Chomsky (2007): "Approaching UG from below" [AUG]

Theme: Presupposes the framework laid out in "On phases" (Chomsky 2008), which had been circulating as a manuscript version before AUG in 2005. Summarises and slightly updates Phase theory.

a. *Language design*[1f.]

The problem problem: identify the specific nature of the human FL ("a cognitive organ"), a species property

Language = I-language = a state of FL

UG = theory of the initial state of FL (S_0)

"Biolinguistics" ⇒ interdisciplinary approach to language initiated by Piatelli-Palmarini in 1974

Metaphysical issues, not epistemological ones

From methodology to empiricism

Language acquisition as theory construction (as of 1974; acc. to Peirce's philosophy of science)

Tension between descriptive and explanatory adequacy ⇒ P&P (Chomsky 1980, 1981; inspired by Jacob: invariant principle + parameters ⇒ surface variability) ⇒ simplification of the format of UG

"Principles of neural organization that may be even more deeply grounded in physical law" (Chomsky 1965: 59; 3rd factor) ⇒ evo-devo

b. *Three factors in language design (constraining the hypothesis space)*[3f.]

1. Genetic endowment

2. External data

3. Principles not specific to FL

3rd factor: efficient computation in generative systems (possibly also properties of the human brain; cf. fn. 4) ⇒ "beyond explanatory adequacy" ("Not only what, but also why?")

MP: a *programme*, not a theory ⇒ a bottom-up approach: "How little can be attributed to UG while still accounting for the variety of I-languages attained, relying on third factor principles?"

SMT: FL is "perfectly designed"

"[A]n account of some linguistic phenomena [is] principled insofar as it derives them by efficient computation satisfying interface conditions. [A]ll phenomena of language have a principled account in this sense, that language is a perfect solution to interface conditions"[5]

c. *Merge*[5ff.]

Merge ⇒ discrete infinity, compositionality, strict cyclicity (computationally efficient), nested (not crossing) dependencies

(1) The men who John V_1V_2… ⇒ *Agree(men,*V_2*)*, *Agree(John,*V_1*)*

Parallel operations: "[A] generative system involves no temporal dimension"[6] ⇒ Top–down grammars possible

Generation ≠ production (competence ≠ performance) ⇒ E.g., bottom–up interpretation (= generation) compatible with an XP's linearly initial position (in performance)

Pritchett (1992, *Grammatical Competence and Parsing Performance*)[fn. 7]

Grammar-derived (i.e. GBT-derived) parsing theory ⇒ head-driven parser ("inputs are stored until a lexical head (verb) is reached" [Reinhart 2006: 9]); *Theta-attachment*: „[A]t any moment the processor attempts to map as many thematic roles onto as many arguments as possible"; *generalized theta attachment:* "every principle of the syntax attempts to be maximally satisfied at every point during processing" (cf. the θ-Criterion; quotes from Altmann's 1993 review; example from Reinhart 2006: 9*f.*)

(2) Max arrived

Max ⇒ STORE SUBJECT ⇒ *arrive*$_{<\theta>}$ ⇒ ATTACH Θ TO SUBJECT ⇒ *Once the θ-grid of a V is exhausted, 'reconstructive' bottom-up generation*: [VP *arrive Max*] ⇒ [TP *Max* [*arrived Max*]] ⇒ …

(3) a. Max knew Lucie would laugh.

Max ⇒ STORE SUBJECT ⇒ *knew*$_{<\theta,\ \theta>}$ ⇒ ATTACH Θ TO SUBJECT ⇒ *Lucie* ⇒ ATTACH Θ TO OBJECT ⇒ *would laugh* ⇒ **LOW COST REANALYSIS**: *[TP *Max* [VP *knew* [***Lucie*** [TP *would laugh*]]]] ⇒ ✓[TP *Max* [VP *knew* ~~*Lucie*~~ [TP ***Lucie*** *would laugh*]]]] (*Lucie* reanalysed within original domain = direct object)

b. Max warned Lucie would laugh. (*Garden path*)

Max ⇒ STORE SUBJECT ⇒ *warned*$_{<\theta,\ \theta,\ \theta>}$ ⇒ ATTACH Θ TO SUBJECT ⇒ *Lucie* ⇒ ATTACH Θ TO OBJECT ⇒ *would laugh* ⇒ **HIGH COST REANALYSIS** ⇒ *garden path effect (proper)*: *[TP *Max* [VP *warned* [***Lucie*** [TP *would laugh*]]]] ⇒ ✓[TP *Max* [VP *knew* ~~*Lucie*~~ [TP ***Lucie*** *would laugh*]]]] (*Lucie* reanalysed external to original domain = oblique object)

High cost reanalysis is better illustrated by:

c. After Susan drank(,) the water evaporated.

Merge(X_1, …, X_n) = Z, Z = 2 (binary Merge, 'unambiguous paths'), Merge(X, Y) ⇒ {X, Y}, NTC, IM vs. EM

Excursus: Chomsky rejects types of Merge beyond set- and pair-Merge[fn. 10]

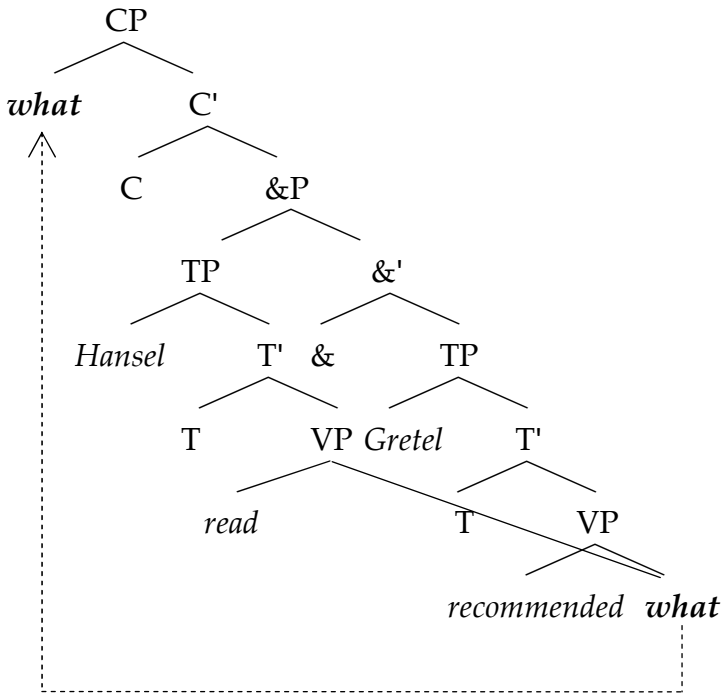
Citko (2005, "On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge")

Parallel Merge creates multidominance (copy-less structure-sharing) trees, where linearisation is achieved by movement (here, of *what*)

(4) *A Parallel Merge Approach to Across-the-Board Wh-Questions*

a. I wonder [what Gretel recommended *e*] and [Hansel read *e*].

b.



Multidominance/Parallel Merge structures provide an elegant account of matching effects in ATB-constructions (Polish ex. from p. 485):

(5) a. Kogo Jan lubi *e* a Maria podziwia *e* ?
who.ACC Jan likes *e*.ACC and Maria admires *e*.ACC
'Who does Jan like and Maria admire?'

b. *Kogo/Komu Jan lubi *e* a Maria ufa *e* ?
who.ACC/DAT Jan likes *e*.ACC and Maria trusts *e*.DAT
'Who does Jan like and Maria trust?'

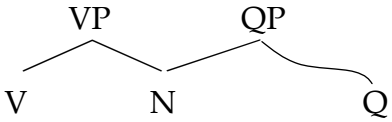
Svenonius (2005, "Extending the Extension Condition to Discontinuous Idioms", manuscript)

Goal: resolving "the apparent conflict between listing discontinuous idioms as syntactic structures and adhering to the Extension Condition" (p. 19)

"Merge cannot project a node which is already dominated" (contra tucking-in) ⇒ Both first Merge and IM can target subconstituents of structures already built by Merge ⇒ **Banyan trees** (an Indian tree growing multiple trunks) ⇒ Multi-rooted graphs allowing restricted form of sideward movement

Thus, Merge(D,VP) may yield

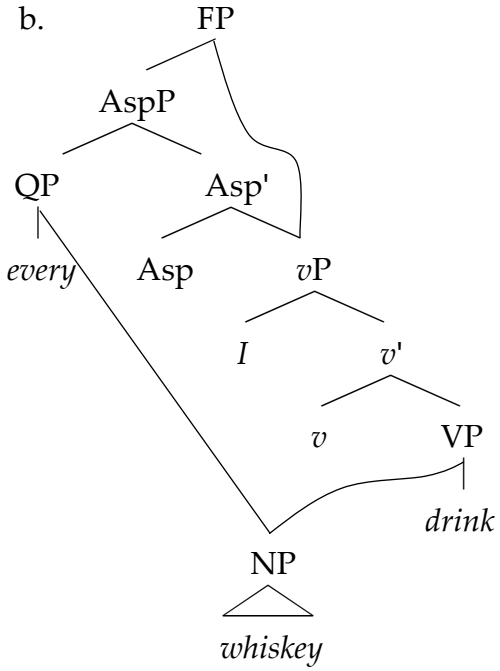
(6)



Squiggly branches indicate that branch is not drawn on the 'correct' side (contrary to (4) above; ex. from p. 21f.). QP (≈ DP) must move to [Spec, AspP] to check ACC (cf. Agr$_O$'s function).

(7) a. [$_{FP}$ $\underline{vP}$ F° [$_{AspP}$ [$_{QP}$ Q $\underline{whiskey}$] Asp° [$_{\underline{vP}}$ *I v*° [$_{VP}$ $\underline{whiskey}$ *drink*]]]]

b.



F° omitted; here, [$_{NP}$ *whiskey*] is first-merged in [Compl, VP] and remerged to [Spec, *v*P] for ACC-Case checking (as [$_{QP}$ *every whiskey*]). *v*P is remerged to [Spec, FP] for linearization purposes (*I drink every whiskey*). *I* will subsequently move to [Spec, TP].

d. *Labels*

Z = SO; Label(Z) = W; W = LI = H° = Probe

No m-command ⇒ No Spec–H relation generally

No c–command ⇒ CI: Agree-based BT (maybe not for scope); SM: LCA (c-command) needed, or another device (departing from SMT)

Fukui & Takano (1998, "Symmetry in syntax: Merge and Demerge")[9; fn. 13]:

Linearisation = Top–down Demerge

Head-parameter (as alternative to LCA): head-last = base position w/o movement; head-first = derived (opposite of LCA) ⇒ W/o appeal to empty dummy FF's

Note: "head movement for checking purposes always takes the form of 'substitution into Spec,' eliminating traditional 'adjunction to head'" (p. 44)

(8) a. [TP John-ga [vP t_{SU} [VP **hon-o** **yonda**] $v_{[-]}$] (koto). (*p. 33*)
John-NOM book-ACC read
'John read a book.'

b. [TP John [vP t_{SU} **read** [VP **a book** t_V] $v_{[+]}$] T]

Residue: Universal Spec ≻ H ordering remains unaccounted for by head-parameter, but supports LCA…

Alternative explanations for properties of ordering[10]

① Parsing-based account (e.g. rightward ordering of complex XP'S; cf. Pritchett's theory)

② Cartographic account: H carries unambiguous information for interpretation of Spec

Kayne (2004, "Antisymmetry and Japanese")[fn. 13]

Explaining typological 'cross-linguistic gaps' (some of Greenberg's negative universals, e.g. mirror German ≈ VOS²; also cf. Newmeyer's possible & probable languages) *qua* antisymmetrical LCA ⇒ 'Head-final/-initial' no primitive notions

Universal S–V–O (Spec–H–Compl) ⇒ Japanese is surface O–V ⇒ O must've raised higher than V

Indirect evidence: many O–V languages (seem to) have *visible* obligatory movement of O out of VP across an arguably VP-external element (so, word order variation *is* 'construction-specific' after all…?) (p. 5f.)

(9) *VP-external focus in Malayalam (Jayaseelan 2001: 40)*

S–O–**Foc**–V
ninne [FocP **aarə** [vP t_S aTiccu t_O]]? (*here*: O–Foc$_S$–V)
you.ACC who$_{Foc}$ beat.PAST
'Who beat you?'

(10) *Negation in West Flemish (Haegeman 2001: 215, simplified)*

S–O–**Neg**–V
da Valère [dat us]$_O$ [NegP **nie**$_{Neg}$ [vP t_S kocht t_O]].
that Valère that house not bought
'…that Valère did not buy that house.'

(11) *Infinitive markers (and separable particles) in German/Dutch*
O–**Prt**–**Inf**–V$_{inf}$
…[das Buch] **mit**–**zu**–[vP bringen t_{Prt} t_O].

e. *Copies*[10]

No distinct operation (just IM) ⇒ Phase-internal IM distinguishes copies (from independent repetitions: *John killed John*) [*Q*: *Only one copy per phase?*]

f. *Merge*[11ff.; fn. 16]

Edge feature licenses/implements Merge

Non-deletable *u*F ⇒ One (or more) Specs, rather than unidirectional H(P)–Compl sequences (thus edge *property* seems to be a more adequate designation…)

Compls = first-Merged; Specs = later-Merged

EF ⇒ Unbounded Merge

What's good for CI is bad for SM:

① CI: optimisation of expressive power (duality of semantics, A/A'-distinction)

② SM: Computational efficiency demands that only the highest copy be pronounced ⇒ Ease of communication (here, parsing) is complicated by reconstructing gaps

Nissenbaum (2000)[fn. 17]

Covert = overt phrasal movement, with *different timings* ⇒ Covert movement is *post*-Transfer ⇒ Covert movement is subject to the same constraints as overt movement is (e.g. the PIC, BT)

(12) *Condition A (pp. 124f.)*

a. Mary$_i$ knows [which picture of herself$_i$] John is looking at ___.

b. *Mary$_i$ knows John is looking at a picture of herself$_i$.

c. * Mary$_i$ knows [which man] was looking at [which picture of herself$_i$].
⇒ But: 'LF' = Mary knows [which picture of herself] [which man] was looking at ___.
⇒ Given LF, *Mary* should be able to bind *herself*, contrary to fact
⇒**Tucking-in at work!**

Primacy of the CI interface (also see Hauser et al. 2002; Chomsky 2005; Chomsky *forthcoming*) in language design (language design dysfunctional for language design; also cf. islands) ⇒ Sometimes resolved by resort strategies (e.g. resumptive pronouns)

Interaction syntactic structure & neo-Gricean communication principles[fn. 18]

Reinhart (2006, *Interface Strategies*)

Reference-set computation: "A hypothesis that got much attention in the 1990s is that the well-formedness of syntactic derivations is not always determined by absolute conditions, but it may be based on a selection of the optimal competitor out of a set of candidates – a reference set. […] [I]n computing a given derivation, an alternative derivation be constructed, in order to determine whether a given step in the current derivation (or the full output) is permitted." (Reinhart 2006: 1)

Other than in OT-syntax, "[i]t applies only as a 'last resort', when the outputs of core syntax operations are insufficient for the interface." (ibid.) ⇒ **Interface strategy** (for resolving QR, stress-shift, coreference, scalar implicatures)

Example QR (pp. 105ff.): Covert movement (QR) doesn't apply freely, only when necessary; thus, not for (*a*), but for (*b*)

(13) a. A doctor will examine every patient. (*Ambiguous*)

b. A doctor will examine every patient, and Lucie will [ ] too. (*Ellipsis; only narrow scope for* every patient)

"[T]he explanation for the ellipsis context must rest on what happens in the elided conjuncts." (p. 107)

(14) *Reference-set for (b) only*

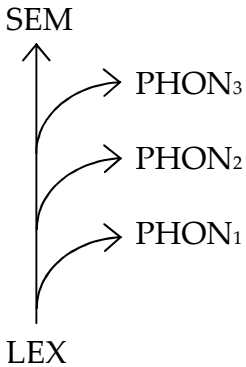
i. ⟨Every patient$_1$ [Lucie will [$_{VP}$ examine e$_1$]]
For every patient x, Lucie will examine x⟩
ii. ⟨Lucie will [$_{VP}$ examine every patient]
For every patient x, Lucie will examine x⟩

Chierchia (2004, "Scalar Implicatures, Polarity Phenomena, and the Syntax/Pragmatics Interface")
"[S]ome of the Grice-inspired pragmatics is probably part of the computational system of grammar. We figure out S[*calar*]I[*mplicatures*]s in tandem with 'core' meaning, as facets of a parallel recursion." (p. 59)

Petitto (2005, "How the brain begets language")[p. 13; fn. 19]
Sign language as a (partially) modality-independent externalisation/development: "Laura Petitto discusses research that localizes tissue in a part of the brain homologous to that found in several primates, tissue that used to be thought of as devoted to sound recognition but that in the case of humans seems to be language-specific, innately "programmed" to recognize, respond to, and lead to production of linguistic syllabic structure – syllabic structure, remarkably, in both speech and sign." (McGilvray 2005: 13)

Hinzen (2006): "A more radical conception of the FL–CI interface relation" ⇒ 'naturalization of meaning' (p. 235) ⇒ "[S]yntax carved the path interpretation must blindly follow" (Uriagereka 2002: 275) ⇒ D = SEM[15]

(15) *Architecture (p. 158)*



g. *Phases*[16ff.]

Strict cyclicity; phases; PIC; multiple Transfer (of Complements); no LF; labels = phase heads drive computation; phases: CP, *v**P

Phases: As large as necessary, as small as possible. RootP's (here, VP's) cannot be phases since they are still unsatisfied on their completion (wrt. category, Case, IM, etc.)

(16) *[v RP] (Marantz 1997: 217, 219)*

[$\sqrt{\text{DESTROY}}$ [the city]]]

i. [$_{v\text{P}}$ AGENT *v* [$\sqrt{\text{DESTROY}}$ [the city]]]

ii. [$_{\text{DP}}$ ('AGENT') D [$\sqrt{\text{DESTROY}}$ [*of* the city]]]

(17) *Argument structure (cf. Marantz 2005)*
[$_{\text{VP}}$ leave [the class]]

i. [The class *v* [left *t*]]

ii. [The teacher *v* [left [the class]]

On Reuland's syntactic BT (see handout on on OP): "Eric Reuland's discovery of locality-bound (hence syntactic) reflexivization in which the antecedent does not c-command the reflexive but both are c-commanded by the head that agrees with the antecedent"[18]

TP?

Richards (2007): "[I]n the *On Phases* model, we are forced to assume that the Agree features are on C but are passed on to T. Valued uninterpretable features cannot be distinguished from interpretable features, so no feature valuation may take place before Transfer. Therefore, only C may introduce Agree features. But because C itself is not transferred until the next cycle, the Agree features must be passed on to T, so that they can be transferred immediately." (Grewendorf & Kremers 2009: 424) ⇒ Conspiracy of uninterpretability & the PIC[19; fn. 25]

h. *C, T & FI*[20]

FI may leave a (morphological?) residue on C

(18) *ϕ-features*

a. Du häs dat Buch jekäuf, isch weiß ever nit [wo(**st**) du et jekäuft häs]. (*Kölsch; Kremers 2007: 4)*

b. …[ob*(-st) du noch Minga kumm-st]. (*Bavarian; Fuß 2008: 4)*

c. …[da(-nk) ik werk-en (*West Flemish; Shlonsky 1994: 353)*

that-1SG I work-1SG

'…that I work.'

AUG: "C never (to my knowledge) manifests Tense"[20]; but:

(19) *Tensed C (Modern Irish; from Chung & McCloskey 1987: 218)*

	Nonpast	Past
Subordinating	*go*	*gur*
Direct relative	*a*	*a*
Indirect relative	*a*	*ar*
Interrogative	*an*	*ar*
Matrix negative	*ní*	*níor*
Embedded negative	*nach*	*nár*

AUG: T's temporal interpretation (in terms of independent vs. dependent tense, finite vs. nonfinite Tense) depends on the presence of C ⇒ *Two options*: ① FI; ② The [T] feature of T "receives only some residual interpretation unless selected by C"[20], rather than being inherited from C ⇒ *Advantage*: T is no empty dummy LI (prior to FI), like Agr° was [*What's with the Irish data?*]

(20) *"English-like modal constructions"?*

Generalised FI: *v**-V ⇒ RtO

Asymmetries: ① V-to-*v** obligatory, T-to-C is not; ② T's semantic content is unclear (unlike V's), C is no categoriser

T is needed to make the sequence C–*v** discontinuous (otherwise FI C–*v** would lead to crash) ⇒ Generalised pattern: +PH ≻ –PH ≻ +PH ≻ –PH (Richards 2007: 570)

What's with TP in ECM/raising? AUG: Preventive mechanism

Nevins (2004, *Derivations without the Activity Condition*)[fn. 31]

The EPP operates autonomously, independent from ϕ/Case ⇒ The Activity Condition does not constrain EPP-movement

Structures that seem to be ruled out by the Activity Condition actually fall under different morphosyntactic constraints ⇒ Already structurally Case-marked DP's satisfy the EPP (as for English, Nevin's argument relies on more subtle/complex data…)

(21) *Finite Accusative Unaccusatives in Russian* (Lavine & Freidin 2002: 258; also OP: n. 65)

Soldata ranilo pulej.
soldier.ACC wounded[-AGR] bullet.INST
'A soldier was wounded by a bullet.'

[TP soldata[ACC] T*def*[EPP] [*v*P ranilo[ACC] [VP $t_{soldata}$ t_V pulej[OBL]]]]

i. *Intermediate copies*[22f.]

(22) *Binding effects*

a. ✓John seems to her [*John* to appear to *himself* to have left].

b. ✓John seems to her [*John* to appear to *her* to have left].

c. *John seems to her [*John* to appear to *herself* to have left].

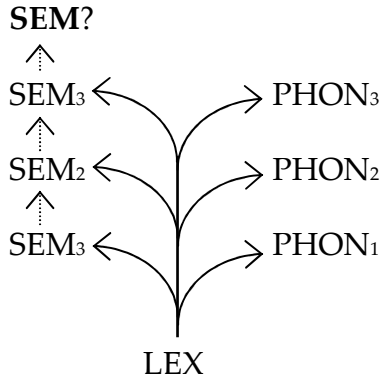
d. *John seems to her [*John*$_i$ to appear to *him*$_i$ to have left].

e. *John seems to her$_i$ [*John* to appear to *Mary*$_i$ to have left]. (Condition C violation)

(23) *Scope (Sauerland 2003: 310)*
Every child$_i$ doesn't *every child* seem to his$_i$ father [to be smart].
NEG >> *every child*

Note: To allow for trans-phasal binding (e.g. the Reuland case [C…T…[XP R]], e.g. *There became a man introduced to self*; see handout on OP) "[t]he structure mapped to CI [*is*] preserved for CI interpretation[24] [*But aren't configurations like Reuland's a single strong phase anyway?*]

(24) *Architecture?*



j. *A/A'-movement*[24f.]

(25) *Principled explanation for Improper movement*
*Who [*who*$_4$ seems [*who*$_3$ C [*who*$_2$ T-is *who*$_1$ smart]]]?
⇒ *who*$_3$ (= head of A'-chain) is inactive, hence invisible to matrix [Spec, TP]

DP?[25f.]

Nominals are ±*definite* ⇒ ±presence of D°: +*definite* ⇒ D°; –*definite* ⇒ no D°

(26) *Indefinite nominals*
a. [$_{n\text{P}}$ *an n*° [$\sqrt{\text{AUTHOR}}$ [(YP)]]] ⇒ [$_{n\text{P}}$ *an* $\sqrt{\text{AUTHOR}}$+*n*° [$\sqrt{\text{AUTHOR}}$ [(YP)]]]
b. [$_{n\text{P}}$ *many n*° [$\sqrt{\text{AUTHORS}}$ [(YP)]]] ⇒ [$_{n\text{P}}$ *many* $\sqrt{\text{AUTHORS}}$+*n*° [$\sqrt{\text{AUTHORS}}$ [(YP)]]]

(27) *Definite nominals*

(28) [$_{n^*\text{P}}$ *n** [$_{\text{DP}}$ D [YP] ⇒ [$_{n^*\text{P}}$ D+*n** [$_{\text{DP}}$ YP D [YP]]] (parallel to *v**P's)

⇒ +*definite* nominals are *n**P's, not DP's ⇒ **All nominals are *n*P's!**

Q: Given that YP's supposed to be an optional Compl (e.g. [PP *of the book*]), where's the root element in strong nominals (= *n**)? D? PP? How (if at all) is *n** a categoriser for X = D°?

(29) *Causative verbal (?)*
[$_{v*P}$ *v** [CAUSE [V Compl]]] ⇒ [$_{v*P}$ CAUSE+*v** [[V Compl] CAUSE [V Compl]]]