

Overview of Chomsky's Binding Theory 1973-1986

A. "Conditions on Transformations" (1973) <The origin of modern binding theory>

(1) *The dog is believed [___ is hungry] (cf. The dog is believed [___ to be hungry])

(2) Tensed Sentence Condition (TSC) (1st version)
 No rule can involve X, Y in the structure
 ...X...[α ...Y...]
 where α is a tensed sentence [i.e., a finite clause]

[[In previous generative work it was assumed that all clause boundaries block most relations, included the NP-movement part of passive. Chomsky was now rejecting this point of view.]]

(3)a The candidates each hated the other(s) (cf. The candidates are hated ___)

b The candidates each expected [the other(s) to win] (cf. The candidates are expected [___ to win])

c The candidates each expected [that the other(s) would win]

(3)a' The candidates hated each other

b' The candidates expected [each other to win]

c' *The candidates expected [that each other would win]

(4) Passive (i.e., NP preposing) and each-movement obey (2). "...one rule that obviously does not satisfy the condition is Coreference Assignment..." as in (6).

(5) John said that he would leave <Much more on this later>

(6) The candidates expected to defeat each other

(7)a *The men expected the soldier to shoot each other

b *John is believed [Mary to like t] ___ [[Remember, 1973 is well before Case Theory was added.]]

These satisfy TSC, so an additional constraint is needed>>

(8) Specified Subject Condition (SSC) (1st version)

No rule can involve X, Y in the structure
 ...X...[α ...Z...-WYV...]
 where Z is the specified subject of WYV in α

<Basically subjects block relations across them. For **Specified** subject': See (18) below.>

(9)a The men saw [_{NP} pictures of each other]

b *The men saw [_{NP} John's pictures of each other]

c *J. is expected [M. to visit ___]

(10)a *I saw me

b *I saw us

c *We saw me

d He saw him

e The soldiers shot the officers (among them)

- (11) RI: "a rule of interpretation applying to the structure NP-V-NP (among other) [that] seeks to interpret the two NPs as nonintersecting in reference, and where this is impossible...it assigns 'strangeness'. cf. Postal's "Unlike Person Constraint". <The ancestor of LGB's Condition B.> <<As far as I know, Postal never used that term. In the 1970's, he called the constraint The Inclusion Constraint.>>
- (12)a We expect [them to visit me]
 b *We expect [me to visit them]
 c We believe [I may still win]
- (13) In (12)a,c, RI is blocked by SSC and TSC, respectively, so intersecting reference is allowed.
- (14) Combining TSC and SSC:
 No rule can involve X, Y in the structure
 ...X...[α ...Z...-WYV...]
 where (a) Z is the specified subject of WYV
 or (b) Y is in COMP and X is not in COMP
 or (c) Y is not in COMP and α is a tensed S
 <<(b) and (c) were to allow long distance wh-movement, where such movement is argued by Chomsky to proceed Comp to Comp. Comp was assumed then to be the target of wh-movement. Later, it would be Spec of Comp.>>
- <TSC and SSC got incorporated into the LGB binding theory, as we will soon see.>
- (15) "...under the analysis proposed here there is no necessity for a rule raising the subject of an embedded sentence to the object position of the matrix sentence..." <Chomsky is here rejecting the then standard view that all the relations in question (passive, reciprocal, RI) obey a 'clause-mate' condition. Under that point of view, ECM subjects had to raise to become matrix objects for the relevant relations to be permitted.>
- (16)a *We persuaded Bill [PRO to kill each other]
 b We promised Bill [PRO to kill each other]
- (17)a I (we) persuaded Bill [PRO to kill us] 'Object control'
 X Z Y <RI between I or we and us is blocked by PRO,
 s since PRO is **not** controlled by I It is a Z.>
 b *I (we) promised Bill [PRO to kill us] 'Subject control'
 X ~~Z~~ Y <RI between I or we and us is **not** blocked by PRO,
 since PRO **is** controlled by I or we. It is not a Z.>
- (18) Z is a **specified subject** with respect to X if it is not 'controlled' by (a category containing) X. (If Z is lexically specified, it is not controlled at all. PRO is controlled in the standard sense. Trace is controlled by its antecedent. <Not the standard usage of the term, which is why I put it in quotation marks.>)

- (19)a *They appealed to John [PRO to like each other] An object control construction>
 X Z Y <PRO not controlled by *They*>
 b They appealed to John [t to like each other] Raising
 X ~~Z~~ Y <trace 'controlled' by *They*>

- (20)a We appealed to John [PRO to like us] <PRO blocks RI, allowing coreference
 RI X Z Y between *We* and *us*.>
 b *We appealed to John [t to like us] <trace 'controlled' by *We*, so not a Z. RI is
 X ~~Z~~ Y **not** blocked, so coref. is out>

A POSSIBLE SIMPLER ALTERNATIVE:

- (21) Specified Subject Condition (Simplified version) <ALL subjects block relations across
 No rule can involve X, Y in the structure them>
 ...X...[_{α} ...Z...-WYV...]...
 where Z is the **specified** subject of WYV in α

AND NOW make the previously 'non-specified' Zs (what I sometimes call transparent subjects) be Xs. Then there never need to be relations **across** a subject. This removes an unstated stipulation: While PRO and trace can be Zs and Ys in relations, they can't be X's.

AND there is independent evidence that traces must, in fact, sometimes be Xs:

- (22) Which men does Mary think [Susan said [t will like each other]]
 X Y **NO WAY!**
 X Y ✓

B. "Conditions on Rules of Grammar" (1976)

- (23)a The men like each other
 b *The men want [John to like each other]
- (24) Reciprocal interpretation assigns an appropriate sense to sentences of the form NP...each other (and is constrained by the conditions). <No *each*-movement. *each other* is base-generated>
- (25)a The men like them
 b The men want [John to like them]
- (26) Disjoint reference (DR) assigns disjoint reference to a pair (NP, pronoun) (and is constrained by the conditions).
- (27)a John seems [t to like Bill]
 b *John seems [Bill to like t]
- (28) "...the relation between NP and the trace that it controls [is] a special case of bound anaphora..."
 That relation is constrained by the conditions. **The conditions thus are all conditions on surface structure applying specifically to rules of anaphora.** Anaphors are reflexives, *each other*, and trace of A-movement. (I have illustrated SSC. TSC is the same.) For examples like (16)-(17) above, Chomsky continued to assume the (1973) definition of 'specified subject'. In retrospect, it is clear that such a complication was unnecessary. Rather, instead of the transparent subjects **not** counting as Z, they **would** count as X as shown in (21).
- (29) The rules of anaphora relate surface structures (enriched to include traces) to LF. Perhaps more generally, surface structure determines LF.

- (30) John thought that Bill liked him <with *John* and *him* coreferential> (cf. (5) above)
- (31) (30) is now not a problem, as it does not involve a rule of sentence grammar at all, Chomsky says, following Lasnik (1976) "Remarks on Coreference". [The problem, of course, is "He thought that Bill liked John", where coreference is blocked even though *John* is inside a finite clause that *He* is outside of, and they are separated by a subject, *Bill*. The problem comes home to roost immediately below.]
- (32)a Who_i t_i said [Mary kissed him_i]
 b Who_i t_i said [he_i kissed Mary]
 c *Who did he say [Mary kissed t] (Wasow's 'Strong Crossover', analogizing (32c) to (33c))
 d *Who_i did he_i say [t_i kissed Mary]
- (33)a John_i said [Mary kissed him_i]
 b John_i said [he_i kissed Mary]
 c *He_i said [Mary kissed John_i]
 d *He_i said [John_i kissed Mary]
- (34) for which person x, he said Mary kissed x
- (35) Taking a variable to function as a name, (32)c then reduces to (33)c. <So now, we also know why TSC and SSC do not constrain wh-movement. wh-trace, unlike passive or raising trace, is not an **anaphor**; it's a **name**.

C. "On Binding" (1980) <<<We won't get into (36-38), nor the complicated new notation Chomsky gave.>>>

- <<<(36) (Certain cases of) SSC and TSC are reformulated as the Opacity Condition:
 If α is an anaphor in the domain of the tense or the subject of β , then α cannot be free in β , $\beta =$ NP or S'.
 The conditions are now strictly on anaphors themselves, not on rules, and "Tense and Subject are 'operators' that make certain domains opaque." [But what about RI? There was VERY fancy notation for doing the work of RI.]
- (37) Which men did Tom think Bill believed [t saw each other]
- (38) In the earlier theories, each other was assumed to take Which men as its antecedent. Aside from the semantic impropriety of that, the conditions would have blocked it. Now each other is coindexed with t and it is not free in any opaque context. [Question: Could an analogous move have been made before? Answer: Yes!]>>>
- (39) *They told me [what I gave each other]
- (40) Chomsky says that (39) displays a certain 'redundancy' in the Opacity Condition: it is excluded by both the SSC part and the TSC part. Hence, Chomsky re-broke it apart into two separate conditions, (41)a,b, with (41)b, the Nominative Island Condition, a narrower version of TSC involving only subjects of finite clauses. (41)a, as before, involves only non-subjects, in the simple examples at least.

(41)a If α is in the domain of the subject of β , then α cannot be free in β .

b A nominative anaphor cannot be free in S' containing S.

<<'Free' means not bound. 'Bound' means c-commanded by and coindexed with some item. 'In the domain of' means c-commanded by.>>

(42) They expected [that [[pictures of each other] would be on sale]]

<(42) violated TSC, evidently incorrectly, but does not violate NIC; each other is not in a **nominative** position.>

2 points about these 2 interesting arguments: Theories of anaphora were always based on the evident complementarity between anaphors and bound pronouns. Where one is good, the other is bad. But sometimes the complementarity breaks down. In configurations like (42), both are possible:

(43) They_i expected [that [[pictures of them_i] would be on sale]]

By letting in (42) this way, (the descendant of) RI rules out (43). The LGB binding theory has the same difficulty.

Next, the 'redundancy' Chomsky worried about is not entirely eliminated:

(44) *They told me [what Mary said [each other gave me]]

As in (39), *each other* is in the domain of a subject (*Mary* this time), so the SSC part is relevant. But *each other* is subject of a finite clause, a nominative position, so NIC is also relevant.

D. Lectures on Government and Binding (1981)

<Chomsky begins his discussion of BT in LGB by criticizing the immediate predecessor in "On Binding". I didn't discuss the details of OB, but the specific criticisms should make the relevant concepts clear enough.>

(45) *Who do you think [that [t left]]

(46) In (45), t is not 'properly governed' (ECP). <More on this in the next course module> <OB tried to treat wh-trace as an anaphor to rule out (45). LGB points out the error of that approach.> The theory of anaphora is no longer at issue. We are thus free to treat the trace of wh-movement as a **name** (rather than an **anaphor**) in accord with the treatment of strong crossover.

(47) The OB system treated PRO as an anaphor, and only an anaphor. This was too weak in that it didn't entail that PRO occurs only in ungoverned positions, and too strong in that 'long distance control' as in (48) would be incorrectly excluded by SSC.

(48) They thought I said [that [[PRO to feed each other] would be difficult]]

(49) In OB there is a sort of redundancy between the theories of Case and binding. They both pick out the subject of infinitives as special, but by totally different means.

(50) In OB the two configurations relevant to binding theory - subject of a finite clause [residue of TSC] and c-command domain of a subject [SSC] - are in no way related.

(51) The OB indexing conventions are complicated. [I didn't give them in this HO, but they are indeed complicated.]

The LGB Binding Theory: [All of this demands that every referential type NP has an index, and that a trace has the same index as the NP it is a trace of.]

(52) α is bound by β if and only if α and β are coindexed and β c-commands α .

[X c-commands Y iff every Z dominating X also dominates Y]

(53)a α is A-bound by β iff β binds α and β is in an A-position

b α is (A-)free if and only if it is not (A-)bound.

(54)A An anaphor is (A-)bound in its GC. 'Condition A'

B A pronominal is (A-)free in its GC. 'Condition B' (A descendent of RI)

C An R-expression (fully lexical NP, or variable) is (A-) free. 'Condition C' (A descendent of the noncoreference rule of Lasnik (1976))

(55) α is a governing category for β if and only if α is the minimal category [i.e., XP] containing β , a governor of β , and a SUBJECT accessible to β .

(56)a SUBJECT = AGR in a finite clause; NP of S in an infinitival; NP of NP in an NP.

b 'Accessible' (preliminary version). X is accessible to Y only if X m-commands Y.

c X m-commands Y iff every ZP dominating X also dominates Y

(57)a *John_i believes [(that) himself_i is clever] <These all violate Cond. A. The GC is the

b *They_i believe [(that) each other_i are clever] embedded clause.>

c *Mary_i is believed [(that) t_i is clever]

<The GC of the anaphors in the following is the matrix clause, as there is no governor of the anaphors in the embedded clause.>

(58)a John_i believes [himself_i to be clever] (cf. J. believes himself)

b They_i believe [each other_i to be clever] (cf. They believe each other)

c Mary_i is believed [t_i to be clever] (cf. M. is believed t)

(59) *John_i believes [him_i to be clever] (cf. *J_i believes him_i) Cond. B

(60) John_i believes [(that) he_i is clever] <No Cond. B violation>

(61)a Mary believes [them to be clever]

b *Mary believes [(that) them are clever]

(62) Infinitivals (at least some of them) are not barriers to government, either for Case assignment (61) or for establishment of governing category (57)-(60); cf. (49) above.

(63) *John_i believes [Mary to like himself_i] <The GC is the embedded clause, with *like* the

(64) John_i believes [Mary to like him_i] governor and *Mary* the accessible SUBJECT.>

(65) *He_i believes [(that) John_i is clever] <These are out by Cond. C. GC is not relevant for

(66) *He_i believes [Mary to like John_i] Cond. C.>

- (67)a *We_i heard [their stories about each other_i] <Out by Cond. A. The GC is the object NP. *about* is the governor, and there is the accessible SUBJECT.>
 b We_i heard [some stories about each other_i] <No accessible SUBJECT in the object NP>

(68) Does (56) successfully address (50)? Or are TSC and SSC simply hidden? Chomsky suggests that (50) **is** addressed, in that SUBJECT of α is the most prominent nominal element of α , taking INFL (which contains AGR) as the head of S. <**But** notice it cannot be the head of NP that counts as SUBJECT of NP, or (67)b will be ruled out alongside (67)a.>

<<<The next section is about more details of ‘accessible SUBJECT’s. Depending on time, we might skip it.>>>

- (69) They₁ expected [that[[pictures of each other₁] would be on sale]]
 (70) They₁ expected [that[[pictures of each other₁]₂ AGR₂ would be on sale]]
 (71) μ is accessible to β iff β is in the m-command domain of μ (and assignment to β of the index of μ would not violate (72).)
 [X m-commands Y iff every ZP dominating X also dominates Y]
 (72) * $[\gamma \dots \delta \dots]$, where γ and δ bear the same index.
 (73) i.e., μ is accessible to β iff β is in the c-command [m-command] domain of μ (and μ is not coindexed with any category properly containing β).
 ((LGB takes agreement to be instantiated via co-indexation, so a subject and AGR are coindexed.))
 (74) *John₁ thinks [that [himself₁ AGR₁ will win]]
 (75) They₁ think [it₂ AGR₂ is a pity [that pictures of each other₁ are hanging on the wall]₂]
 (76) *They₁ think [it₂ AGR₂ bothered each other₁ [that S]]
 (77) They₁ think [it₂ AGR₂ is a pity [that pictures of them₁ are hanging on the wall]₂]
 (78) They₁ expected [that[[pictures of them₁]₂ AGR₂ would be on sale]]
 (79) Problematically, anaphors and bound pronouns are not in full complementary distribution. (80) is a further illustration.
 (80)a They read [each other's books]
 b They read [their books] <<<End of section on accessible SUBJECT>>>

The distribution of PRO -the “PRO Theorem”

(81) John tried [PRO to leave]

(82) a *I like PRO

b *Susan spoke to PRO

c *John believes [PRO to be intelligent]

d *John's belief [PRO to be intelligent]

e *John believes [PRO is intelligent]

(83) Proposal: PRO is a pronominal anaphor [see (84)], hence, it must obey both Condition A (54)A and Condition B (54)B. That is, it must be both bound and free in its governing category. If it has a governing category, this is a contradiction. Any sentence where a PRO has a GC will necessarily violate either (54)A or (54)B. Therefore PRO must have no GC. This (almost) entails that it must be ungoverned, the descriptive generalization covering (82). This deduction is standardly called the PRO theorem. It also has the effect of permitting long distance control, by virtue of divorcing control from binding theory. See (47).

(84) "...PRO is like overt pronouns in that it never has an antecedent within its clause or NP. PRO also resembles anaphors in that it has no intrinsic referential content but is either assigned reference by an antecedent or is indefinite in interpretation, lacking specific reference."

(85) *[Pictures of each other₁]₂ AGR₂ are on sale

(86) *[Pictures of PRO₁]₂ AGR₂ are on sale

(87) Addendum to (55): A root sentence is the GC for a governed element that otherwise has no GC.

(88) **The Binding Conditions (54)A-C are purely syntactic:** They filter out structures based solely on their formal properties. Do we need any associated semantics? (Recall that Chomsky (1973) and Lasnik (1976) had semantic interpretive rules of (non-)coreference, or non-overlap in reference.) Chomsky, in the 1979 draft of his 1981 “Markedness and Core Grammar”, indicates that we no longer need such semantic rules, saying “... we.. need introduce no mention of the property of disjoint reference”. Well, let's see:

(89) "John likes him" can't mean that John likes himself.

(90) *John₁ likes him₁ Out by Cond. B

(91) John₁ likes him₂ In by Cond. B. What prevents it from meaning John likes himself?

NOTHING so far. Indices are just pieces of notation. They have no meaning apart from what the theory provides for them. The quote from Chomsky above suggests that the theory need not provide them any meaning. That can't be right.

WE NEED SOME SEMANTIC RULES. A first attempt:

(92) If two NP's have distinct indices then...

(93) They like him John likes him

(94) ...then they are disjoint in reference. Call this Semantic Rule B

- (95) *We_i like myself₂
 (96) →We_i like myself₁
 (97) If two NPS have identical indices, then they are coreferential. Call this Semantic Rule A
 (98) We_i think [I_i will win] ? can't be 1, since *We* and *I* can't corefer. And it can't be anything other than 1, since *We* and *I* can't be disjoint in reference.
 (99) The problem is that NP's have at least three referential relations: disjointness; identity; overlap. But two numerical subscripts are either identical or distinct. See (51). (98) is the cost of addressing (51) in the way that LGB does. Notice that Chomsky (1973) and Lasnik (1976) had no such problem, **because they didn't use indices at all.**

Where in the derivation do binding conditions apply?

- (100) [Which book that John_i read did he_i like t_j
 (101) *He_i liked [every book that John_i read] with LF, via QR, [every book that John_i read][He liked t_j
 (102) *Who [t_j thinks that he_i read [which book that John_i likes] <LF wh-movement moves the direct object to join *Who* in Spec of matrix C, thus taking *John* out of the c-command domain of *he*.>
 (103) "...these examples provide prima facie evidence that the binding theory applies at S-structure, a conclusion that I will now adopt."

E. Some Concepts and Consequences of the Theory of Government and Binding (1982)

- (105) Explicating the LGB theory, Chomsky proposed a typology of nominal expressions with respect to BT, based on 2 binary features:
 (106) [+anaphor], [+pronominal]
 (107)a Anything with the feature [+anaphor] must obey Condition A.
 b Anything with the feature [+pronominal] must obey Condition B.
 c So something that is [+anaphor] and [+pronominal] must obey both conditions: The PRO Theorem, as in LGB.
 d Something that is [+anaphor] and [+pronominal] need not obey either condition, but must, instead, obey Condition C.

2 binary features yield 4 combinations. When we cross this with 'lexical' and 'non-lexical', we get 8 combinations. Chomsky claims that 7 of the 8 exist, and the 8th can't exist.

(108)	<u>non-lexical</u>	<u>lexical</u>
[+a,-p]	A-trace	reflexives, reciprocals
[+a,+p]	PRO	<can't exist; must be governed to be Case marked, but then would violate Condition A or Condition B>
[-a,+p]	pro (the silent subject in null subject languages)	overt pronouns
[-a,-p]	\bar{A} -trace	NPs that aren't pronouns or anaphors

