a CP and that only a truncated TP serves).

Working bottom-up:

    a.   Bottom …TP   b. Top CP…

They believe [TP Him to [VP be innocent]]   [CP [TP They [VP believe]]] him to be innocent
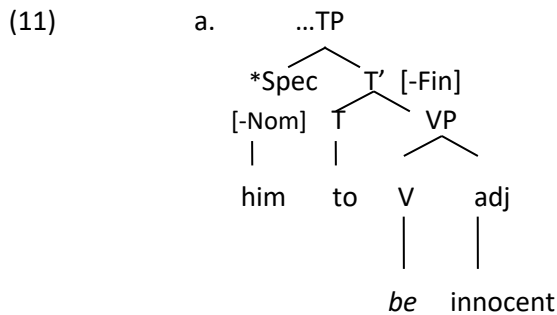
(11)   a.   …TP   b.   CP

```
   a.       ...TP                          b.          CP
        ⌐‾‾‾‾‾|                                    ⌐‾‾‾‾|
      *Spec     T' [-Fin]                         C      TP
     [-Nom]   ⌐‾‾|                             ⌐‾‾|  ⌐‾‾‾|
             T    VP                          ∅  Spec   T'
             |    |  ⌐‾‾|                           |  ⌐‾‾|
            him   to V   adj                      They T   VP
                     |    |                            V   ⌐‾‾|
                    be  innocent                       V    TP...
                                                    *[-Nom]
                                                       |
                                                     believe
```

117
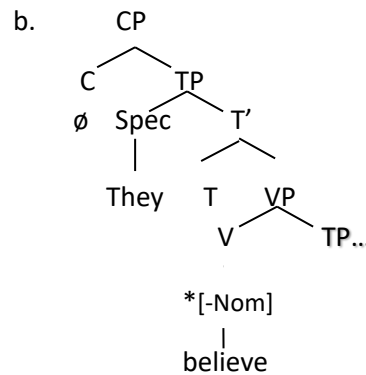
*ECM (Exceptional Case Marking) transitive verb *believe* assigns accusative [-Nom] case to its subject *him* lower down in the tree.

(Note on alternative account on case: ECM seems to make sense if we posit that all spec positions above vP are in fact adjoin positions and case gets assigned via a [+/-Fin] spec-head relation).

=> Ex. 9 Case marking and the Null Constituent: 'Case Revisited'

[1] **Case: The Local Probe-Goal Domain: where above (Functional Head) probe searches for the closest c-commanded noun expression (just below it). That the probe hands-over its AGR/Case features to the head immediately below it (AR: 400)—so:**

> **(i) Head of vP hands-over its features to V of VP (Accusative Case)**

> **(ii) Head of T hands-over its features to v of vP (Nominative Case)**

> **(iii) With an added stipulation on Subjects that ϕ-AGR features PERSon and NUMber get handed-over by C of CP.**

In Sum: we follow Radford and assume that Case is via Probe-Goal relation with highest functional Head just above, and that AGReement on subject is handed-down from C (AR: 398).

a.   **FH** (**functional Head**, above)   **Head>Spec relation**

```
        [H]        XP (TP, vP)
          ⌐→ Spec
```
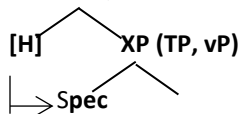
[2].'**Object-to-Subject raising' (Passives)**

'They were arrested'   [CP [C Ø] [TP They [T were] [VP [V arrested] they]]]

a.   T'   (AR: 286, Ex (9))   b.   TP

```
     BE      VP                                  Spec        T'

 [past]   V      Prn                            They    T        VP

 [3-pers] arrested  THEY                               were   V      Prn

 [Pl-Num]          [3-pers]                           [EPP]  arrested  T̶h̶e̶y̶

                   [Pl-Num]

                   [Nom-case]
```

118
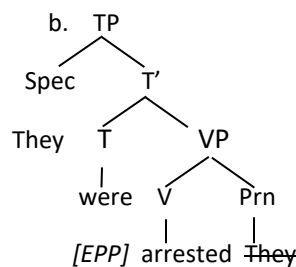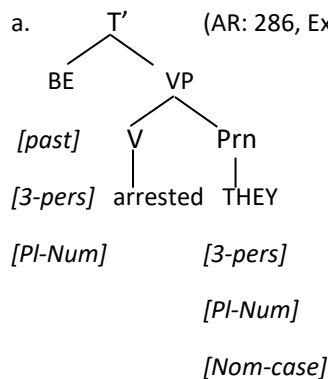
We note above that example (c) would have a CP layer so that T of TP would receive its AGR features [3-Person], and [Plural-number] from C (cf. 1, iii).

At this point of the derivation (reduced by not showing split-VP or light verb projections) , the following has already happened (with CAPITALIZED Aux and Subject indicating pre-spell-out bare forms as pulled from out of the lexicon)

(i)   The auxiliary 'BE' gets spelled out as *'were'* with person/number ɸ-AGR features,
(ii) The subject 'THEY' gets case-marked by the probe in T.
(iii) The resulting spell-out projects ' __were arrested they'
(iv) However, the derivation is not yet terminated: The EPP feature of T (shown in (2b)) will need to trigger 'They' into Spec of TP. This movement is not due to Case or AGR, but rather exclusively owing to EPP.

[3] 'Subject-to-Object raising' (subject of lower infinitive TP becomes subject of higher VP)

a. The DA proved the witness conclusively to have lied (AR: 400)

b. [vP The DA proved [VP the witness conclusively ~~prove~~ [TP ~~the witness~~ to have lied]]]

The fact that the adverb 'conclusively' projects after 'the witness' suggests that what originated as the subject of the TP has in fact raised up to become the object of the VP(the object of the transitive light verb *prove-Ø.* Consider the tree below:

[4]
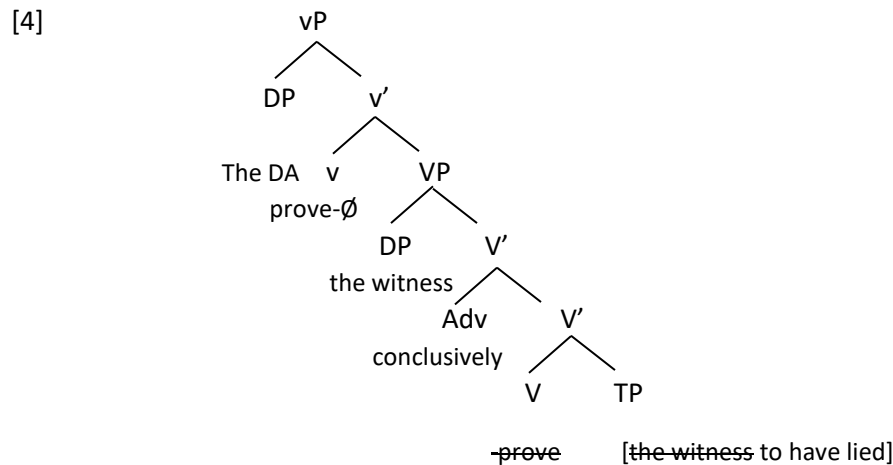
We could extend the same analysis for all defective clauses: namely, that all subjects of defective TP-clauses raise up and become objects of the above VP. In this manner, we could extend the EPP feature to V as well as T. Consider how this might play out below and how we might define ECM (Exceptional Case marking) of defective clauses as not so exceptional anymore (in the sense that the subject of ECM (a lower clause, is actually now raised inside of the same clause which is doing the case marking).

 (We recall that ECM was defined as 'exceptional' since the matrix verb doing the case assigning was in a different (higher) clause than the element becoming case-marked).

[5]   a. The DA proved him to have cheated on his wife.

        b. [vP The DA proved him [TP him to have cheated on his wife]]

[6]

```
                        vP
                       /  \
                    DP      v'
                   /       /  \
              The DA    v      VP
                     prove-Ø   /  \
                            DP      V'
                           Him     /  \
                                  V    TP
```

        -prove   [HIM to have cheated on his wife]
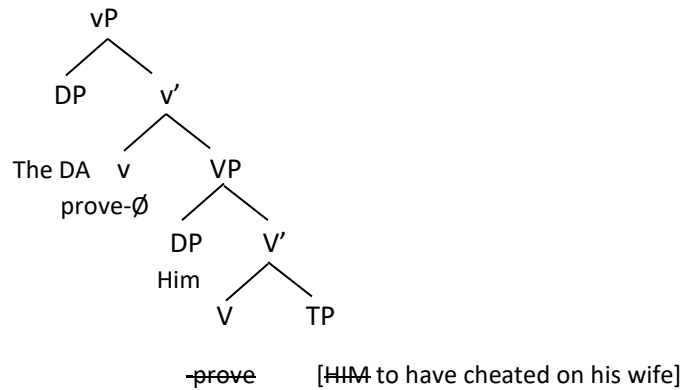
(i)   'HIM' (unvalued case subject of infinitive TP) raises to spec of VP (object) to be close to probe the
        light verb (probe) just above so as to receive Accusative case ('him').
(ii)   Hence ECM/defective clauses force a split vP>VP projection on top of the infinitive TP.
(iii)   What we had already assumes of pure object raising (out of lower VP) for case, the same holds
        for subjects of lower TPs.
(iv)   Split projections are required throughout the derivation.

Summary of Case Marking (AR, p. 124, 304)

[7] **Given (VISH) subject internal verb hypothesis, subjects originate internally within VP, we can
assume that they are Case-marked by the closest functional head above them (AR: 286, ex. (9))**

[8] **Case marking is done by the above closest functional Head:**

        a. Nominative Case for spec of VP: (Probe is T).
            Nom case if c-commanded by Intransitive finite complementizer ((that, if), or null finite
            main clause {T = null}. (In English, main finite verbs don't occupy T since they don't raise
            out of the light verb [v]. Hence, T would have no phonological manifestation, it would be
            null. Only English Auxiliary verbs (*Do, Be, Have*) and Modals (*can, could, will, would*, etc)
            would raise out of [v] into[T]).

                (i)   [CP [TP He T [Ø] [vP He….]]]   [*He* is nice]
                (ii)   [CP [C that/if] [TP He [vP He…]]]   [I think [*that*/if [*he is* nice]]]

b. Accusative Case for Spec of VP (VP2 of a split VP): Probe is *light verb v*.
Acc Case if c-commanded by a transitive Head (*break*), or a transitive preposition (*with*), or a transitive complementizer (*for*).

(i)	[vP [v break-Ø] him]	[She will [*break him*]]
(ii)	[PP with [DP him]]	[She will go [*with him*]]
(iii)	[CP [C for] [TP him [vP ~~him~~…]]]	I want [*for [him to meet the*m]]

Case Assigning Condition (Radford: 124)

A nominal expression (Noun, Pronoun) is assigned case by the closest case-assigner which c-commands it (**Earliness Principle**) and is assigned:

Let's pause and see how in (ii) both 'If/that' and the null complementizer Ø introduce a finite main clause:

[9]
```
                 CP          (Radford: 112, showing a reduced tree)
                /  \
              C     vP
                   /  \
          If/that PRN   T'
                       /  \
                  he  T    VP    (showing VISH)
                     /  \
                   had  [HIM  resigned]
```
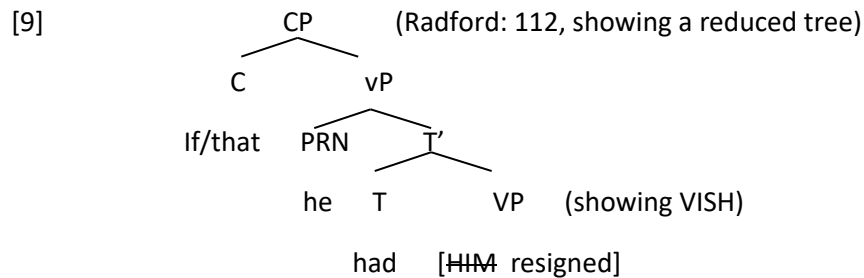
Null Case if c-commanded by a null intransitive nonfinite complementizer Ø.

[10]
```
                 CP          I will arrange [PRO to see a specialist]  (Radford: 118)
                /  \
              C     TP
                   /  \
            Ø    PRN   T'
                 /  \
           PRO  T    VP (VISH)
               /  \
             to   [PRO see a specialist]
```
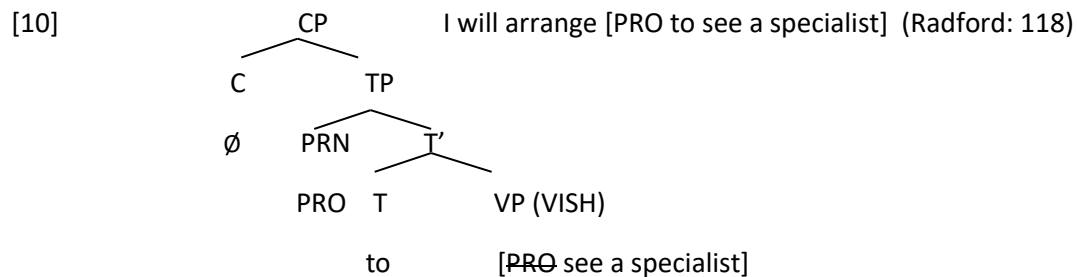
In both examples, we show HIM and PRO respectively as originating within VP (VISH) with subsequent raising in order for element to be close to the probe (feature hand-over).

So, as Miyagawa suggests, there is a kind of probe-goal relation whereby the closest c-commanded nominal becomes the goal and receives Acc case if probe is 'for', the goal receives Nom case if the probe is 'that/if', etc.

Subject Case: probe is intransitive Head　　　　**Object Case: probe is transitive Head**

[11]　　T'　　　　　　　　　　　　　　　　[12]　　　v'

[T] Probe　　vP　　　　　　　　　　　　**[v] Probe**　　VP

　[+Nom]　　**Goal** = Spec of vP　　　　　　[+Acc]　　**Goal** = Spec of VP

[Per/Num]　'*he*'　　　　　　　　　　　　[Per/Num]　'him'
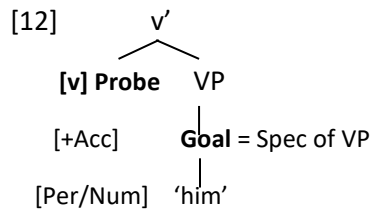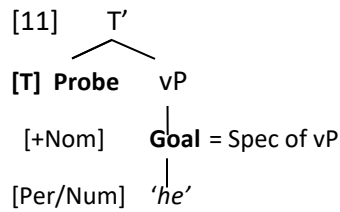
> • **All Case checking is done before the Subject raises into Spec of TP.**
>
> • **All Case is valued/deleted inside the first functional category above VP which is vP.**
>
> • **Noun expressions are case-marked by the <u>closest functional head above them</u>**: both T and light verb [v] are functional heads.

[13]　<u>Earliness Principle (AR, p. 282)</u>

Operations must apply as early as possible in a derivation.

Agreement and the Null Constituent

[14] **AGReement: Null [C]: All declaratives are CPs (AR : Chapter 7)**

[15]　　　a. [CP Ø [TP there [T were/*was] [VP awarded [QP {ϕ-F} several prizes]]]]

　　　　　b. [CP Ø [TP several prizes [T were/*was] [VP awarded [QP {ϕ-F} ~~several prizes~~]]]]

```
                    CP
                  /    \
                 C      TP
                [Ø]    /  \
                    Spec   T'
                    there  / \
                          T   VP
                        were  / \
                             V   QP
                        awarded  several prizes
```
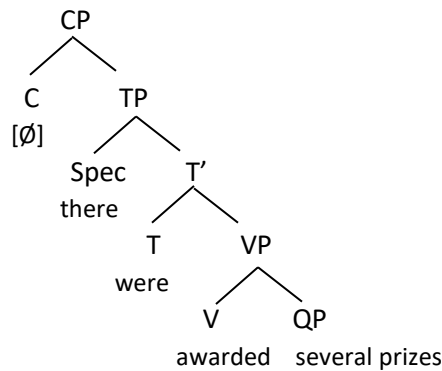
　　　QP/DP phi (pronounced 'fie') ϕ-features:　'*several prizes*': [**3P, +Plural**]

　　　Person and number become the AGR (phi) ϕ-features.

　　　'there' [+Pl, 3P] = ϕ-features which establish 'subject-verb AGR'.

[16] A CP Null constituency is required in order to create a probe-goal relation of AGReement down the tree (between 'several prizes' and 'there' [+Pl]).

　　One account for [-Nom/-Fin] small clauses is that they are in fact defective TP clauses (and not full CP clauses). Consider the analysis:

[17] **In Sum:**   **Above functional Head assigns Case to lower Spec below (TP, vP):**

  (iv)   A finite complementizer such as 'that' assigns [+Nom] case,

  (v)   A nonfinite complementizer such as 'for' assigns [-Nom] case,

  (vi)   Or a defective TP by default assigns [-nom] case (similar to what we find with inherent case [*give [to him*]] for datives).

[18]   **FH** (**functional Head**, above)   **Head>Spec relation**

  [C]   **XP (TP, vP)**

  ⤷ S**péc**
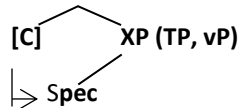
⇨ **Exceptional Case Marking (ECM)**).

We can assume that clauses which lack a CP (such as small clauses) are exceptional in that case gets presumably assigned via default (otherwise known as **Exceptional Case Marking (ECM), or via a probe-goal of the ECM verb**). Verbs like *believe, intend* seem to be ECM verbs. (What is exceptional about this is that the verb is in a different clause from the subject which it assigns accusative case to. But his 'above-to-below' <Head to Spec >configuration is exactly what we find regarding our more general Probe-Goal relations as discussed throughout our class notes.

⇨ **Defective Clauses (TP-embedded declaratives)**

  Defective clauses
  b.   They <u>believe [him</u> to be innocent]   = >(ECM)

[19]   VP   No CP case assigner.

         V   TP
      believe Spec   T'

            | T   VP
          Him   to   V   N

                      be   innocent
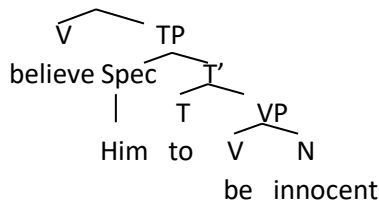
  Null Determiners
  Consider the sentences below: a Probe from above Functional Head

[20] a. [TP [Det A]   number of students   [T *is/are] [V dropping]]. **[-Def]**

  b. [TP [Det The] number of students   [T is/*are] [V dropping]]. **[+Def]**
  c. [TP [Det Ø]   Numbers of students   [T *is/are] [V dropping]]. **[-Def]**

The DP [**A number**] vs . the DP [**The number**] differs in [+/- Def] and affects the feature Number correlating to Subj-Verb agreement. An 'AGReement-based' Probe-Goal relation must be established from a Functional Head above the goal

But why doesn't the probe [Det] <u>A number</u> require the agreeing verb <u>are</u> to be close (as with normal Probe-Goal relations)?

e.g., *<u>A number are</u> of students <s>are</s> dropping.

Well, it is close, considering that the DP [DP the/a number of students] forms a single constituent and nothing can break constituency(as shown in (21) below:

[21]   [DP Which_] does she like [DP _ films]?

[CP Which films [C does] [TP she [T <s>does</s>] [VP <s>HER</s> [V like] <s>which films</s>]]]

  (i)   Aux inversion of 'does' (from T to C)   (**Head to Head movement**)
  (ii)   Wh-movement of 'which films' (Comp of VP tp Spec CP) (**A-bar movement**)

Functional Features must be specified under a Functional Head, hence a null D must be projected within a DP:

[22]   DP

     D   N

      ∅   John
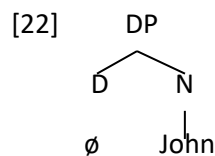
  [+Def]

  [+Nom]

   [3p]

   [-Pl]

  [+masc]

The Extended Projection Principle (EPP) basically is a reflex of the simple fact that all declarative sentences must have a subject (in forming a TP max-projection). However, in some languages, the stipulated subject can be phonologically Null (not pronounced)—e.g., Spanish, Italian.

[23]   Italian: Maria é arrivato?   The null subject is labeled as *pro.*

        Si, ____ é arrivato.
        Si, *pro*   é arrivato.
      (yes, she has arrived).

124

[24] English allows Imperative null subjects;

      b. *pro* Push the car!
           (You push the car)

[25] A third type of Null constituent is found in non-finite structures:

      c. I would like [you to stay the night]
      d. I would like [Pro to stay the night]

        We say that PRO is **controlled** by the pronoun I.

        Let's consider the syntactic tree for the bracket clause in (3) above:

[26]
```
              TP
         You/PRO    T'
                   T     VP
                   |   V     DP
                   to  |   D    N
                     stay  |    |
                          the  night
```
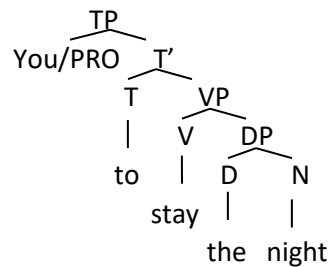
[27] Strong evidence of Pro is found amongst **'reflexive anaphors'.**

Coreferential within same Clause.

The first point is that in order for features (such as person/number/gender) to work in a coreferential manner, the two items (**pronoun/antecedent** *he* and the **reflexive/anaphor** *himself)* **must remain within the same clause**. 'Movement and Distance traveled' now becomes a defining aspect—'closeness is preferred over distance'.

[28]   a. John wants [Jim$_j$ to prove himself$_j$]

        b. *John$_j$ wants [Jim to prove himself$_j$]

 Hence, a PRO must be inserted within the bracketed 'prove-clause' in order to maintain feature coreferential control.

[29] John wants [TP PRO$_j$ to prove himself$_j$]
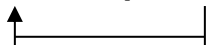
        Pro is controlled by John.

This same kind of covert empty category/blocking was what was behind our 'wanna contraction' example:

Empty Categories / wanna contraction

[30] a) Who do you __ want to help ___? => (Who do you *wanna* help?)

  You (do) **want to** help who? => (base order)

  c) Who do you __ want__ to help you? => *(Who do you *wanna* help you?)
   You (do) **want who to** help you?   (want & to are separated by an empty category)

[31] TP

    Spec   T′

    You   T   VP   a. Who do you want [T′ [T *to* [VP help you]]]

    do   V   TP   b. [TP She [T *can* [VP help you]]]

    want spec   T′

    a. ___   T   VP
    b. who | V   N

     to   help

     a. Who
     b. You

Null T   **(Tense affix (af) inserted under null T)**

[32]   TP

   Spec   T′   => Null T hosts af/features

    | T   VP

   He | V   N   => 'affix hopping'
   a. af | |

   [3sgPr] enjoy*s* syntax

   {s} ————┘

   b.   doe*s* enjoy syntax   => T is filled with verb/inflection
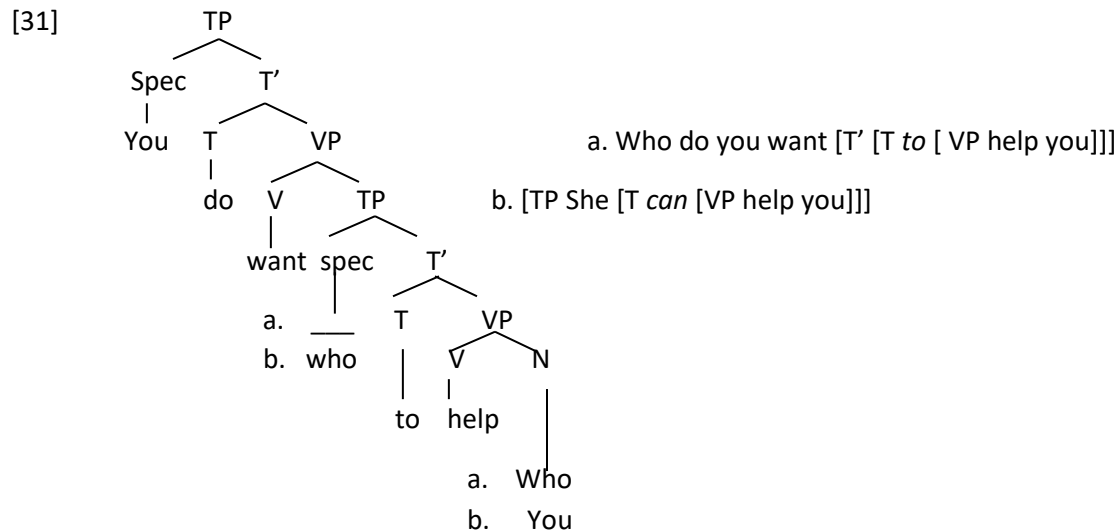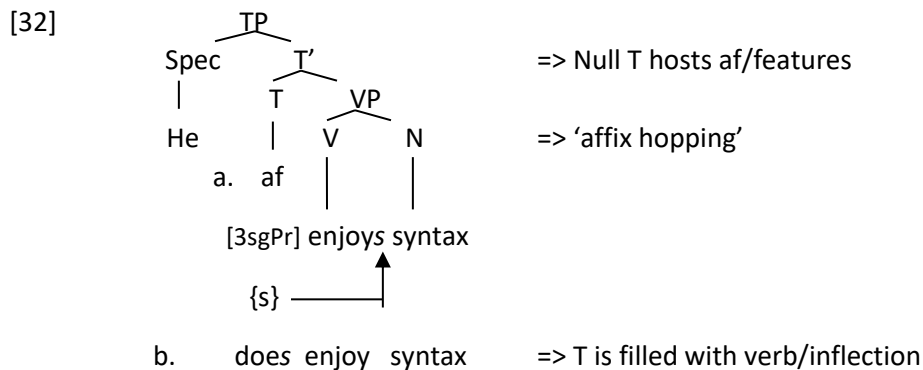
Null T in Subjunctive Clauses

[33] a. He suggests [that I have a physical exam]
   b. He suggests [that I ~~should~~ have a physical exam]
   c I've an exam (=> clitic 've')
   d. *He suggests [that I've an exam] => clitic blocked due to null T ~~should~~

A theoretical note: (See *References/Other Sources* for link to paper 'A Dual Probe-Goal relation Miyagawa'. Also see No. 10 *Snapshots on Syntax*).

I'd like to advance the notion that TP is affixal in nature (it is not a true phrase/phase)…but due to EPP and the fact that affixal properties must be housed under T, a TP is projected. Hence, the vP is the first real phrase/phase projection…on top of which an extended spec is adjoined onto a T-bar (intermediate projection). Recall, that T-bar is not a max-projection—hence, the EPP condition.

Null T in Infinitive/Small Clauses
a.   I have never seen   [Tom __ speak to anyone].
b.   I have never seen *[Tom to speak to anyone].
c.   I have never seen *[Tom   speaks to anyone].

[34]
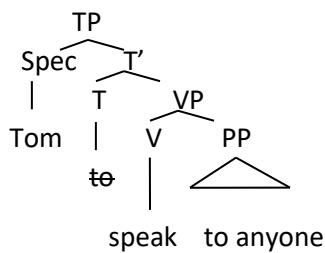
```
                TP
             /      \
          Spec      T'
           |       /  \
          T      T    VP
          |          /  \
         Tom  |    V     PP
             to    |     /\
                         /  \
                speak   to anyone
```

Null complementizer and Case marking

It can be argued that Case (nominative, accusative, genitive) is a result of a case-assigner feature which is associated with C (of CP). (An alternative approach, one which I pursue, is that the light verb vP assigns nominative case). For instance, complementizers such as '*that/for*' assign specific case:

[35] A CP to TP Case assigner

(vii)   I think [that [he speaks French]]
(viii)   *I think [that [*him* speaks French]] =>   'that' assigns [+Nom] case.
(ix)   I want [for [*him* to speak French]]
(x)   *I want [for [*he* speaks French]]   =>   'for' assigns [-Nom] case.

If case is assigned by C, then all declarative sentences must be CP projections by default due to case.

[36]

```
                CP
             /      \
            Ø        TP
                  /      \
        [+nom]Spec       T'
                |       /   \
               T      T      VP
               |            /   \
               He          V     N
                           |     |
                         speaks French
```
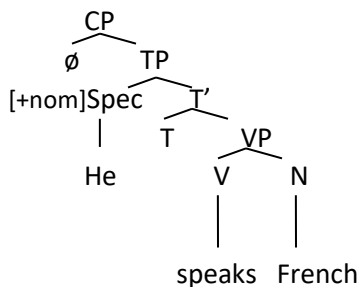
Now note both null C (for case)as well as a null subject (PRO) in spec of TP:

(xi)   I will arrange   [Ø [PRO to see a specialist]]

(xii)   (I will arrange [for [me to see a specialist]]

[37]
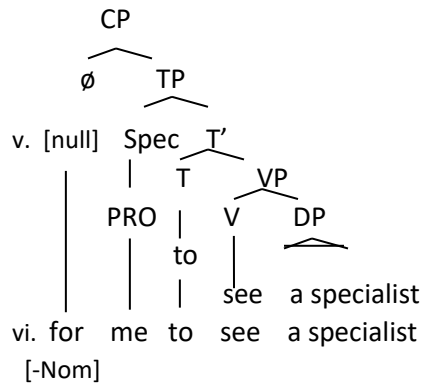
⇨ **Ex.10 Why Move? AGReement doesn't trigger MOVE**

Contrary to much theoretical work in the 1980s and early 1990s, **it is now thought that AGReement does not require movement** (displacement up the tree) of the element triggering agreement. In other words, the former tight constraint on **locality of agreement** (thought to be locally configured within a Spec-Head relation) has been relaxed and replaced by a (potentially) 'long-distance agreement relation' having to do with **c-command**. One of many empirical facts which led to this conclusion on agreement comes from the following facts in (i). We could equally extend a non Spec-Head locality condition on Case assignment as well by observed facts as presented in (ii):

(i)   Where it was clear that the agreeing DP had not moved anywhere close to a spec-head domain for agreement, but rather where the agreement held seemingly through c-command. (cf. [1], [2]).

(ii)   Where it was clear that Case assigned DP (GOAL) had not moved above the Head (PROBE) for Case as would be necessary for a spec-head relation. (cf. [3], [4]).

For example, regarding agreement, consider the existential construction:

AGReement

[1] There *do* seem (to appear to Mary) to be *two men* in the garden.

[2] *There *do*es seem (to appear to Mary) to be *two men* in the garden

(i)   [TP Spec [+Pl] Two men [T [+pl] do   [vP [v] seem]]]…. [+AGR]

(ii)   *[TP Spec [+Pl] Two men [T [-pl] does [vP [v] seem]]]…. [-AGR]

In (1) it's clear that existential <there> is plural in number (as it agrees with <do>) and that it correctly agrees with the plural true subject <two men> very much lower down in the tree. The problem, however, with this kind of 'long distance agreement' is that it breaks 'close locality' as would be required by a very local Spec-Head relation.

Case

[3] [TP They [vP [v **believe**] [TP *him* [T to] [VP be innocent]]]]   (AR, 128, ECM, Defective Clause)

[4] I need [CP [C **for**] [TP *him* [T to] [VP be quiet]]]

[5] [vP John [v *break-Ø*] [VP2 **windows** [V ~~break~~] [VP1 ~~John/him~~ [V ~~break~~] ~~windows~~]]]]

 (where VP2 'windows' received Acc case (*windows = them*) via light verb probe).

Regarding Case assignment, similar to 'long distance AGR' as cited in [1-2] above, we note that there is also no apparent Spec-Head relation for Acc case assignment—namely, <John> raises up but remains within vP to be in union with higher probe found in T. This calls for a Head>Spec relation where two phrases are involved. For a potential interpretation of a (non)-Spec-Head Probe-Goal Union (PGU), see Miyagawa: 35, though Miyagawa 's analysis keeps to notions of the traditional Spec-Head relation for PGU).

Head>Spec relation for Probe-Goal: (between Spec and closest above functional Head):

[6]

```
                    X'
                  /    \
                X        YP
                |         /
              probe    Spec
                        |
                      Goal
```
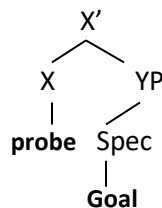

Chomsky (2006) offers theoretical arguments in support of positing that agreement features originate on a phase head and then percolate down onto the head beneath it (whereby the Spec of the lower Head enters into a checking domain). (Italics belong to Radford : 402).

Merge Over Move may also be considered as an example of a constraint on movement:

[7]   a. There seems __to be a man in the garden

          b.   *There seems a man$_i$ to be __ t$_i$ in the garden.

Summary

In the case of (1) above, the subject of the existential sentence is the DP 'two men' ('there' must take on the AGR features of 'two men'). But the two elements are NOT in a Spec-Head configuration.

In the case of [3-5] it rather seems Case can enter into a PROB-GOAL relation driven by C-command (not Spec-Head) where the Goal is not required to enter into Spec above the probe head. One further assumption made (Radford : 284) is that **the probe-goal relation only involves the highest functional ahead above the goal** (and not the forcing of the goal into the Spec position above the probe. What this means is that e.g., if T is the closest functional head above a vP-subject, then the subject needn't raise out of spec vP for case. Therefore, subsequent subject raising out of Spec vP and into Spec TP must be motivated by AGR (other approaches cite the EPP, or an Edge feature in Spec of TP which requires checking of a raised subject into Spec TP).

In the case of (7) above, we assume feature-checking is required on the infinitive T (to be) , but then why doesn't this motivate movement of the indefinite DP as found in (7b)? What are the mechanisms that suppress movement in (7a) yet block movement in (7b)?

> In sum: The problem with a Spec-Head AGR relation is that the DP-subject 'two men' (the GOAL) doesn't seem to get displaced and move up the tree so to enter into a local Spec-Head relation with the [T] Head 'do' (the PROBE). Such a local configuration would necessarily yield the derivation found in (8) below where the DP-subject would be required to move up the tree and position as Spec of TP. While this is a possible derivation involving argument (A-movement) raising, it rather seems such movement is not forced and the expletive 'there' can remain in Spec TP with the true DP-subject remaining vP in situ below.

[8] Two men do seem to be in the garden.

> ⇨ **Question: based on the above observations,   if movement doesn't seem to be Agreement-motivated, what other types of factors might we employ for getting an element to move into a local configuration—'e.g., regarding a Probe-Goal relation as found in (9) below, where 'which windows' must receive Acc case via Probe 'break'. How might AGR be different from Case in this respect? What other factors might be at work in motivating the wh-expression to move?**

> **Possible references:**

> 1. **Probe-Goal Union (PGU): (Miyagawa 2010: 35, see no. 10 'Snaps on syntax'): Miyagawa makes reference to the observation that a goal must move in order to unit with its probe. While Spec-Head is no longer a union configuration, what else could we say creates locality for PGU? (See *Note below).**
> 2. **Chomsky (2001: 13) remarks that P-G must be local in order to minimize search.**
> 3. **(Recall here that Radford suggests that the Goal is attracted by the closest above functional head (but may not need to move above that functional head, Radford: 286)**

[CP Which window did [TP __ [vP t ~~which window~~ [VP t John break t ~~which window~~]]]?

[9]

```
                        vP
                      /    \
                  Spec      v'
                          /    \
                         v      VP2   assume an ergative structure here:  (' (which) windows break')
                       /      /   \
                 break-Ø  Spec     V'
                          |       /   \
                 which windows   V     VP1
                                 |    /   \
                              break  Spec   V'
                                     |     /  \
                                John/him  V    N
                                          |    |
                                       break  which window
```
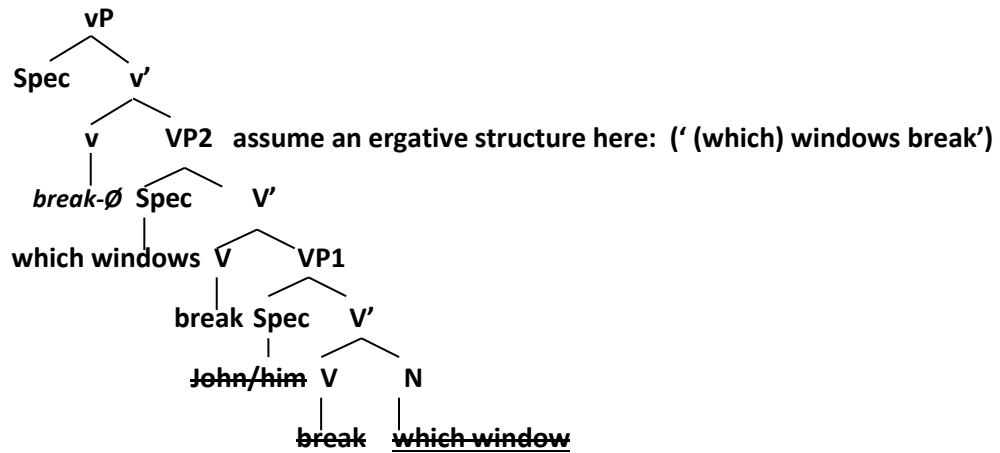
*Note: Probe-Goal relation (PGU)** (using MP terminology) for the checking-off of features will follow a dual-mechanism route:

  (i)   Probe-Goal relations of a case/thematic/semantic nature secured by local movement (e.g., vP—handling *case & argument structure*), and

  (ii)   Probe-Goal relations of a syntactic nature secured by distant movement (e.g., CP—handling *expression structure*). In addition to these two important phrases (known has *phases* since material from each phase must be independently transferred to PF and LF interpretation), we will consider how the two movements work together informing a cohesive syntactic structure. Any putative lack of distant movement at early stages of child syntax must surely impact the speech of a child.

⇨ **Ex.11 'Four Sentences'** (Link to 'four sentences' is found in the end section 'Other Sources').

Sentence-1: 'Can eagles that fly swim?

                        v1   v2

        i. [Can eagles that fly swim?]

So, if we are simply scanning strings via a process which only adheres to the 'adjacency-factors' of the string, then we should interpret that we are asking 'can eagles fly?' But let's consider now what sentence-1 looks like under a recursive structure:

        ii.   [$_x$ Can eagles [$_y$ that fly $_y$] swim $_x$]

Now, if we consider the nature of recursive structures (as found with embedded strings), then we can see that indeed the closest verb to the subject [Eagles$_x$] (found within the x constituency, or unit of structure), is in fact [swim $_x$] and not [fly $_y$]. As Chomsky puts it, it rather seems that it is due to some unique design of our human brain (a brain which gives rise to language) that allows us to instantiate immediately upon recognition (an innate recognition) the underlying recursive structure of [[ ]] over a flat structure [ ].This recognition is knowledge not learned in school, nor is it taught to us by our parents at an early age, but rather, comes for 'free' out of the *human design* of language.

Sentence-2: 'Him falled me down' (1960s child language studies)

In considering sentence-2, the item we are interested in here is the **over-regularization** of the verb 'fall' => 'falled' (fell). If we were, again, to take the *naive* flat assumption that all words are memorized, stored and retrieved as holistic chunks, in other words as [falled], then the immediate problem surfaces as to where and how the child ever came across such the word, being that it is not supported by the input. This very question goes to the heart of what Chomsky referred to as the **creativity of language.** Berko's work on child language quickly saw that such errors in fact proved that the child was working under a rule-based design of language, and that at roughly the point where over-regularizations take place in the child stages of acquisition, we find that the over-regularizations align with the acquisition of the rule—viz., [[N] + s], for plural, and [[V] + ed] for past), noting that such 'errors based on rules' supports recursive structure. Hence, what we have here with such errors is a decomposed item of [[stem]+affix] e.g., [[fall]ed] whereby the two parts of the words must be stored in distinct units or constituencies as found in the morphology (stem, inflectional morphology).

Sentence-3a: 'The horse raced past the barn fell'.

Sentence-3a is also known as a '*garden-path'* sentence. (The classic sentence and its first use is attributed to Tom Bever. In such designed constructs, readers are often lured into parsing (processing) the structure of a given sentence in a certain way, and by doing so is actually led down a wrong syntactic reading of the sentence (i.e., down a 'garden path')—viz., in believing that a grammatical element should follow based on what came prior. In other words, the erroneous assumption is tied to a processing which reads the first verb parsed '*raced*' as a past tense main verb of the matrix subject '*horse*', rather than how it should alternatively be processed, as an embedded passive past-participle of an covert embedded clause (The horse—that was **raced** past the barn—fell). This nice parsing trick shows how the brain seeks to parse and process pieces of syntactic phrase structure in systematic ways, in ways which speak **to phrase-structure rules,** (and in more current theory) **X-bar syntax.**

When the reader first hears and confronts the designed parsing of an initial DP, say 'the horse' (in the above garden path sentence), the DP immediately gets assigned as subject—this is done in concord and under syntactic X-bar theory, assuming that the syntax of the given language is SVO (subject-Verb-Object). Fine, but what this also means is that the following verbal item usually gets assigned as a Tense verb which then, due to phrase-structure rules, determines the Tensed verb to be a matrix predicate of the subject. The phrase structure design would read as follows:

S (sentence) ➜   DP, TP…

But this reading is false. The first Tensed Verb item **raced** does not relate to the predicate of the subject, but rather is part of an embedded structure which should rather be parsed accordingly:

[S [DP The horse] [~~that was~~ raced past the barn] [fell]]

```
                  TP  (Tense Phrase = Sentence)
                 /                    \
              DP                        T'        (= predicate)
             /  \                      /  \
          DP     CP                  T      VP
           |    /  \                        |
    The horse  C     VoiceP                fell
               |    /    \
            that  Voice    VP
                   |      /  \
                 was    V      PP
                        |     /    \
                     raced   past the barn
```
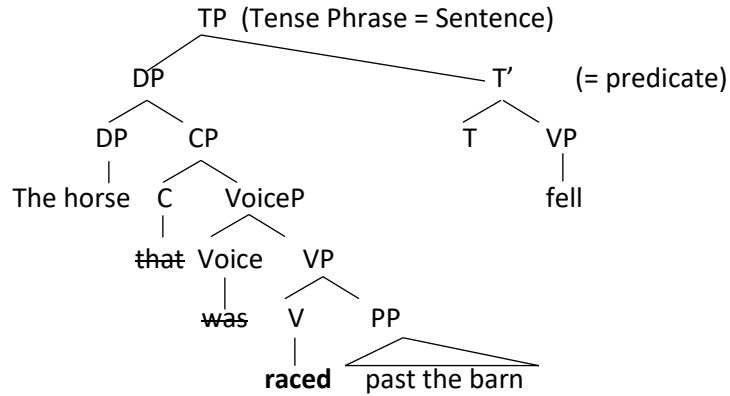
(where Voice P = Voice Phrase for passive voice *was raced*).

There is a question of **binding & licensing** here which closely relates to C-command. Although binding and licensing is usually called upon to show anaphor/antecedent relations, (as well as polarity expressions), what the structure above shows is that the same types of conditions and constraints which speak to binding & licensing can equally serve us here in explaining how garden-path constructs come to be analyzed. For example, let's slightly extend the garden-path sentence to read 'The horse raced past the barn fell to the ground'. Now what we discover is that the Preposition Phrase (PP 'to the ground') can only be *bound & licensed* by the verb 'fell' (as an adjunct/argument of the verb <fall <to the ground>> , and not the embedded verb 'raced' *<race <to the ground>> (e.g., …fell/*raced to the ground)). (See structures below).

```
 i.        VP <fall <to the ground>>        ii.        * VP      *<race <to the ground>>
          /    \                                      /    \
         V      PP                                   V      PP
         |     /   \                                 |     /   \
       fall   P     DP                             race   P     DP
              |    /  \                                    |    /  \
             to  the ground                              to  the ground
```
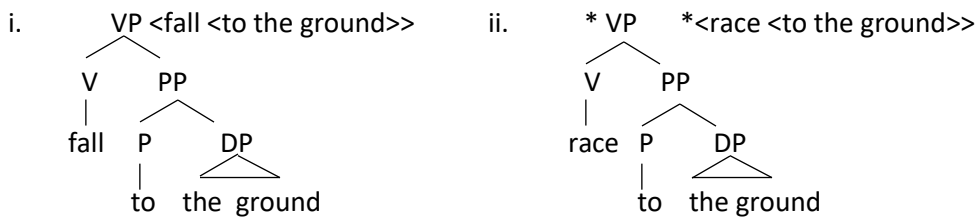
Sentence no. 3b 'The boy Bill asked to speak to <u>Mary thinks</u> he is smart'

In sentence-3b, consider how we actually find the opposite effect from that of sentence-3a (the garden-path sentence). In 3a the closest adjacent verb (*raced*) as pronounced in the utterance took precedence over a more distant verb (fell), hence the wrong assumption was made that the '*horse fell*' rather than the '*horse raced*'. Adjacency wins out in processing in such garden-path structures. On the other hand, sentence-b, **when read outload**—as opposed to simply reading it silently which doesn't give you the wanted effect (recalling that the natural skills, 'speaking' and 'comprehension' provide the underlying structure while the artificial culture-bound skills are rather learned and do not necessarily provide the underlying language structure)—instantly leads to the correct assumption that it is *The boy* who is doing the *thinking* rather that *Mary* who is doing the *thinking,* despite the fact that <u>Mary thinks</u> is adjacently placed together which might otherwise trigger a frequency-effect. Consider how Sentence-3b must be structured below:

[The boy [Bill asked to speak to *<u>Mary] thinks</u> he is smart].
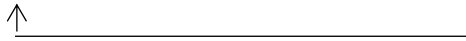
a. *Mary thinks he is smart.
b. The boy thinks he is smart.

When we read the sentence aloud, we go against adjacency of <u>Mary thinks</u> and rather, via an instinct level of processing, we naturally understand that it is 'the boy' further down and far removed in the tree that is the subject of the verb 'think'. Such types of examples give linguists evidence that native speakers of a language at times (actually quite often) go against surface-frequency or statistical-probability analyses as would be presented in the actual surface data. In other words, native speakers of a language go beyond surface data made available in the input, and rather rely on deep-hidden structures which may not always be evidenced in the pronunciation.
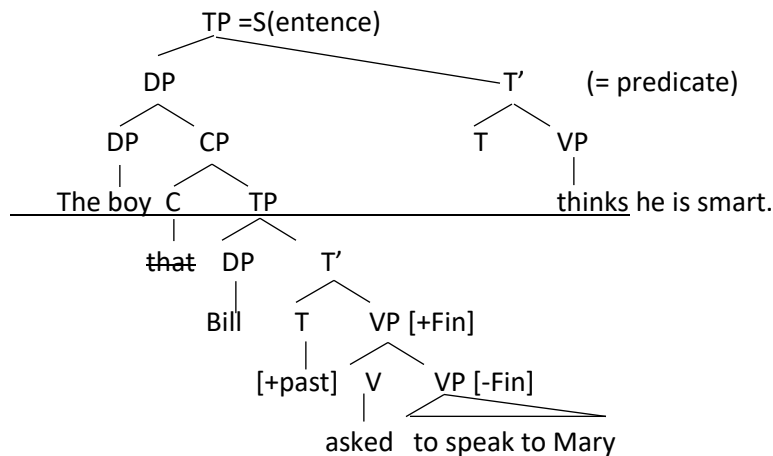
Sentence-4: 'I wonder what that is ___ up there'

David Lightfoot (2006, p.52) beautifully shows how a recursive-movement analogy of [[ ]] is both psychologically and indeed physically captured by the following simple illustration, showing the merge/move sequence as discussed in the *Overview*. Consider the 'is-what' phrase in the sentence 'I wonder what that is up there'. The base-generated structure first looks something like: I wonder [__ [that [VP *is what*]]] up there.

…and where the Wh-object 'what' begins as the object/complement of the verb 'is' and then gets displaced by moving above 'that' in the surface phonology (PF), yielding the derived structure. But if we take a closer look, we see that after such movement of 'what' out of the [VP 'is-what'] phrase, the VP survives only as a head [VP is ø] and is without its complement 'what'—thus the phrase 'partially projects'.

But *partial* phrase projections are allowed given that their Heads still remain (in *situ*) within the constituent phrase, hence, we get the licit structure in (a):

a. I wonder [what$_j$ [that [VP is __$_j$]]] up there?

b.*I wonder [what$_j$ [that's$_k$ [VP __$_k$__ $_j$]]] up there?

But movement has an effect: note how the head 'is' must remain phonologically intact as a head of the VP and can't become a (phonological) clitic attached to the adjacent 'that', as in [that's]. In other words, at least one of the two lexical items within a phrase (P) must be pronounced (be projected). Hence, as we see, when both items [is] as well as [what] move out of the VP ('What' moving into a Spec of a higher P along with the item [is] moving out of its head (H) position of the P and (forming itself as a clitic) piggy-backs onto the item [that] of the higher P, we see the result that the VP becomes vacuous (completely empty) and so the structure cannot survive (it becomes ungrammatical).

> **Moved-based** *[[that]'s] is an illicit structure while **Merge-based** of the two words [that] [is] is the only licit structure. It seems simultaneous movement of both head 'is' along with its complement 'what' of the [VP is-what] renders the verb phrase vacuous [VP ø] (i.e., phrases can't be both headless and complementless). In this sense, MOVE-based *[[that]'s] is barred and only Merge-based of the two items [that] [is] is allowed to project—the former (move) being affixal in nature, the latter (merge) lexical. This 'merge vs. move' treatment is similar to what we find with the distinction between (merge-based) **Derivational** vs. (move-based) **Inflectional** morphology, where the former is an affix process, and the latter a word-forming process.

Progression of structure:

(a)   'is-what' = VP (Verb Phrase)

VP   When object 'what' moves up, it leaves Complement/Obj of Head V

still intact, still allowing a licit projection of VP.

V   Obj
| |
is   ~~what~~

135

(b)   XP
         / \
        Y   X'
      /   / \
   what   X   VP   (VP head is filled with V 'is', so VP projects)
          /   / \
        that V   Obj   But note how 'is' must remain as a full word and not as a clitic.
              |
             is   ~~what~~

(c)   * XP
         / \
        Y   X'
      /   / \
   what   X   VP   (When V 'is' is reduced to a clitic ['s], the VP becomes vacuous
          /   / \
       that's V   Obj   (i.e., both V and Comp are empty) and so the VP can't project).
              |
             ~~is~~ ~~what~~   (* Vacuous/Empty VP becomes illicit: [*VP [v _] [Obj _]]).
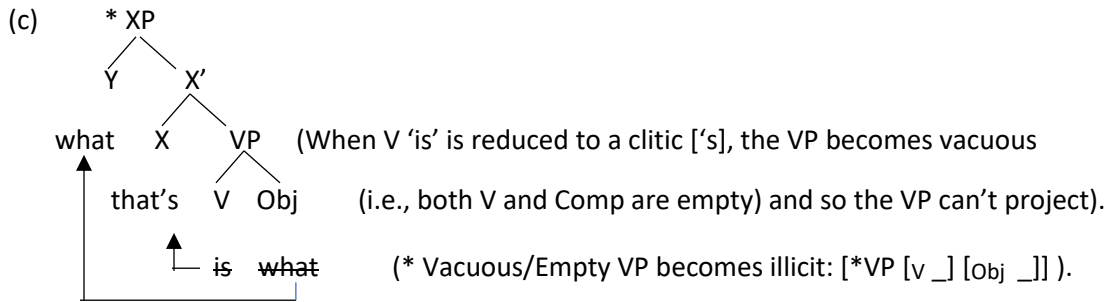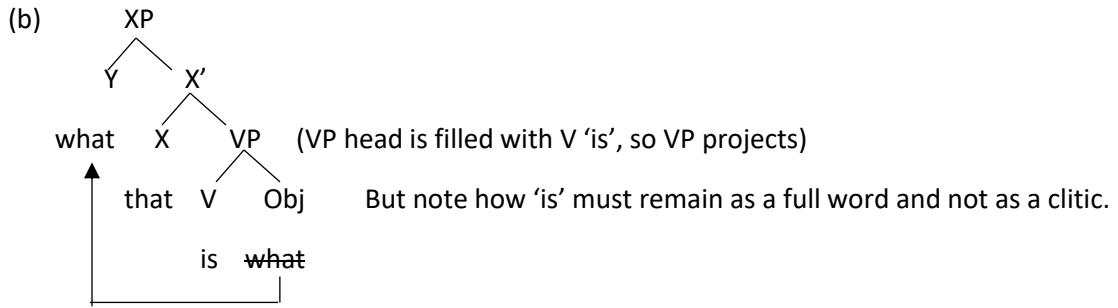
In the example above, now both words have moved: a) 'what' up to a higher position (of XP), and 'is' up to the adjacent word 'that' (say, as Head X of a higher phrasal projection XP).

Appendix: Snapshots on Syntax

1. X-Bar Theory: Fragments and Co-ordination:

https://www.academia.edu/5761529/Ling_610_Seminar_in_Syntax_Syllabus

2. Snapshots on Movement: Locality vs Distance:

https://www.academia.edu/10344372/Snapshots_on_movement_opening_remarks_Ling_610

3. Clauses, Phrases & Phases:

file:///C:/Users/jgjos/Downloads/Classs_Notes_on_Phrases_Clauses_and_Phas.pdf

4. C-command:

https://www.academia.edu/6021575/A_Note_on_C_Command_weeks_3_4_cont

5. Verb Phrase Internal Subject Hypothesis (VPISH)

https://www.academia.edu/10850538/Week_4_Notes_on_article_Subject_Position_VPISH_Koopman_Sportiche_the_Position_of_Subjects_Lingua1991_click_on_scribd_link_to_view

6. Merge

https://www.academia.edu/6108862/A_Note_on_Merge_week_5

7. VP-Shells

https://www.academia.edu/10943661/Week_5_Larson_1988_VP_Shells_split_VPs_click_on_Download_PDF_larson88do_pdf_to_view_paper

8. Antisymmetry of Syntax

https://www.academia.edu/11278860/Week_6_Kaynes_Antisymmetry_of_Syntax_Merge_vs_Move_Unaccusative_Verbs_and_Split_VP_Shells

9. Merge, Tense

https://www.academia.edu/6191790/A_Note_on_Merge_Tense_week_6

10. A Dual Probe-Goal

https://www.academia.edu/11430217/Weeks_7_8_Miyagawa_A_Dual_Probe_Goal_Status_A_word_on_Case

11. Case and Movement, EPP and ECM

https://www.academia.edu/6273853/A_Note_on_Case_and_Movement_EPP_and_ECM_weeks_7_8

12. Case Movement, Double Objects

https://www.academia.edu/6424451/A_Note_on_Case_Movement_and_Double_Object_Structures_week_9

13. Case marking and Null constituents

https://www.academia.edu/6557476/A_Note_on_Case_marking_and_Null_Constituents_weeks_11_12

14. Head to Head Movement

https://www.academia.edu/6822675/A_Note_on_Head_to_Head_Movement_week_13

15. Clitic Climbing

https://www.academia.edu/6940983/Clitic_Climbing_week_14

16. Deletion

https://www.academia.edu/6965503/A_Note_on_Deletion_week_15

17. Why Move: Reference paper

https://www.academia.edu/6508606/Why_Move_Reference_Paper_Exam_1_

(§§18-23 taken from monograph From Merge to Move, Galasso 2016).

18. Overview on Theory.

https://www.academia.edu/34403452/Working_Papers_1_Minimalist_Perspectives_on_Child_Syntax

19. Opening Remarks on the Minimalist Perspective.

https://www.academia.edu/34403450/Working_Papers_2_Minimalist_Perspectives_on_Child_Syntax_Opening_Remarks_and_Introduction

20. Why Move?

https://www.academia.edu/34403445/Working_Papers_3_Minimalist_Perspectives_on_Child_Syntax_Why_Move

21. The Minimalist Program Framework

https://www.academia.edu/34403441/Working_Papers_4_Minimalist_Perspectives_on_Child_Syntax

22. 'Merge over Move' Child syntax:

https://www.academia.edu/34403440/Working_Papers_5_Minimalist_Perspectives_on_Child_Syntax_Merge_Over_Move_Movement_Application_in_Child_Syntax

23. Labels & Phases: Ex. For multi-level phase intersections of a sentence 'We are breakfasting'

https://www.academia.edu/34403438/Working_Papers_6_Minimalist_Perspective_on_Child_Syntax_Other_Theoretical_Implications_Labels_Phases_and_Treelet_structures

24. Paper on Developmental implications that the lack of movement has on Phase and Spell-out: 'Small children's sentences are dead on arrival'

https://www.academia.edu/15155921/Small_Children_s_Sentences_are_Dead_on_Arrival_Remarks_on_a_Minimalist_Approach_to_Early_Child_Syntax_Journal_of_Child_Language_Acquisition_and_Development_vol_3_no_4_Dec_2015_

References:

Abney, S. (1987) The English noun phrase and its sentential aspect. Ph.D. Diss. MIT.

Bertram et als. (2000) The balance of storage and computation in morphological processing. *J of experimental psychology: learning, memory and cognition. 26, 489-511.*

Chomsky, N. (1957). Syntactic Structures. The Hague: Mouton.

___(1995) The Minimalist Program. MIT Press.

___(2001) Derivation by Phase. In Ken Hale A life in language, ed. M Kenstowics, 1-52. MIT Press.

___(2002) On Nature and Language. Cambridge University Press.

Diamond, J. (1992) *The third chimpanzee*. New York: HarperCollins.

Fitch, T., Huaser, M. & N. Chomsky (2005). The evolution of the language faculty: clarifications and implications. Cognition 97, 179-210.

Galasso, J. (2013). Minimum of English Grammar: An Introduction to Feature Theory, Vol. II. Cognella Publications: San Diego.

___(2015) Small Children's sentences are dead on arrival. *Journal of Child Language Acquisition and Development*. Vo.l.3, issue: 4 184-226.

   https://www.academia.edu/15155921/Small_Children_s_Sentences_are_Dead_on_Arrival_Remarks _on_a_Minimalist_Approach_to_Early_Child_Syntax_Journal_of_Child_Language_Acquisition_and_ Development_vol_3_no_4_Dec_2015_

___(2016). From Merge to Move. LINCOM Studies in Theoretical Syntax, 59. Munich.

___(2019). Recursive Syntax. LINCOM Studies in Theoretical Syntax*, 61*. Munich.

__ (2019) A Note on Artificial Intelligence. Ms CSUN.

   https://www.academia.edu/39578937/Note_4_A_Note_on_Artificial_Intelligence_and_the_critical_ recursive_implementation_The_lagging_problem_of_background_knowledge_1

___(2021). Reflections on Syntax. Peter Lang: Oxford.

Kayne, R. (1994) The antisymmetry of syntax. MIT Press.

Lasnik, H.___(1995) Verbal Morphology: Syntactic Structures Meets the Minimalist Program. MS. University of Connecticut.

Lasnik, H. & M. Saito (1992) *Move alpha: Conditions on its Application and Output*. MIT Press.

Marantz, A. (2001) Phases and Words. Ms, NYU.

Marcus, G. (2001) The algebraic mind. MIT.

McClelland & Rumelhart (1988) Explorations in parallel distributed processing. MIT

Moro, A. (2000) Dynamic Antisymmetry. MIT Press.

Pinker, S. (1999) Rules and Rules. Basic Books.

Radford, A. (1990) Syntactic theory and the acquisition of English syntax. Oxford: Blackwell.

_____(2019). Relative Clauses. Cambridge Studies in Linguistics, 161. CUP.

Seidenberg & McClelland (1989) A distributed developmental model of word recognition. *Psychological Review, 96, 523-568.*

Uriagereka, J. (2000) Rhyme and Reason. MIT Press.

White, R. (1989) Visual thinking in the ice age. *Scientific America* 260(7): 92-99.

Other Sources:

•Optional links to PDF lectures: 'Seminar in Syntax' (CSUN, 2014):

https://csun.academia.edu/josephgalasso/Seminar--Syntax-(Ling-610)

•For an alternative account on Case marking, see link:

https://www.academia.edu/6192388/Ling_610_Alternative_Case_Marking_galasso_Case_Theory_and_Phrase_Expansion

•On a Dual Probe-Goal relations (Miyagawa)

https://www.academia.edu/11430217/Weeks_7_8_Miyagawa_A_Dual_Probe_Goal_Status_A_word_on_Case

•Subject position:

http://www.scribd.com/doc/247431800/Koopman-Sportiche-the-Position-of-Subjects-Lingua1991

Author's Autobiographical Note.   Joseph Galasso published two books in 2016 and 2013 and a recent 2015 article in addition to a 2013 commentary on the historic changing attitudes toward higher education in Academe, the publication of the American Association of University Professors. A well-respected syntactician in the tradition of Noam Chomsky, Joseph Galasso has two recent publications on young children's utterances that postulate the nature of the innate language capability based on actual data: "Small Children's Sentences are 'Dead on Arrival': Remarks on a Minimalist Approach to Early Child Syntax" in Journal of Child Language Acquisition and Development and the book "From Merge to Move: A Minimalist Perspective on the Design of Language and its Role in Early Child Syntax" (LINCOM Studies in Theoretical Linguistics, 59). A second book is a text on the "Minimum of English Grammar", an approach to syntax based on insights from language acquisition, language disorders, and brain functions that has been adopted by several universities, including Hofstra University, NY. A follow-up monograph in the same LINCOM series was published in 2019 entitled "Recursive Syntax: A Minimalist Perspective on Recursion as the Core Property of Human Language and its Role in the Generative Grammar Enterprise: (with special notes on Dual Mechanism Model, Problems of projection, Proto-language, Recursive implementation in AI, and the Brain"). His last book "Reflections on Syntax" (Peter Lang: Oxford) was published in 2021.

His main research involves issues surrounding early child language development. He is interested in pursuing certain "Minimalist Program" assumptions (Chomsky 1995) and to ask how such assumptions might explain observed early stages of morphosyntactic development in children. His research specifically asks how/when the requirements and conditions placed on 'Merge' over 'Move' operations come on-line in child language and whether or not these operations are open to maturation factors having to do with a brain-to-language corollary. His work has appeared in "The Oxford Handbook of Developmental Linguistics" (2016) (eds. Jeffrey L. Lidz, William Snyder, and Joe Pater).

Index