

Li9 Syntax: Notes

Ashna Ahmad

April 30, 2022

Contents

1	Introduction	2
1.1	Levels of adequacy in syntax	3
1.2	X' revision	3
2	Hierarchy and binding	5
2.1	Relational grammar vs phrase structure	5
2.2	Binding theory	8
2.3	Passives	10
3	Head movement, c-command and locality	12
3.1	A-movement and intervention	17
4	Clause-internal structure	18
4.1	Splitting VP	19
4.2	VP-shells	20
5	Nominal structure	21
5.1	Defining 'Case'	23
5.2	Motivating A-movement	24
6	Agreement and locality	25
6.1	Unaccusatives and unergatives	25
6.2	Formal features	27
6.3	Agree	29
7	Hierarchy and linear order	31
7.1	Antisymmetry in syntax	31
8	Bare phrase structure	33
9	Wh-movement	37
9.1	Island constraints	38
9.2	Subjacency	38

9.3	Phase theory	40
9.4	Barriers	41
10	Covert wh-movement	41
10.1	Quantifiers	42
10.2	Copy theory	43
11	Phases	44
11.1	Phases and A'-movement	45
11.2	Relativised Minimality	46
12	Silent syntax	46
12.1	Ellipsis	46
12.2	Indexation	48
12.3	Null pronouns	49
13	Parameters	50

1 Introduction

- Adult L1 speakers have a remarkable amount of rich, intricate knowledge of which they are largely not conscious (e.g. grammaticality, ambiguity, interpretation etc judgements). These abilities can seem obvious at first sight but are in fact more intricate.
 - Words have to go in some type of order; most permutations will not be grammatical
 - Grammaticality: the difference between 'it is unlikely to be' and 'it is improbable to be'
 - Constituency and structure dependency: the difference between phrasal verbs and directions ('looked up' has two meanings)
- All these facts of interpretation and ambiguity are easy for L1 speakers to apprehend: largely unconscious grammatical knowledge
- The above is known as I-language: internal language
- Competence is an abstract property and is necessary but not sufficient for a speech situation to function (other things are required such as short-term memory and wakefulness). Performance is actually actualising competence.
- E-language: external or extensional language, which is socially and culturally mediated
- What is the nature and constraints of crosslinguistic variation given that some aspect of the system/competence is inherited?

1.1 Levels of adequacy in syntax

- Observational accuracy is when a sentence is observed and accepted as well-formed. A grammar contains only core constructions verified by observational adequacy
- To be descriptively accurate, we need generalisations/motivatable structural descriptions (e.g. X is the logical subject, Y is the direct object; indexation and coindexation in specifiers). In other words, "structural descriptions which capture a difference in meaning."
- The ultimate goal is not only descriptions, but explanatory accuracy - we want to know the nature of these descriptions and what exactly the symbols and categories explain, in order to integrate observations within a broader understanding of cognition.
- Prior to the instigation of phrase structure grammar, there was not much more than observational accuracy - no good formal way to describe sentences.
- The principles and parameters approach was an attempt to move from descriptive to explanatory adequacy. On paper, the account of principles and parameters as values set during language acquisition) was a substantial improvement on the idea that children acquiring language just had to somehow acquire phrase structure rules.
- Minimalist program: the notions we use in our explanations (principles plus parameters) should be reducible in some sense to very minimal, optimised specifications.
- Strong minimal thesis: everything in syntax is driven by 2 constraints: optimality (parsimony) and legibility (this means being legible to both the phonological and semantic levels of representation, which read the syntactic input in different respective ways). An optimal solution gets us the simplest conditions for the sound-to-meaning interfacing process.
- SMT can be seen essentially as 'keep the syntax as simple as possible while ensuring it provides the information needed by the phonological and semantic components'

1.2 X' revision

- Assumptions are category-neutrality and endocentricity: all syntactic categories have the same basic internal structure, and for every XP there must be a head (X) of the same category which determines the nature of the phrasal constituent (XP is the maximal projection)
- Complements are sisters to heads; sisterhood is potentially the closest possible structural relationship. The complement is determined by the head's categorial selection properties (sublexical category)

- The specifier is the sister of the X', i.e. the [head+complement] constituent, where the latter is X'.
- X' is the intermediate projection of the maximal projection (XP) and minimal projection (head)
- Specifiers are not directly determined by lexical properties of heads
- This X' template is always applied (simple NPs still have all levels; the noun still projects upwards even without a complement or specifier, i.e. unary branching)
- Adjuncts are non C-selective categories which are *optionally* connected to the head. Adjuncts carefully modify phrases with specific, optional information. e.g temporal adverbs and nominal adjectives (separate APs which can modify a noun)
- Adjuncts, unlike specifiers, are stackable. If we take possessors and articles to be specifiers of NP, the ungrammaticality of some sentences can be accounted for by the idea that only one specifier can occupy the position (maybe this constraint is a parameter, i.e. language-specific) (in contrast to arguments; adjectives can be unlimited, for example)
- Adjuncts can normally be adjoined at different parts of the structure, often N' in the case of adjectives
- X is a variable over categories (can be any category)
- Complements are always closest to the heads that select them
- Constituency tests allow for the isolation of particular constituents:
 - Clefting:
 - * The Party Chairman sent a present to John.
 - * It was to John that the Party Chairman sent a present t. (the PP 'to John' is clefted)
 - * It was the Party Chairman that t sent a present to John. (the NP 'the Party Chairman' is clefted)
 - * It was a present that the Party Chairman sent t to John. (the NP 'the present' is clefted)
 - * It was the Party Chairman sent that t present to John. (ungrammatical because 'the Party Chairman sent' does not form one single constituent)
 - * It was a present to that the Party Chairman sent t John. (ungrammatical because 'a present to' does not form one single constituent)

- Ellipsis: John ate the cake and Mary did [eat the cake] too. ('eat the cake' can be elided because it forms a single VP constituent)
- Pro-forms: [The man who wears glasses] hopes that he will win. ('the man who wears glasses' can be replaced by the pro-form 'he' because it forms a single NP constituent; the word 'man' alone cannot be replaced by 'he' because 'man' alone is not a full constituent). The pro-form to identify VP constituents is 'do so'; 'the man hopes that he will win, and the woman does so too' (isolates the VP starting with 'hopes...')
- Wh-questions/fronting: in both cases, the constituent to be isolated is moved to the front of the sentence. 'Mary hopes that John will like her friends' → 'Which friends does Mary hope that John will like t?' or 'Her friends, Mary hopes that John will like.'
- X' accounts for three things - hierarchy, projection, and linear order - but is category-neutral (there is a way of generalising rules to X rather than accounting separately for V, N and other categorial specifications as in PS-rules)
- Parametrised X' theories cannot be rewritten as PS rules because they do not inherently specify order; in other words the X' constituent decomposes as a set (with order-independent elements) but does still hierarchically dominate the daughter nodes which project up to it, so hierarchy and projection seem to go together (as per the endocentricity assumption).

2 Hierarchy and binding

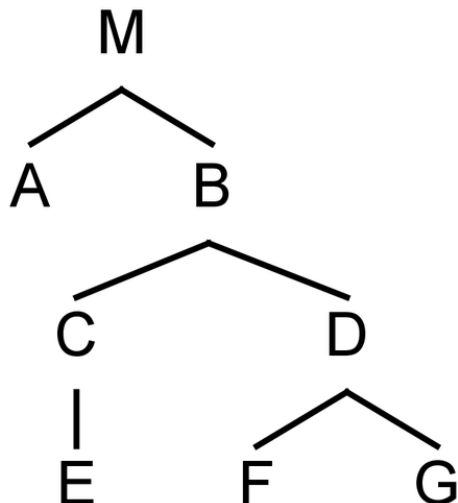
2.1 Relational grammar vs phrase structure

- Central questions: do grammatical relations take primacy, or does phrase structure?
- Based on a phrase structure NP...V NP, we can define grammatical notions of subject, object and so on ('the object NP is dominated by the VP'). This is known as a configurational approach. But can we also define phrase structure based on relational grammar, or does the former take primacy?
- It is canonical to define grammatical relations on the basis of phrase structure as follows: only subjects are nominative, agree with the verb, can be controlled ('convince' and similar verbs) in non finite clauses, are usually the antecedents of a reflexive, and generate the syntactic properties of a passivised object. These rules defining the subject can all be redefined in PS terms as conditions applying to [NP, S] or passivisation involving [NP, VP] to [NP, S]. As a result Chomsky gave primacy to phrase structure as providing information about grammatical relations.

- The **relational grammar critique** resulted from the earliest attempts to generalise this relation beyond English, and the observation that VP-fronting, VP-ellipsis and so on are not cogent concepts in other languages.
 - There is no VP pro-form in Japanese (SOV is ungrammatical) and VP-fronting is simply ungrammatical.
 - In French, “Kai a mangé du pizza et Lina aussi” is grammatically correct, but is I'-ellipsis as opposed to VP-ellipsis; the lack of a verbal pro-form like ‘do’ in many languages, including French, makes VP-ellipsis more ambiguous. Fronting the VP, moreover (et mangé du pizza elle a ??), would be unacceptable.
 - Meanwhile, many European languages which have more complex inflectional morphology than English do not seem to have configurationally distinct subject and object NPs; in Croatian and Russian, fronting would make no sense because almost any element of the sentence can be fronted. Ellipsis in Croatian is also ungrammatical: ‘Kai je jeo pizzu i Lina [takodjer]??’ makes no sense.
- Despite this lack of configurationality, passivisation in Japanese/French are very similar to English - word order constraints do not apply, but the notion of a direct object becoming subject-like in the passive absolutely does
- Advocates of relational grammar also claim that their representational schema is applicable to more languages than phrase structure. According to the phrase structure account, passivisation is a movement operation wherein an NP moves out of the VP which would have dominated it in the active. However, in some languages such as Cebuano, passivisation is better described as an affixation operation than as movement (cf. Perlmutter and Postal 1977). This means that the phrase structure description of passivisation works well for English, but not so well for other languages which do not incorporate movement into their passive constructions. On the other hand, an account of passivisation based on grammatical relations would simply describe the passive as the displacement of the subject to yield an intransitive clause. Arguably, this is more widely applicable than the phrase structure account.
- On the other hand, constituency doesn’t work that well all the time in English. Binding theory is a more general way of finding structural relations.
- It is quite possible that VP-fronting, VP-ellipsis and so on are Anglocentric, but not the constituents they seek to describe. Baker (2001) backs up this argument by pointing out that some constituents even in English cannot be identified by constituency tests: the NP complement of a determiner is a constituent, but cannot be clefted (“It’s [picture of John] that Mary bought that...” is the example provided.)

- Relational prominence refers to the prominence of subjects over objects, and direct objects over indirect objects; embedding prominence, meanwhile, is the prominence of elements of a main clause over those of a clausal complement (embedded within the main clause). Relational grammar sets up a clear distinction between these two types, wherein a process like passivisation which raises an object to a subject position is termed ‘advancement’ and the raising of a subject of an embedded clause to the matrix clause is termed ‘ascension.’ In a phrase structure account, no such distinction exists: both types of prominence are c-command relations, sharing the same “conceptual core” in Baker’s words.
- Baker suggests that treating both prominence types as the same type of relation allows for more successful predictions across languages. In Mohawk, matrix subjects are prominent over direct objects and embedded subjects, but not the subject of the same clause. If the two types of prominence are independent, there is no natural answer as to which takes primacy when the two types of prominence conflict with each other; it is impossible to predict whether the matrix object will have prominence over the embedded subject or vice versa. However, using the phrase structure explanation, we know **the matrix object has prominence because the first phrase that contains the matrix object – the VP – also contains a phrase that properly contains the embedded subject (the CP complement) and therefore the matrix object c-commands the embedded subject.**
- In short, prominence can be reduced in a very simple and general way to c-command and phrase structure rules reflect this intuition while relational grammar does not.

2.2 Binding theory



- C-command holds over the sister node and all the nodes it dominates:
 - M does not c-command any node because it dominates all other nodes.
 - A c-commands B, C, D, E, F, and G.
 - B c-commands A.
 - C c-commands D, F, and G.
 - D c-commands C and E.
 - E does not c-command any node because it does not have a sister node or any daughter nodes.
 - F c-commands G.
 - G c-commands F.
- X binds Y iff X c-commands Y and X is co-indexed with Y.
- The binding domain of A (where A must be an anaphor or a pronoun) is the smallest maximal projection containing both A and [a subject other than A or finite T.]
- A pronoun must be bound in its binding domain; an anaphor must be free in its binding domain; an R-expression (name or definite noun) must be free in any case.
- Examples:
 - 'John hates him': 'him' and 'John' cannot coindex because him is a pronoun, the binding domain is the VP containing 'him' and the

finite tense 'hates', and 'him' is dominated by the VP which is a sister node of 'John.' If they are coindexed, then the two conditions of binding - coindexing and c-command - would be fulfilled, and this cannot hold because the pronoun needs to be free in the binding domain.

- 'John believes him to be the best': 'him' cannot coindex because the binding domain is the whole sentence ('to be' is not a finite T) and John c-commands him. This contrasts with 'John believes he is the best' because finite T ('is') means that the CP 'that he is the best' is the binding domain; John in this case is outside the binding domain, so the pronoun is free even if it coindexes. This applies to any indirect statement.
- 'I asked John about him' - 'him' cannot coindex because the binding domain is the VP (finite T) which contains 'John', 'John' c-commands everything in the PP, and the pronoun must be free in its binding domain
- 'He left' - 'himself' doesn't work because it would have had to be bound
- 'Why do those people think to be smart?' (vs. 'Why do those people think they are smart?') - 'seem to be smart' is correct because seem is a raising predicate, however 'thinks' is not a raising predicate ('It thinks that those people are smart' would have to be the start sentence in order for raising to apply, and that is illegal). They is a pronoun which is free in its binding domain, i.e. the subclause introduced by 'that', so this is correct.
- More examples
 - 'He loves John' - John cannot coindex because it is an R-expression and they have to be free
 - 'His mother loves John' - 'John's mother loves him' is ok because the binding domain is the VP, which 'his' is not part of, so the pronoun is free in the binding domain; however, R-expressions need to be free full stop, regardless of whether their c-commander is in the binding domain.
 - 'He thinks that John is smart' vs 'John thinks that he is smart' - coindexing does not work in the first case because an R-expression always has to be free, but does work in the second case because the binding domain of 'he' is the CP; 'John' is therefore outside of it and therefore free
- Crosslinguistic exercises:
 - Principle C solutions are basically always the same; R-expressions

must always be free, pronouns within indirect statements are ok because of complementiser phrases

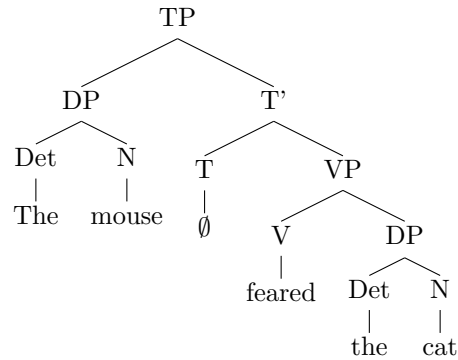
- Malayalam (see lecture notes) - "The king pinched his own wife" here 'his own' is an anaphor, the binding domain is the whole sentence, and it must be bound i.e. c-commanded and coindexed by the subject
- 'His own wife pinched the king' does not coindex because 'the king' is an R-expression
- Subject-object asymmetries in binding can be used as a constituency test (although it doesn't specify much info about the syntactic objects, neither do standard constituency tests). This is because the binding domain is necessarily a maximal projection, so if an anaphor/referent pair or pronoun/referent pair are in the same binding domain and have the expected coindexation properties, we can locate the constituent represented by the maximal projection which defines the binding domain.
- Other cases in which c-command reveals structural asymmetries:
 - Bound-variable interpretations, e.g. 'Every boy persuaded his mother that video games are good for you.' This is syntactically ambiguous, but reversing the c-command relation to yield the ungrammatical 'Her son persuaded every mother that...' (ungrammatical due to the bound R-expression) shows that 'his mother' doesn't corefer with the quantifier over the subject (rather with each individual instance of the subject).
 - Negative polarity items; the negative element must be in a c-commanding position ('no one persuaded anyone' is correct, but 'anyone persuaded no one' is not, because the latter places the negative element in a daughter node).
 - Superiority, e.g. 'Who persuaded whom that [...]?' is possible, but reversing the positions of the wh-elements is not, because the c-commanding wh-element is subject to wh-movement - unlike the other element. (This is essentially a subject-object asymmetry.). The non-c-commanding wh-element remains in the post-verbal subject position, rather than being moved higher.

2.3 Passives

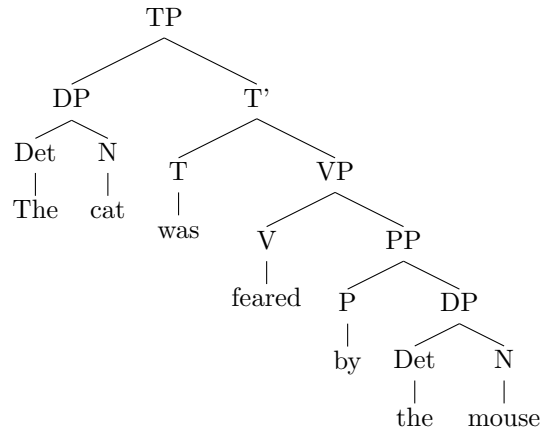
- The object appears to 'move to the subject position' in passives: it agrees with the verb, takes the nominative case feature, and can take antecedent position.
- *All* c-command asymmetries discussed in 2.2 also apply to passives, where the passivised noun takes the c-commanding position.

- So on the syntactic level, the passive takes the subject position, but on the semantic level, it bears θ -roles which could be considered typical of active objects (patient, cause, recipient/instrument [like an active indirect object], experiencer, location).
- The passive reaches this syntactic position via a movement operation, which leaves a trace in the complement of VP (which in the example below is also occupied by PP):

1.



2.



- In (2), the new position of 'the cat' c-commands the former position which is now occupied by the trace/PP. The new θ -role is determined by the start position of the movement operation. The new position determines all of the syntactic and binding/c-command-related properties (listed in section 2.2).
- The relationship between the passive and the trace it moved from is similar to the relationship between an anaphor and its antecedent, in the sense that it is an instance of movement where the new position (analogue of the antecedent) c-commands the old position (analogue of the anaphor) and must also be in the same binding domain as the old position (so no CPs/indirect statements may intervene). In other words, the passive/trace

pair (and the anaphor/antecedent pair) must be in the same clause. This is known as the locality property ('local movement').

- So does passivisation (or argument-movement in general) always have to take place within a binding domain, like the antecedent/anaphor relation? Not quite:
 - 'The students were believed t to have been arrested t' ' - the two traces here stand in for the two original positions which the passive word was moved from, i.e. the object position of each verb in the sentence (specTP of the main verb and specTP of the embedded verb). This is cyclic movement, because the passivised phrase moves through an intermediate specTP position (of the sub clause) to its final specTP position (of the main clause)
- Cyclic movement cannot create sentences like 'The students were believed the professors to have been arrested (the students)' because the specTP position of the subclause is already occupied here, so the intermediate position is blocked and therefore cyclicity cannot enable movement to the main clause's specTP.
- 'The students were believed that the professors were arrested' - **the reason why the only way to create a sentence with this sense is by using the infinitive is because a non-finite tense does not create a binding domain.** Think about it using the anaphor equivalence - an anaphor cannot be free in its binding domain, so neither can the trace of a passive, and therefore the start position of the trace is ungrammatical because the binding domain of the sentence-final trace is the CP and there is nothing within this domain for it to be bound to.
- Compare 'the students believed the professor to have saved themselves' - this is the same situation as above, but the illegally free element is the anaphor, because its binding domain begins with the subject 'the professor' (recall that a binding domain can be generated by an XP dominating either a finite tense *or* a subject).

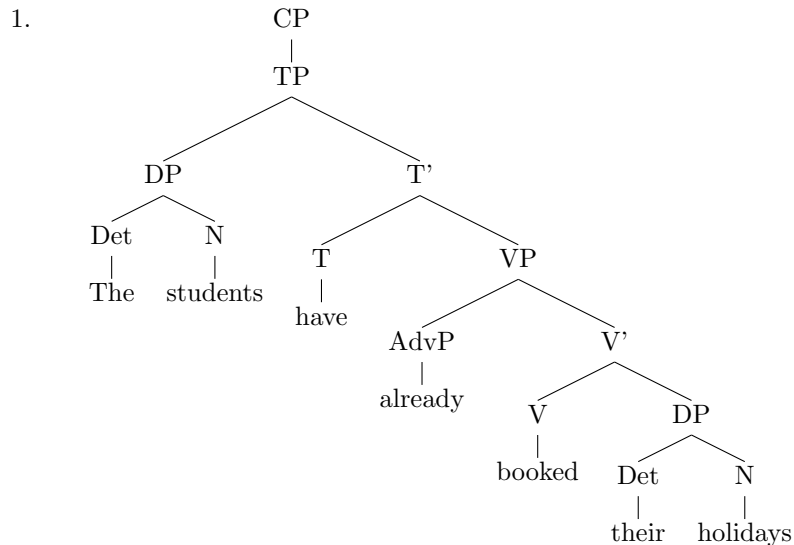
Conclusion Phrase structure is capable of defining the framework of apparent object-to-subject conversion, even though this appears to be a relational grammar concept, because binding and c-command constraints govern the movement of the passive (which is an instance of A-movement/argument movement).

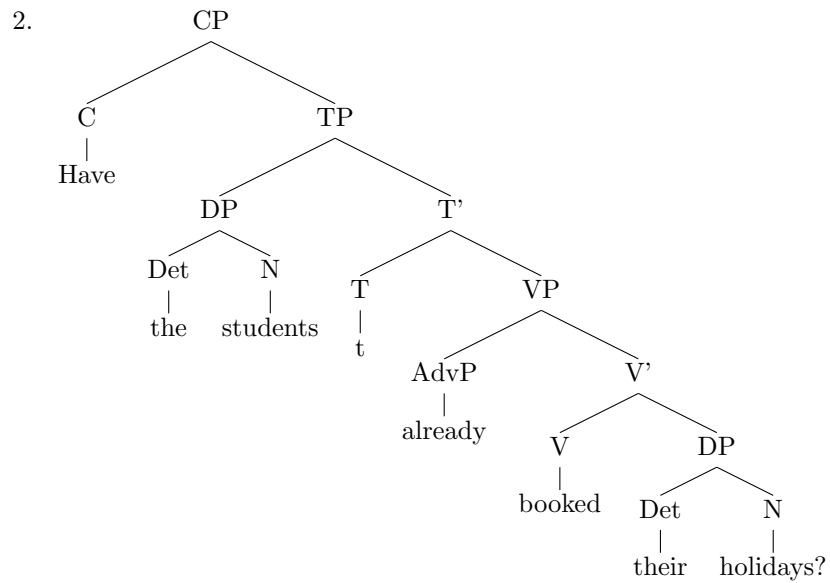
3 Head movement, c-command and locality

Locality of selection Every argument that α selects must be projected within the projection of α . If α selects β , then β depends on α . If α selects β , and if locality of selection is satisfied, then α and β are in a local dependency. If α selects β , and if locality of selection is not satisfied, then α and β are in a

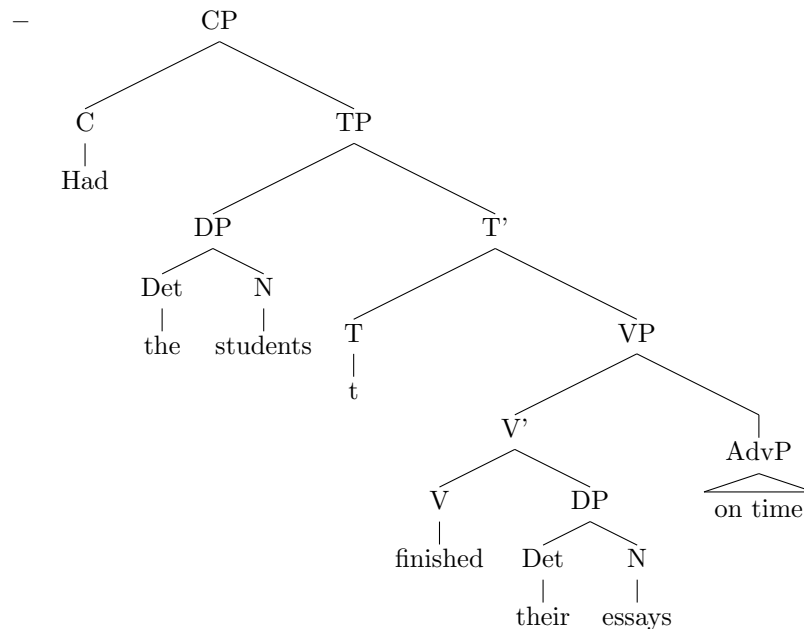
non-local dependency. The existence of a non-local dependency indicates that movement has occurred.

- Example:
 - In the sentence ‘John studies the report carefully’, the verb, ‘studies’, is the Head of the VP projection, the DPTHEME (the report) is projected onto the Complement position (as sister to the head V), and the DPAGENT (John) is projected onto the Specifier (as sister to V’). In this way, (1a) satisfies Locality of Selection as both arguments are projected within the projection of the head that introduces them. The adverb phrase, ‘carefully’, attaches as an unselected adjunct to VP; structurally this means that it is outside of the local projection of V as it is sister to and dominated by VP.
 - In contrast, in the sentence ‘*John studied carefully the report’, the introduction of the AdvP carefully as sister to the verb study violates Locality of Selection; this is because the lexical entry of the verb study does not select an AdvP, so the latter cannot be introduced in the local projection of the verb.
- Head movement is another operation which is local, cyclic and sensitive to c-command. It involves the movement of a head to a position c-commanding its old position (i.e. c-commanding the trace).
- Subject-auxiliary inversion, for instance, moves the auxiliary to specTP (or T) at the start of the sentence such that it precedes the sentence-initial DP, as follows:





- In (2), the trace *t* in Tense position signals the perfect tense/aspect, while the C position - normally considered the clause-typing position - signals the clause as interrogative.
- A similar type of inversion is conditional inversion, where the element in C labels the clause as a counterfactual conditional. For this particular use, T must be past tense (the tensed component of the element in C is still in the trace left in T after movement):



- We can think of the clause-typing function in C as attracting certain elements in T (i.e. the tensed elements which are left in the trace post-movement). There are a few different so-called 'abstract affixes' which can exist in C and induce corresponding overt realisations of tense in T; Q (interrogative) and Irr (conditional) are just two examples.
- In English subordinate clauses, Q and Irr are overtly realised; this is represented by indirect questions ('I wondered whether...') and regular conditionals ('if...had') respectively. Since these are free morphemes and not affixes, they block subject-auxiliary inversion (because inversion cannot occur out of a subclause), and this is why main clause syntax is preserved in the subclause of English indirect questions ('I wondered was he coming' is incorrect).
- Subject-auxiliary inversion and conditional inversion show that T-to-C movement is clearly possible. However, **V-to-C movement is not**, and this is why we require do-support in cases where the verb lacks an auxiliary but still needs an interrogative sense. In these situations, some tensed form of 'do' moves to C, leaving a trace in T to mark tense just as other auxiliaries do.
- Why exactly does T block the movement of main verbs in English?

Head Movement Constraint A head cannot 'skip' an intervening head position.

- V-to-C movement is therefore illegal because T is an intervening head

between C and V. Cyclic movement via T is not possible if the intermediate position is blocked (the passives section also contains illegal examples of cyclic movement).

Intervention In the configuration $X \dots Y \dots Z$, Y intervenes between X and Z iff Y and Z are in an asymmetric c-command relation and Y does not c-command X.

- For **intervention** to apply, Y has to be a daughter node of X. V and T are in an asymmetric c-command relation because T is at the same level as the VP, not V, so T c-commands V.
- V-to-C movement is illegal regardless of the properties of the word occupying the V position; if an auxiliary happens to be occupying V ('how tall can John be?', 'should the students have booked their holidays?') it still cannot be moved to C due to the **intervening head**, whereas **any word occupying T** (the modals 'can' and 'should' in this case), can move to C.
- Thus, the head movement constraint encodes the local nature of head movement, where locality is defined in terms of non-intervention.
- Why does 'to be' not do-support?
 - 'How tall is John?' satisfies the HMC because we can see that 'be' has moved through T before arriving at C, since it has acquired present tense morphology. This makes this sentence an instance of cyclic movement (V-to-T-to-C). This is possible because of the auxiliary properties of 'be?' ('be' is essentially using its **auxiliary** properties to stand in where the modal stood in the previous example).
- From these constraints, head movement - despite having little in common with passivisation in terms of grammatical construction - is cyclic and local. This hints that a general constraint, independent of relational grammatical categories, may be able to capture the patterns of A-movement.
- Crosslinguistic evidence:
 - Ils réservent déjà leurs vacances - why is V-to-T and V-to-C movement allowed here? It seems to be because Pres is an affix in French, but not in English. **Like the abstract Q and Irr affixes, the Pres feature does not block inversion** (it does not occupy the T node or constitute a subclause which the verb cannot move out of). This is why the verb can move to T (preceding and leaving a trace in VP, where the adverb is located) and **therefore cyclically move through T to C** (i.e. specTP) to the very start of the sentence in an interrogative.
 - Auxiliaries can also move in French ('ont-ils déjà ... ?') - but the verb cannot move from V to sentence-initial position (C) if the auxiliary is already in C.

- In V2 languages like German, we see XP-V-Subject order, which seems to be due to the presence of an abstract, declarative affix in C - i.e. German **marks declarative sentences** in a similar way to the clause-typing Q element which permits T-to-C auxiliary movement in English. This seems to be why e.g. **V-to-T-to-C** movement seems to be permitted in German questions ('schon buchen sie das Urlaub?')
- 'Dass' sends the verb to the end for the same reason as 'if/whether' in English: because it is a free morpheme and so blocks V-movement.
- V2 clauses require **cyclic movement of the verb via T**, but we never actually see it in T...so this analysis of German is unproven.

3.1 A-movement and intervention

- Recall that passives are a specific case of the broader generalisation known as A-movement (where an entire NP, as opposed to just a head, moves). With intervention in mind, we may be able to come up with a definitive answer as to whether moved arguments pattern like anaphors with respect to binding (i.e. must a trace always be bound in its binding domain?).
- In fact, **A-movement appears to have stricter locality conditions than anaphors? Intervening expletives (where 'it' comes under T) seem to be illegal:**
 - 'John and Mary believes that it appeared that each other were bankrupt'
 - 'John was believed that it seemed t to be strange' (the expletive in T makes the sentence completely wrong due to nonlocality)

A-Movement Constraint A-movement, e.g. in passivisation, cannot skip *any* intervening subject position (including an expletive subject).

- The AMC therefore explains why the intervening subject 'the professors' in 'The students were believed the professors to have been arrested' renders cyclic A-movement impossible and the sentence ungrammatical.
- Raising, which is introduced by raising predicates (it seems that, it appears that etc.), enables the subject of the subclause (i.e. the left branch of the TP in the subclause) to move to subject position in the main clause, creating sentences like 'John appears to...'. This is subject-to-subject movement, rather than object-to-subject movement like passivisation, but is still argument movement because it moves an entire NP to an asymmetrically c-commanding subject position in a cyclic, local manner.
- As an instance of A-movement, raising abides by the AMC and **does not allow any intervening subjects**: 'John seems to appear to be strange' (from 'It seems that John appears to be strange') is allowed, but 'John seems that it appears to be strange' does not. Case?

Conclusion In the configuration $X \dots Y \dots Z$, movement from Z to X (of heads by HMC or of subject/object positions by AMC) is illegal if Y and Z are in an asymmetric c-command relation and Y does not c-command X . Intervening heads block head movement and intervening arguments block A-movement.

4 Clause-internal structure

- We have seen from the Q and Irr morphemes that C contains clause-typing features which attract certain T-elements; fronted and wh-elements are also in SpecCP in main clauses. This is in addition to the standard complementiser element in subclause C.
- T, meanwhile, contains auxiliaries and inflectional features which attract certain V-elements (e.g. in the French passé composé - in English, all agreement features are in T, not V).
- V is just the lexical verb plus complements.
- So:
 - C interfaces with the next clause up (e.g. the main verb when C is a subclause complementiser) or extrasyntactic (semantic?) discourse (e.g. clause type)
 - T contains information about tense, mood and aspect
 - V determines thematic structure by assigning θ -roles to arguments.
- The asymmetric c-command relations between C and T, and T and V, encode semantic as well as syntactic dependencies; the thematic domain (V) takes cues from the inflectional domain (T) (e.g. when there is inflection in V as well), and the inflectional domain takes cues from the discourse domain (C) (e.g. a conditional can cue a tense). Since the three domains are so fluid and interdependent, it makes more sense to suggest that each clausal domain (discourse/inflectional/thematic) covers fields of categories (C/T/V). This could mean that thematic, inflectional etc. information is split over multiple iterations of the same category, or that it is expressed by both C and T, both T and V etc. within a sentence.
- E.g.
 - ‘Should the students have been booking their holidays?’ - which category is ‘been’ in given that it expresses tense/aspectual information and tense is already covered by ‘have’?
 - We can split the T domain into the different inflectional categories which it serves to express. ‘have’ and ‘been’ express perfect and progressive aspectual features respectively. The T-domain in English consists of at least T, Perf and Prog.

- Modals can be part of the T domain when finite (as they always are in English), so ‘should’ etc occupies T as a modal. Other languages have modal affixes.
- In non-finite English clauses, ‘to’ occupies the space within T which the modal would have occupied: ‘To have been writing their essays was a good thing.’ Sometimes modality, alongside the perfect and progressive features, is considered essential to the T-domain.

4.1 Splitting VP

- Recall that raising is triggered by a raising predicate such as ‘seem ...’ and raise a subject of a subordinate clause (e.g. the clause introduced by ‘it seems that’) to subject position in the main clause.
- Some raising predicates do not take finite complements (e.g. ‘[subject] tends to...’ is correct, but ‘it tends that [subject]...’ is not). However, we still want to treat such verbs as raising triggers; so movement from a subclause CP (as opposed to a TP) should not be a necessary condition. We define raising as any A-movement (i.e. cyclic, local, and theta-role preserving) from a subject position to a higher subject position. (why does T become nonfinite in the ‘seems’ raising predicate case? Because the tense feature is on the main verb, i.e. ‘seems’?) Ask about this
- Can we prove that auxiliaries are raising predicates?
 - Auxiliaries take a non-finite complement - in this case, a participle. E.g. ‘enjoying their lunch’ in ‘the students should have been enjoying their lunch.’
 - The θ -role of ‘the students’ is determined by the verb ‘enjoy’, of which they are the subject, and is **preserved** even when the subject is ‘raised’ by the ‘should have been’ construction.
 - Like ‘tend’, auxiliaries do not take ‘that’-CPs as complements (selectional restriction) - which is not an issue because we have broadened the definition of A-movement.
 - If VP is the thematic domain, we can suggest that the subject originates there, and is raised to the subject position in specTP by the auxiliary - whence the idea of a VP-internal subject.
 - Successive cyclicity: ‘John appears t to have been believed t to have left’
 - Locality (the expletive blocks, as in other cases of A-movement): ‘John seems that it has been t eating his lunch.’
 - Floating quantifiers: the quantifier can be in the distinct positions SpecVP (‘should have all [been]’), SpecPerfP (‘should all have [been]’) or SpecProgP (‘should have all [been ...ing]’). This shows that the

subject must move cyclically through these positions when triggered by the auxiliary.

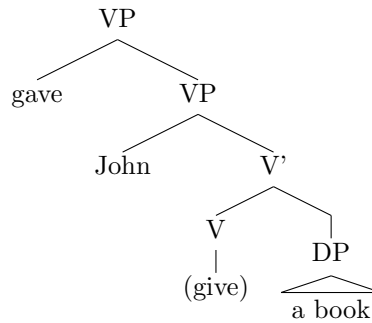
- Even non-overt auxiliaries can be raising triggers: ‘the students’ is still in SpecTP in ‘the students all t ate their lunch’ (it has been raised from the position immediately preceding the VP - ate - and preserves the agreement with the verb ‘ate’ as well as the agent semantic role).
- Not just auxiliaries, but any T including ‘to’ is capable of triggering raising (i.e. an instance of argument movement, specifically subject to subject argument movement).

Conclusion All subjects are generated, and assigned a θ -role, in specVP - i.e. within the thematic domain. The surface subject position, specTP, is always the final position of subjects in finite clauses.

- Accounting for VSO order:
 - If the object evacuates the VP before the VP moves into a higher position within the clause, it will derive VSO order instead of VOS. Massam looks at the difference between VP-raising accounts of VOS versus VSO by investigating the Niuean language. Massam argues that whether a NP object or a DP is selected by the verb will determine if there is VP or VP-remnant raising. In order to get the structure of a VOS clause, an NP object is selected by the verb; the selected NP does not need case so it will stay in the VP. Consequently, the object is pseudo-incorporated into the verb in the VOS clause. If a DP object is selected by the verb, VP-remnant movement occurs and creates VSO order.

4.2 VP-shells

- Ditransitive sentences contain a direct and an indirect object (without a prepositional phrase). In these sentences, the indirect object always asymmetrically c-commands the direct object. This means that the direct object must be lower in the tree structure than the indirect object, in order for the c-command to be asymmetric. We therefore get a structure where the indirect object is part of a (second) VP, while the direct object is in the noun phrase following V:



- Generalising this, we get the agent in the first VP, the goal in the second VP, and the theme in V. 'gave' moves from the lower V position (immediately before the direct object) to the higher V position (immediately before the indirect object). Note that the intermediate and minimal projections of the highest VP are omitted above.

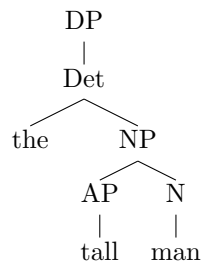
Uniformity of Theta-Assignment Hypothesis Specific θ -roles are associated with specific structural positions, whose hierarchy - e.g. the precedence of indirect over direct objects in ditransitives - is determined by c-command relations.

- Since we now have a notion of lower and higher VPs, we need to account for a cyclic transit of the lower verb through the higher VP by HMC. If we find a language where the lower VP does not move to the higher one, or where there is not a verb in each slot, it may be problematic.
 - Haitian Creole has a construction with the same order as the prepositional dative in English, but with no preposition - i.e. the reverse of the goal-theme order (the theme, or direct object, comes first).
 - This seems like an open problem whoops

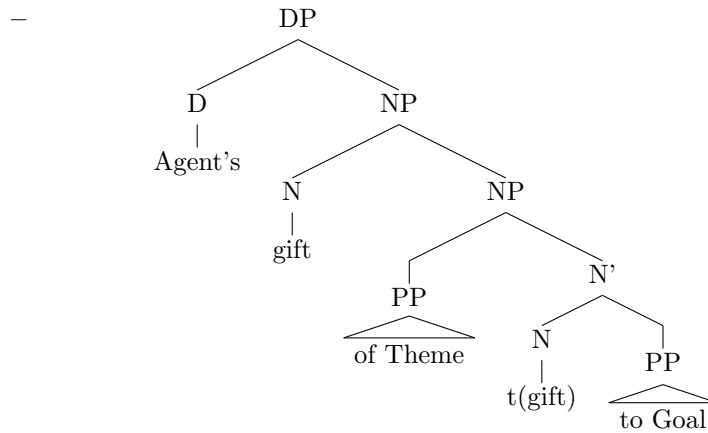
5 Nominal structure

- One-replacement (as in the quantity, not the pronoun) is the English constituency test to isolate nominals:
 - 'These [expensive and illegal bottles of wine] → [ones] are to be smuggled from Hungary.'
- The argument for considering D (rather than N) as the head is that:
 - N agrees with D (see demonstratives in English, and any articles in German [and most European languages])
 - D is not always in 'specNP' (West African languages)

- D is in complementary distribution, hence the ungrammaticality of multiple article, article and demonstrative etc. sentences
- D encodes the property of definiteness (which appears to be part of the discourse domain - see previous lecture)
- Quantifiers describe set relations (pretty self-evident). This suggests that simple nouns are sets, which may be evidenced by the ungrammaticality of singular count nouns (quantifiers and plural affixes both serve to turn set-denoting nouns (predicates) into argument-denoting nominals)
- Changing the semantic type of a nominal allows it to occupy an argument role - so a singular noun which has been affected by a semantic type change (via pluralisation or quantification) may occupy that argument slot in a DP.
- The idea is not that NPs do not exist at all, but they are adjoined to DPs, as follows:



- NB. adjectival modifiers go in specAP.
- Italian has a rule whereby the only valid DP configurations have AP following NP, or have an article preceding AP (so the order DP(Det, AP, NP) - hence the idea that D must be overtly filled in arguments (either by a determiner or by N, which moves to the head of the DP - preceding AP - when this is not the case). Some languages, such as English, allow a null D position (in English, this occurs in the case of proper names, mass nouns, and the indefinite plural).
- Within the DP, as we noted earlier, it seems that D is the clause-typing domain (analogous to C) since it encodes definiteness. Do there exist analogues to T and V within the DP?
- Recall the asymmetric c-command relations within ditransitives from week 4; similar asymmetric c-command relations in prepositional datives reveal similar relations within DPs (i.e. certain semantic roles within DPs are obliged to be in certain syntactic positions else the sentence is ungrammatical):
 - Specifically, the order seems to be Agent, Theme, Goal.



- NP-shells here seem analogous to VP-shells in that they are the thematic domain, where information about the θ -roles of the theme and the goal in the sentence originate. This can be extended beyond the ditransitive idea: once the determiner in D has specified definiteness, semantic elements (e.g. something like animacy and its link to the θ -role of agency) are encoded in N.
- Meanwhile, the intermediate position between N and D - analogous to the intermediate position between V and C - is represented by Num. Observe that numerical quantifiers must always be positioned between the article and the first AP in a sentence ('the two detailed studies' as opposed to 'the detailed two studies') and must also agree in gender and number with D and AP. Hence, the inflectional domain - T - of the nominal is NumP.
- However, nominals are obviously not perfectly analogous to clauses because without a verb, the obligatory features of the argument are different: nominative/accusative relations and verb agreement do not apply to nominals.

5.1 Defining 'Case'

- Case can be relationally defined as follows:
 - The subject of a finite clause is nominative and agrees with the verb.
 - The direct object of any clause is accusative and does not agree with the verb.
 - The subject of the infinitival complement of a verb like 'believe' ('we believe him to have...') is accusative and does not agree with the verb. Verbs of this category are known as Exceptional Case-Marking, or ECM, verbs.

- The above also applies to the complement of a transitive verb/preposition/prepositional complementiser (‘for, acc, to’ constructions are an example of a prepositional complementiser phrase).
- Syntactic relations are always marked by case. This can take the form of case-based inflection, a preposition, or no case marker at all.
- We can state case rules in a non-relational manner, as follows:
 1. the DP in the Specifier of a finite T is Nominative;
 2. the DP locally c-commanded by active v is Accusative;
 3. the DP in the Specifier of D is Genitive;
 4. the DP in the complement of an N is Genitive, marked by of;
 5. the DP locally c-commanded by a “Dative v/V” is Dative, marked by to.
- Subjects of infinitives (specTP where T is [-finite]) are not Case positions, hence the lack of overt DPs as infinitival subjects.
 - This is why ‘John seems to be nice’ is grammatical, but as soon as we insert another NP before ‘to’ (‘John seems Bill to be nice’) the sentence is completely ungrammatical.

The Case Filter DP, if phonologically overt, must be in a Case position.

5.2 Motivating A-movement

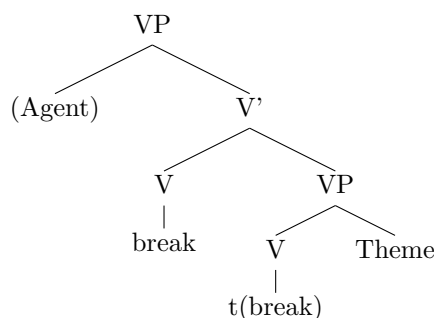
- In the sentence ‘John seems to be nice’, ‘John’ cannot be the subject of the infinitive because it is an overt, nominative-marked DP, so by A-movement it moves to specTP of the main clause. This position can take Case-marked NP because T is finite. Therefore movement has occurred by the Case Filter.
- For passives, we claim that the passive verb fails to license a Case- (accusative-) marked DP (by rule 2 above since V must be active) and so once again A-movement is motivated by the case filter.
- Subjects of ECM verbs also end up being in the subject position of an infinitive (‘John is believed (John) to be...’) and so must be raised in order for nominative to be licensed.
- Raising verbs like ‘seem’ are like passivised ECM verbs.

Conclusion A-movement is driven by the Case Filter: DPs move from Caseless positions to Case-licensed ones (locally, cyclically and obeying c-command).

6 Agreement and locality

6.1 Unaccusatives and unergatives

- Intransitive verbs only assign one θ -role, and so appear with just one DP (argument), which appears in subject position regardless of its θ -role. The subject DP in ‘John smiled’, ‘John arrived’ and ‘the vase broke’ has slightly different characteristics in each example.
- Alternating unaccusatives:
 - ‘The vase’ always has the theme/patient role, even when it is in subject position in the sentence ‘The vase broke.’ This may appear at first to contravene UTAH, because the patient is in a structural position which is not associated with its θ -role.
 - However, looking at other examples with this verb, we see that any sentence which appears to assign the theme/patient role to two words is ungrammatical: ‘The vase broke a nasty break’, ‘The vase broke its way to the museum’ and ‘The vase outbroke the glass’ are all ungrammatical because they attempt to assign multiple themes, whereas ‘John broke the vase’ is grammatical because ‘John’ is in agent position.

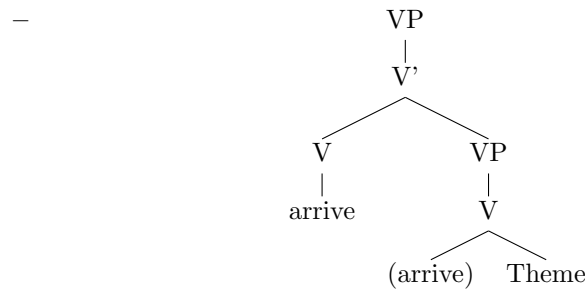


- Note that the tree above describes the property of alternating unaccusatives known as ‘transitive alternation’, wherein adding an agent means that movement of the theme to specTP is not required and so the sentence looks like a normal transitive sentence. Normal unaccusatives, as detailed below, lack this property.
- When the Agent isn’t present, the theme A-moves to specTP (via specVP - which is the intervening subject position - by AMC). This is because in the absence of an agent, the theme - the vase - is the nearest element with a case-marking, and so by the Case Filter the DP must take case in order to be grammatical. T’ is not shown in the tree above, but this is movement to specTP, so the agent is the sister of the VP node, in a mutual c-commanding relation, and both of these nodes are dominated by TP.

Burzio's Generalisation If no θ -role is assigned to the external argument (specVP), then there is no Accusative Case for the internal argument (complement of V).

- Unaccusatives:

- Unaccusatives like 'arrive' only have an internal argument, which would have been in the complement of V. By Burzio's, they cannot have an accusative case in direct object position because there is no agent role, so they must raise to specVP.
- The only argument of an unaccusative is a theme/patient - e.g. the theme/patient which 'undergoes the arrival motion.' This is why there is no agent in the sentence below (and alternating transitives are not possible).



- In 'there' sentences, 'there' assigns Case to the argument, but it has to be in a Case position to do so. 'There goes the train' is ok because it is in the subject DP position and assigns the theme role to 'train' (undergoing motion) in the complement of V, eliminating the requirement for raising. 'There to arrive a train would be a surprise' is ungrammatical because if 'there' immediately precedes an infinitive, it does not have Case (Case is assigned for finite T).
- So the conditions for the 'there' construction are:
 1. The theme/patient θ -role must be assigned by 'there' to the element in the complement of V. In other words, the verb must be an unaccusative.
 2. Burzio (1986) proposes the following list of verbs as unaccusative in English: arise, emerge, develop, ensue, begin, exist, occur, arrive, follow.
 3. 'There' must be in a Case position, i.e. precede finite T.
- Unergatives:
 - Unergatives, such as 'smile' have a single agentive external argument. They are unlike all types of unaccusatives in that *they have an agent*

- Therefore, they are a bit freer than unaccusatives. The following sentences are all grammatical:
 - * John smiled a cheerful smile.
 - * John smiled his way across the room.
 - * John outsmiled Mary.
- However, they must only have a single argument, so the transitive alternation would be completely ungrammatical here, as would ‘there’-constructions because these only assign theme roles.
- Unaccusatives are our final example of argument-movement. The three types of argument-movement are:
 - Passives (object position to subject position, with object-like θ -roles preserved)
 - Raising (subject position of a subclause to subject position of a main clause, triggered by a raising predicate or an auxiliary: ‘the students should have been t enjoying their lunch.’)
 - Unaccusative (object position to subject position, with an active verb that does not assign an agent role)
- All of these are local, cyclic, case-driven, and move the argument to a c-commanding subject position.

6.2 Formal features

- Syntactic operations generate and manipulate phrase structure, which then gets ‘interpreted’ at the sound (PF) and meaning (LF) interfaces - hence the parallels between phonological features (e.g. [+son]), semantic features ([+anim]) and syntactic (aka formal) features ([+nom]). Each entry in the lexicon, or lexical item, is a bundle of all of these categories of features.
- Syntactic features apply within the VP- and NP-shells we identified in sections 4 and 5. Broadly, for VP- and NP-shells respectively, C/D is the discourse domain, T/Num is the inflection domain, and V/N is the thematic domain (all of these categories are functional heads).
- E.g. adjective phrases (which are NP-adjuncts) contain degree (‘very’, ‘how’ etc) and the adjective itself
- Person, number, gender, case and interrogative/declarative etc aspect are all licensed by verbal and nominal functional heads (e.g. in the sentence ‘the boys talk’, the plural number feature is applied to both the NP and the VP).

- **T and the DP in specTP agree for the values of their Person and Number formal features.**
- Formal features can be intrinsic or optional. Intrinsic formal features are those stored in lexical entries and are therefore ‘called’ whenever the lexical item is used. Their value is always determined before the lexical item is used - e.g. ‘el gato’ is always masculine. Optional formal features, meanwhile, are determined in context with each particular use of the lexical item. In the sentence ‘the cat drank milk’, the optional features of ‘the cat’ are [3rd, Sg., Nom.] and the optional features of ‘drank’ are [Sg., 3rd, past].
- Interpretable features are visible at one or more of the interfaces; they are syntactically-relevant semantic features. Uninterpretable features do not play a role in the word’s representation beyond the syntax. E.g. ‘the cat’ is visibly 3rd person and singular, so the corresponding features are interpretable, while the nominative case feature is not. Meanwhile, ‘drank’ is visibly past tense, but not visibly 3rd person or singular.
- Formal features tend to occur within sentences in interpretable/uninterpretable pairs - so the NP-shell might have an interpretable feature while the same feature is uninterpretable in the VP-shell or vice versa. The exception in English is case features, which are always uninterpretable.

Principle of Full Interpretation A derivation converges if and only if all of the features that arrive at the interface levels PF and LF are interpretable at that level of representation.

- In other words, interpretable formal features survive to reach the interfaces (where they are interpreted), but uninterpretable formal features must be eliminated by LF else the derivation will crash. A correct derivation will trace the syntactic features which are semantically relevant and represented at LF back to the syntactic level, *and no other features*.
- Uninterpretable formal features:
 - φ -features - i.e. person, number and gender - in VP
 - Case (always)
- Interpretable features:
 - φ -features in NP
 - Tense, aspect and mood in VP (note that these are not uninterpretable in NP, but nonexistent in NP, since these are purely verbal properties - person, number and gender can and do reflect on verbs, however, even if they originate in NP)
 - Clausal typing, e.g. Q and Irr

6.3 Agree

Activity Condition Given two heads H_1 and H_2 , H_1 has an uninterpretable feature $[uF]$ distinct from the F matching H_2 .

- For instance, ‘el gato bebió leche’ satisfies the Activity Condition because the masculine feature is interpretable on N and not on V .
- For a pair of heads, the $[uF]$ is known as the Probe and the matching $[iF]$ as the Goal.
- Conditions for Agree on internal arguments (e.g. the theme and instrument in ‘Tom opened the door with his key.’):
 - V has $u\varphi$
 - D has $i\varphi$ (singular etc.)
 - V asymmetrically c-commands D
 - Any VP-shells intervening between V and D lack φ -features (to preserve locality).
- In the internal argument case, Agree eliminates the $u\varphi$ in V and values the uninterpretable Case feature in D as accusative. The new accusative feature in D may or may not be realised in inflectional morphology (i.e. on PF); it is not realised in English.
- Conditions for Agree on external arguments:
 - T has $u\varphi$
 - D has $i\varphi$
 - T does not asymmetrically c-command D in this case. We claim that asymmetric c-command is required in order for T to contain the probe, as it does in the above internal argument case. Because T (probe) does not c-command D (goal), Agree cannot simply eliminate the uninterpretable feature of the probe as it does in the internal argument case. For this reason, a movement operation must occur to satisfy the c-command requirements, alongside Agree (in this case, A-movement to specTP from what would originally have been a VP complement position c-commanded by T). Before this movement occurs, the DP (containing the goal/interpretable feature) is in an internal argument-like position.
 - Any head intervening between T and D lacks φ -features (to preserve locality).

VP-internal subject hypothesis All subjects, even the subjects of active sentences, originate in VP and are moved to their surface position.

It follows that in an active-passive coordinate structure the subject of the active verb in fact originates inside VP.

- T is meant to asymmetrically c-command D in order for it to have a Probe corresponding to the Goal in D; the VISH accounts for why this is not the case on a surface level, because when we see a sentence with an external argument, movement has already occurred to place DP in the specTP position which we see.
- However, we still need to motivate this movement. The Case Filter does not motivate movement in this situation, because this is not to do with the licensing properties of the verb (e.g. in passives); it is motivated by c-command relations, not case relations.

EPP-feature A hangover from the now-disproven Extended Projection Principle, the idea that clauses universally have a subject in specTP, which dictates that an instance of Agree co-occurs with movement [to specTP].

- The EPP-feature triggers A-movement to specTP and also triggers the uninterpretable feature in T to become interpretable in order to interface with LF.
- If V has an EPP-feature, the argument goes from V to specTP along with the uninterpretable feature it carries (if the verb does not move past V, this generates SOV word order).
- Agree+EPP gives a general account of A-movement and agreement: local, cyclic, to a c-commanding position, and driven not by case but by the c-command requirements of $u\varphi$ -elimination. Since passives, unaccusatives etc. all agree, they can be explained by this account.

The Stray Affix Filter Where T lacks an affix, we have Agree without movement (i.e. ‘lowering’ of the affix from T to V, or ‘affix-hopping’).

- Motivating affix-hopping/its difference from movement:
 - In a language like French (which has obligatory V - \bar{i} T head movement), adverbs can separate the (main) verb from its object. In English (no V - \bar{i} T movement), the facts are exactly the reverse (je fais souvent... vs. I often do...). Lowering cannot jump across an intervening projection, whereas Head Movement can¹. So in a negated verb structure, we predict that T cannot lower onto V, because of the intervening Neg. This is the motivation for do-support – do is inserted to “support” the features in T (which would otherwise be flagged ungrammatical by the Stray Affix Filter, which says basically that you cannot pronounce a bound morpheme in isolation).

Featural Relativised Minimality This states the HMC and AMC in terms of features. No relation can be hold between X and Z if there

exists an intervening Y where Y has the same featural properties as X. A minimality domain for Z is determined by the interaction of X and Y's features.

7 Hierarchy and linear order

- Principle of phrase structure: heads and complements form a constituent which excludes everything else (X') (always sister nodes under X')
- Parameter of phrase structure ('Head Parameter'): heads either precede or follow their complement in X' (head-initial or head-final order). This is a linear order distinction, not a hierarchical one.
- Harmonic orders, i.e. a combination of head-final order with postpositions or head-initial order with prepositions, is significantly more common across the world's languages.
- However, some languages - such as German - do not have a consistent harmonic pair and vary orders depending on the clause type. If the Head Parameter has to be relativised to categories, is it really a single parameter?
- The phonological interface certainly requires information about linear precedence, since it is manifested in phonology - however, head-initial and head-final languages do not differ at LF (they have the same clausal domains). So we claim that only hierarchy matters to the LF, and linearisation (mapping of the elements in the hierarchical representation to their linear precedence) occurs at some point in the PF interface. **The Head Parameter applies at the point of linearisation.**

7.1 Antisymmetry in syntax

- Theory proposed by Kayne (1994). Core proposals:
 - Asymmetric c-command always maps into linear precedence.
 - Constituent order is alwaysn fundamentally Specifier-Head-Complement [so movement operations cause head-finality??]
 - Principles of X'-theory are derived from asymmetric c-command
- This enables linear order to fit into a configurational (i.e. defined by c-command) framework
- Definitions:
 - $d(X)$ (where X is non-terminal) := the set of terminals that X dominates.
 - A := set of all pairs of non-terminals in an asymmetric c-command relation

- ‘Total Ordering’ is the ordering of constituents (categories) which themselves are made up of linearly ordered components

Linear Correspondence Axiom For a given phrase marker P , $d(A)$ is a linear ordering on T where T is the set of terminals.

- The LCA is violated if orderings within constituents are not given.
- A slightly different definition of c-command is used in this framework, wherein X c-commands Y iff every category dominating X also dominates Y .
- Below are representations of two different linear orders:



- In both of these orderings, A includes the pairs (S,H), (H,T) and (S,C); the only difference is that Hn (head) and C (complement) are in different orders.
- By LCA, the linearisations of both structures are the same because the asymmetric c-command relations are the same; this eliminates the need for the Head Parameter (no non-surface difference between head-initial and head-final linearisation).
- Kayne (1994) evidences this typologically:
 - Specifier-head order dominates strongly across the data (e.g. wh-elements raise to specCP, and never move rightwards)
 - Spec-TP is almost always clause-initial
- Word order differences must therefore be captured by postulating differences wrt the amount of leftward movement (of heads and XPs) that takes place:

- SVO: default/no movement
- SOV: leftward movement of a category containing the object over V
- VSO: V moves over the subject
- VOS: VP moves (pied-pipes) over the subject
- Meanwhile, **iterated roll-up** yields head-finality by reversing the first-merged order of elements:
 - Head-initial: [.VP V O]
 - Move O around V: [.VP O [V (O)]]
 - Move (higher) VP around V: [.VP [.VP O [V (O)]] V (VP)]
 - Surround with TP: [.TP [.VP [.VP O [V (O)]] V (VP)] T (VP)]
 - Surround with CP: [.CP [.TP [.VP [.VP O [V (O)]] V (VP)] T (VP)] C (TP)]
- The number of iterations controls a great deal of word order variation.
- The degree of harmony is controlled by the presence or absence of movement triggers (i.e. generalised EPP features).

The Final over Final Condition A head-final phrase XP cannot immediately dominate a head-initial phrase YP (in a given local domain), so sentence-final complementisers are not found in VO languages (unlike in OV languages in which there has been movement).

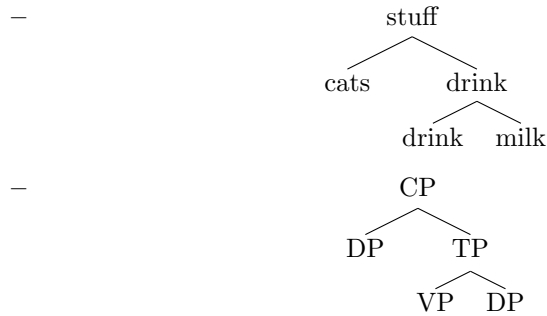
- This is a dependency between two different heads, and so is not captured by the Head Parameter.
- By LCA, X' only needs to encode hierarchy and projection, not order.

8 Bare phrase structure

- c-command encodes hierarchy and projection (the requirement to encode order is eliminated by the LCA). Ideally, we can separate out hierarchy and projection and derive c-command from a hierarchical primitive.
- $\text{Merge}(X, Y) = K(X, Y)$ where K is either X or Y (i.e. either X or Y projects)
- Whichever of X and Y projects can be defined as the head of the object formed by Merge, e.g. $\text{Merge}(\textit{drink}, \textit{milk}) = \{\textit{drink}, \{\textit{drink}, \textit{milk}\}\}$ because 'drink' is the head of the VP.
- Merge can be iteratively, recursively applied to yield progressively higher nodes in the structure: $\{\textit{cats}, \{\textit{drink}, \textit{drink milk}\}\}$ represents the DP which is a sister node to the previously merged head (the VP) (so now Y

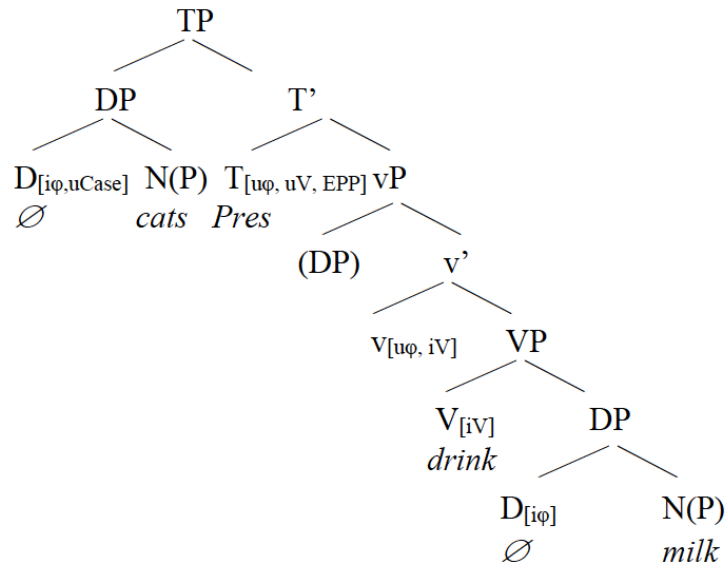
= {*drink*, *drink milk*}). This can be reapplied until the highest node is reached.

- Merge defines c-command relations as follows:
 - If X merges with Y, X and Y are sisters.
 - If X merges with Y forming {X, Y} with label K, K contains (dominates) X and Y [and is either X or Y].
 - X asymmetrically c-commands Z iff Z is contained in Y, the sister of X (so in the example above, ‘cats’ c-commands ‘milk’ because milk is contained by the merged sister constituent Y)
- C-command can thus be derived from Merge as the transitive closure of sisterhood and containment.
- {*cats*, {*drink*, *drink milk*}} is equivalent to



- X'-theoretic notions defined in terms of Merge (in terms of bare phrase structure):
 - A complement is whatever is first merged with a given head (the first Y, and therefore found to the right of the head), i.e. merged with a minimal constituent.
 - A specifier is whatever is merged second (X in relation to the merged Y), i.e. merged with a non-minimal constituent.
 - X is a maximal projection (XP) iff the label of the category immediately containing X (the output of Merge(X,Y) for some Y) is distinct from that of X.
 - X is a minimal projection (X) iff X contains no category distinct from itself.
 - (Else X is an intermediate projection (X'))
- ‘head’ is a projection-based label rather than a hierarchical one; a head H is the category that provides the label for the object formed by merging H with something else.

- Merge is binary, therefore all branching is binary. There can be multiple specifiers, merged in succession.
- Terminal nodes are equivalent to the category they are inserted under (i.e. not constituents).
- In summary, if we add all formal features to the tree generated by bare phrase structure, we get:



- Features like Pres and φ can be merged, just like categories.

The Extension Condition Merge is strictly bottom-up and serial. Exactly one of the merging pair must be the root of the tree.

- Merge and Move:
 - In the tree above, T has an EPP feature and a $u\varphi$ corresponding to the $i\varphi$ feature in DP (which is in specVP).
 - $\text{Merge}(T, D) = \{T, \{T, D\}\}$ where T is the label of the category formed by merging T and D.
 - This is known as ‘internal merge’: the root is merging with something it contains. This is triggered by the EPP or by affix-hopping.
- Merge and the LCA:
 - Asymmetric c-command always maps into linear precedence; the hierarchies generated by Merge are linearised by LCA.

- The linear order is directly dictated by asymmetric c-command relations in any tree - not much more to it
- Note only categories matter to the LCA, not features like Pres, because only PF-realised features are important to linear precedence
- OV word order is generated by moving the final DP leftward over V into some position which asymmetrically c-commands V (e.g. specVP). This can be triggered by an EPP feature.
- To make Merge and the LCA compatible, X' cannot c-command its specifier.
- In Kayne (1994), the head always linearly precedes the complement, and this follows from an asymmetric c-command relation - the head c-commands its complement, fundamentally. This is a problem because the complement must be hierarchically lower than the head, but this creates a unary branch because the complement will therefore be on its own level (nothing can intervene between the head and the complement). This makes Kaynes' linearisation account incompatible with the binary branching rule.
- To make Merge a purely hierarchical primitive, rather than covering hierarchy and projection under Merge (arguably unparsimoniously), we define labels separately from Merge.

The Labelling Algorithm Either:

1. X is a head, YP is not a head, therefore X provides the label
 2. Neither XP nor YP are heads, therefore movement occurs to leave a trace behind which will not count for labelling. For example, when a DP originates in VP and is raised to specTP by VISH, the remaining VP is labelled V.
 3. Both X and Y are heads, therefore the complex head provides the label
- The Labelling Algorithm can be used to motivate the raising of a formerly VP-internal subject to specTP: when DP (a complex head) is in VP, it is unclear how to apply the Labelling Algorithm, and so when DP is raised to specTP, we know that VP is in fact a VP.
 - Any successive-cyclic movement (including movement to specCP, as in wh-questions) is subject to the same condition: the wh-element (or DP) moves until the algorithm no longer fails.
 - Recap:
 - Hierarchy: by merge $(V,D) = \{V,D\}$ there exists a syntactic object which contains V and D and nothing else.

- Projection: D is not a head but V is a head, so we can say by the Labelling Algorithm that the VP is the maximal projection (as opposed to the DP on its right branch)
- Linear order: V asymmetrically c-commands and therefore linearly precedes D, by the LCA
- Category: the categories V and D (and others) bear formal features

9 Wh-movement

- We have accounted for the essential elements of phrase structure - hierarchy, projection, linear order and category - using Merge, the Labelling Algorithm, the LCA, and formal features respectively. We must now account for locality.
- A few movement facts (recall movement is implied by the existence of a nonlocal dependency) can be derived from Agree:
 - Two heads may agree if they are in an asymmetric c-command relation, where the c-commanded head has an interpretable feature matching the uninterpretable feature of its c-commander, and nothing intervenes between them.
 - An EPP feature on the Probe (the c-commander) causes argument (NP) movement; the association of an abstract affix ([-Af]) like Q or Irr with the Probe causes head movement, as in conditional inversion.
 - Movement is local because an intervening head is not allowed (which means that long-distance movement must occur cyclically), and any moved constituent will move to a position which c-commands the trace it leaves behind.
 - E.g. to give conditional meaning to the sentence 'The students had handed in their work on time', 'had' (which is in T) is c-commanded by 'the students' (specTP) and has the interpretable feature represented by the abstract affix -Irr which is uninterpretable on the DP in specTP. Therefore, 'had' in T moves to C driven by Agree and [-Af]. The new position C is the sister of TP and (remember specTP, which is the DP, is the sister of T'!) and asymmetrically c-commands the old position T.
- Featural Relativised Minimality describes the instances where intervention blocks movement (like intervenes for like).
- Wh-movement is the final main type of movement, apart from A-movement and head-movement. Like A-movement and head-movement, it leaves a trace at the extraction site and moves the XP containing a wh-element to

a c-commanding position. However, it is apparently not local. It invariably moves the phrase containing the wh-element to specCP, regardless of whether there exists an intervening c-commanding head.

- Because wh-movement can be fronted over an arbitrary quantity of intervening material, **it is an unbounded dependency**. The task that follows is to develop a theory of unbounded dependencies.

9.1 Island constraints

Complex NP constraint It is impossible to extract from relatives (“which politician did you write a play [which was about (which politician?)])” or sentential complements, e.g. indirect statements (“which writer did you believe the rumour [that we had met (which writer?)]”))

Coordinate Structure Constraint It is impossible to move a conjunct or an element thereof (“the lute which Henry plays (which) and sings madrigals is warped”). Note that ‘the students who fail the final exam or who (the students) do not do the reading will be executed’ is legal because movement out of each element in the conjunction relation occurs in exactly the same way, i.e. the two conjuncts are identically affected by wh-movement.

Left Branch Condition No NP which is the leftmost constituent of a larger NP can be moved out. From ‘I played Mick’s friend’s favourite guitar’, the only legal wh-movement is ‘whose friend’s favourite guitar did you play?’ because splitting the NP and wh-moving the leftmost constituent(s) is illegal (‘whose did you play friend’s favourite guitar?’). This seems not to be applicable to languages without determiners.

Sentential Subject Constraint It is impossible to extract from a clause which is in the subject position of another clause. This was extended to all subject positions (no movement out of a subject): ‘which politician were the supporters of (which politician) arrested?’

Wh-island constraint No extraction out of a clause introduced by another wh-phrase (‘which car were you wondering how you should fix (which car)?’)

Adjunct island constraint It is impossible to extract from an adjunct (‘who did you go to London without talking to (who)?’)

9.2 Subjacency

- How exactly do we generalise all of the island constraints?

- Claim: no rule can relate X to Y when X asymmetrically c-commands Y and *more than one bounding node* (TP or DP) separates X and Y. This is known as subjacency.
- In fact, **the apparent unboundedness of wh-movement is simply cyclic local movement.**
 - In the sentence ‘which books did you force Bill (which books) to buy (which books)?’, the wh-element moves to specCP of the indirect question/subclause, then to specCP of the main clause. Each step of movement crosses just one TP (exactly one bounding node), therefore subjacency is respected.
- Subjacency rules out islands because movement out of islands requires movement of a wh-element to the main clause specCP position, thus crossing two TPs and violating subjacency.
 - In other words, if a subclause exists, it must be provable that the wh-element moved *through the subclause specCP position* before moving to main clause specCP. The wh-island constraint applies because this intermediate movement is not possible when the subclause specCP condition is already occupied by a wh-element.
 - So in ‘Who were you wondering which book we gave (which book) to (who)?’, ‘who’ cannot move out of the wh-island introduced by ‘which book’ because the intermediate specCP is blocked.
 - It is impossible to reconcile subjacency and wh islands by moving each wh-element in succession, because by the Strict Cycle Condition it is impossible to apply a rule to the most deeply embedded cyclic domain *having already applied a rule to the whole CP*.

Comp-to-Comp Condition A phrase in a CP can only move upwards, to a higher specCP position.

- More examples of subjacency constraining movement:
 - Complex NP (relative clause): “Which race did you hear the announcement that John had won?” Originally ‘which race’ is in the object DP and the postulation is that it moves to specCP of the subclause, then to specCP of the main clause. This is impossible by subjacency because there is an intervening TP (T=did) and DP (the announcement).
 - Wh-island: “Who did Mary read the book which we gave to?” Originally ‘who’ is sentence-final and movement would require it to go to the sentence-initial specCP position in one step. There are two intervening TPs (did read, gave) and one intervening DP (the book).

- Sentential subject: "Which politician was that the police would arrest expected?" Originally 'which politician' is in the object DP position of 'arrest'; movement requires crossing the DP 'the police' and two TPs (was, would).
- Generalised subject: "Which politician were the supporters of arrested?" Intervening TP 'were', intervening DP 'the supporters.'
- Left branch: "Whose did you play favourite guitar?" Intervening TP 'did (play)', intervening DP 'favourite guitar.'
- Movement constraints which are not accounted for by subjacency:
 - Coordinate structure: "Who did John talk to and?" The original sentence here would be "John talked to (who) and (who)", so the required wh-movement here would be movement out of a VP, which is not directly accounted for by subjacency because VPs are not bounding nodes.
 - Adjunct: "Who did you go to London without talking to?" The original position of 'who' is in the PP 'to who?' - it is not clear how crossing a PP (which)

9.3 Phase theory

- A phase is essentially a more modern conceptualisation of a cycle, i.e. a complete domain in which syntactic computation can occur and which abides by the strict cyclicity principle. Chomsky appeals to the features which come under different maximal projections in order to determine whether or not they constitute phases.
- For instance, a CP contains both the arguments with attached nominal lexical features (e.g. case) and those with attached verbal features (tense, aspect), and these arguments are organised within the CP in a complete hierarchical structure. While subjacency did not account for the Coordinate Structure Constraint since it did not afford any special status to VP nodes, phase theory holds that a VP is a phase because it contains a complete hierarchy of the arguments in a verb's lexical semantics.
- The phase-theoretic account of constituent movement involves successive-cyclic progression of a constituent through the 'edges' of a phase, where an edge is defined as a domain through which a constituent can transit on its way to an upper clause. This clearly accounts for comp-to-comp movement: the intermediate positions are the edges corresponding to the intermediate CPs.
- It also accounts for the difference between CPs/VPs (which are phases and have edges through which moving constituents can transit, yielding longer displacement) and the 'bounding node' TPs and DPs (which do

not, and can therefore only yield one-step displacement – e.g. in the case of the complex noun phrase constraint).

9.4 Barriers

- Claim: Any category can be a bounding node (aka a barrier) for subadjacency, depending on the structural environment and relations it enters into.
- For instance, adjunct islands are not completely accounted for by subadjacency (since PPs are not bounding nodes) and the bounding status of APs and PPs seems conditional. A barrier (essentially a generalised bounding node) is therefore defined as follows:
- A barrier is a head which is either non-lexical or immediately dominates the extraction site. Alternatively, a barrier may be separated from the extraction site by a category which is either not a lexical head *or* dominates the extraction site. (E.g. T' intervenes between a subject DP and the object DP extraction site, so the intervening component is a non-lexical head barrier; the verb in T is a sister to DP and so does not count as an intervening element). If a category is separated from an extraction site by a lexical head or a sister node to the extraction site, it is not a barrier.
- Subjects (barrier DPs), adjuncts (barrier PPs), CPs in relative clauses and TP in wh-islands (which are intrinsic barriers) are all barriers. In sentential complements under the complex NP constraint, however, the CP is lexically marked by (i.e. corefers with) the noun which the relative pronoun refers to and therefore cannot be a barrier. This leads to a weaker violation because only the CP is crossed and it is necessary to claim that it is an inherent barrier. The same applies to the wh-island constraint; moving across a wh-island only requires crossing one barrier (CP).
- Barriers can be voided by adjunction. For the sentence 'Which books did Bill buy?' we assume that the wh-element adjoins to VP and so VP does not actually dominate the extraction site (i.e. the wh-element originates in the sister of VP).

10 Covert wh-movement

- Recap: how a grammar is constructed
 1. Merge, i.e. external and internal merge (where the former is 'normal merge', and the latter does not interact with PF and "combines already-formed syntactic operators") builds syntactic structure by creating asymmetric c-command relations. Internal merge is the same thing as 'move' apparently.

2. Certain parts of the structure are transferred to the PF to be phonologically represented.
 3. The LCA linearises the structure.
- Merge, PF-realisation and the LCA together give rise to a legible structure at the phonological level. Meanwhile:
 1. The remaining (non PF-realised) parts of the structure are read at the LF.
 2. The syntactic representation at LF specifies elements of semantics.
 - **Covert movement** does not affect the PF-representation, but may be deducible from the way a sentence is interpreted or from cross-linguistic similarities which are not immediately evident from linear order (i.e. from the PF).
 - Using the covert movement device, linguists can suggest that Chinese - despite not having instances of wh-movement which are obvious at the PF-level - has covert movement which is analogous to the overt movement in English, as follows:
 - Indirect questions involve wh-movement to intermediate specCP; direct questions involve movement to main clause specCP.
 - This covert movement takes place *after* translation to the PF has already occurred.
 - Correspondingly, we can describe the selectional restrictions of the wh-elements in the same way as we can for overtly-moved elements (selectional restrictions of wh-elements are descriptions like ‘the complementiser phrase introduced by ‘that’ is an obligatory complement’)
 - We should also be able to expect that islands exist, and they do generally appear to.

10.1 Quantifiers

- Quantifiers can be viewed as an instance of covert movement; wh-elements and quantifiers are operators, and the elements they govern are variables.
- Scope ambiguities in the sentence ‘every student dreads some lecture’:
 1. Wide scope for ‘every student’: ‘every student dreads at least one lecture.’ (this is the default reading because the first quantifier in the linear order is interpreted as having wider scope)
 2. Inverse scope, i.e. wide scope for ‘some lecture’: ‘every student dreads a particular lecture.’

- Linear precedence does not completely capture scope, because inverse scope readings are not ungrammatical and linearisation only takes place after transfer to PF.

Configurational Scope A quantified phrase has scope over everything it c-commands at LF (i.e. before any PF-realisation has taken place).

- It follows from the above definition that the quantifier phrase must c-command different elements (i.e. precisely the elements in its scope) under each interpretation, and the reason both interpretations eventually converge to the same realisation is due to a movement operation which converts them to their final, PF-realised form.
- In the original, LF order, the quantifier with the widest scope is placed first: ‘every lecture some student dreads’ gives order 1 above, and ‘some lecture every student dreads’ gives order 2. This occurs via TP adjunction (a new TP in the sister of the quantified DP).

10.2 Copy theory

- Why is the trace of a movement operation, which is left in the extraction site position, not realised at PF?

Inclusiveness Condition No new objects are added during the course of the computation apart from rearrangements of lexical properties.

- This is why we prefer to think of traces as simply ‘copies’ of the moved item, and simply hold that only the highest version of a copied item is realised at PF (however, both are present at LF). ‘Which songs did Paul write which songs?’ is a correct LF representation.
- This is also why a sentence like ‘Which picture of himself does John like (which picture of himself)?’ is still grammatical; although ‘John’ does not c-command the element it is coindexed with at the PF level, it does at the LF, and the anaphor is simply not realised within the binding domain at the PF (but because it is at the LF, it still counts).
- For anaphors, the LF seems to prefer the lowest copy for coindexing; for names/R-expressions, it is the highest copy.
- LF c-command relations provide semantic interpretation in other cases too:
 - See the section on the HMC, ‘to’ and modals. Movement to T on the LF level seems to denote concrete existence, or a non-hypothetical tense feature.
 - For instance, ‘it is likely that a hippogryph will be executed’ does not imply the existence of a hippogryph, but ‘a hippogryph is likely

(a hippogryph) to be executed' might and 'a hippogryph is anxious to be executed' definitely does.

- This is because there are also copies left behind by A-movement (e.g. raising). **There is no copy left behind on LF by '... is anxious to be ...'.**
- The above scenarios, where copies provide semantic interpretations, are known as reconstruction. Quantifiers/covert movement operators have the lower copy realised at PF; reconstruction has the higher copy realised at PF.

11 Phases

- In 'there-' sentences, 'there' occupies the subject position, blocking any A-movement to the subject position.
- In the sentence 'There seems to be someone here', both T and V must precede N, because 'there' "merges in the embedded clause" (combines with the verb in the embedded clause such that it moves as a unit with the verb??). This merging within the embedded clause precedes the application of the movement rule, i.e. 'there' is in its position relative to the verbs in the sentence before it is moved to the subject position. This is an instance where "external merge precedes internal merge."
- 'There' is merged as soon as possible where it exists, because it is the primary candidate for satisfying the requirements of the EPP feature in T (=moving to subject position). When 'there' is not present, the internal argument of the unaccusative (the DP, 'someone') raises to subject position.
- By this account, it should be impossible for the DP to intervene within any part of TP; however, 'there was somebody seeming to be here' is fine. Why?
 - Claim: the computational system only works with sub-parts (phases, which are cycles but better) of the structure at a given point in the derivation.
 - Phases include CP (any and all) and VP where V is accusative-assigning (INFINITIVES CANNOT ASSIGN CASE!) OR does not trigger A-movement (not unaccusative, passive or raising). So any T, or a V accompanying non-finite T, does not introduce a phase, however agreeing/case-assigning V does.
 - A phase is structured as $[\alpha[H\beta]]$ where α is the phase edge and β is the complement.

Phase Impenetrability Condition Only the head and the edge which introduces the phase are accessible to operations from outside the phase.

- So the reason we say ‘there was somebody seeming to be here’ is because ‘seeming’ is essentially a nominal modifier and so the DP is not splitting up T at all.
- Phases can account for island effects, e.g. a wh-island is caused by a wh-element blocking off a phase edge and disallowing movement via the edge.

11.1 Phases and A'-movement

- Argument movement is associated with particular theta roles and grammatical functions, e.g. subject movement to specTP. This movement is **phase-internal**. A'-movement, meanwhile, is movement to the edge of a phase.
 - A-movement: ‘John seems (John) to have been arrested (John).’ The EPP-feature in the finite T requires an argument in specTP; the argument originates as the theme in VP, moves cyclically to the second specTP (subject position of the nonfinite T) and must move further because this T is nonfinite. SpecTP, the subject position, is not a phase edge because TP is not a phase.
 - A'-movement: ‘What did Sarah say (what) Jenny was sending (what)?’ This is cyclic A' movement via the intermediate specCP position, i.e. via the edge of the phase denoted by CP.
- The above example of wh-movement (i.e. an instance of A'-movement) shows that copies can exist in specCP position. They can also exist in certain (accusative assigning, non-A-movement-licensing) specVP positions, as this example shows:
 - ‘Which of the papers that he gave Mary did every student ask her to read carefully?’ - the wh-phrase must be below ‘every student’ so that the pronoun is bound, and ‘Mary’ must be above ‘her’ so that it (an R-expression) is not bound. ‘ask’ marks the edge of the VP, and the wh-element moves from sentence-final position through the edge preceding ‘ask’ in order to satisfy the binding constraints. This shows that specVP, like specCP, can be an intermediate position in successive-cyclic movement.
 - If a wh-element or a copy thereof directly precedes an agreeing or non-auxiliary verb, movement to specVP has occurred.

Antilocality Condition Movement of A targeting B must cross a projection labelled distinctly from B.

- Who did you hear [DP the [(who) [NP rumour [edge (who) that a dog bit (who)]]] ? - the intermediate movement from the edge of the complement NP to another NP is illegal by antilocality, and any further movement violates PIC.
- This rule does not apply to VP because an external argument can be merged in the target VP.

11.2 Relativised Minimality

- Essentially the equivalent of the general intervention constraint (like intervenes for like) but for features as opposed to heads: when X asymmetrically c-commands Y and Y asymmetrically c-commands Z, no relation can hold between X and Z when Y has the same featural properties as X.
- ‘Which problem do you wonder how to solve (which problem)?’, where ‘which problem’ in specCP c-commands its sentence-final trace, and ‘how’ (which has the Q feature) intervenes, is more acceptable than ‘how did you wonder which problem to solve (how)?’ This exemplifies how relativised minimality applies to adjunct wh-movement (when the wh-element is a left-branch quantifier within DP), but less strictly to argument wh-movement (when the wh-element is a direct object).
- This is because arguments are +Q, +N, whereas adjuncts are only +Q: arguments have a superset of the features of the wh-element introducing a direct question.
- To generalise: barriers/bounding nodes/phases produce islands and prevent movement to a c-commanding position, which is known as impenetrability, whereas some constituent Y in X, Y, Z which has shared properties with X blocks direct relations (movement or agreement) between X and Z, and this is known as intervention.

12 Silent syntax

12.1 Ellipsis

- Ellipsis is an instance of silent structure, which is syntactic structure invisible to the phonological level of representation. Copies are an obvious example, as are the ‘dummy pronouns’ in the subject position of control infinitives (‘John tried to be nice’; the second specTP is occupied by a dummy pronoun).
- English is fairly generous with respect to ellipsis; it allows ellipsis of verb phrases following auxiliaries, modals, do/pro-forms (this is called pseudogapping) and in relative clauses. Most European languages do not allow VP-ellipsis after auxiliaries (but they do after modals), and have limited possibilities for pseudogapping.

- It seems that **elements in the T-system** license ellipsis:
 - ‘Mag wants to read Fred’s story and I also want to [VP]’, ‘John’s leaving but Mary’s not [VP]’ - ‘not’ (neg), aux and infinitival ‘to’ seem capable of permitting VP-ellipsis
 - ‘Sally made John laugh and then Bill made’/‘Sally started learning French, but only after John did’ (in the latter case, either ‘did’ or ‘to’ is needed in order to render the sentence grammatical) - aspectual and causative verbs, which *take VP complements* (since their complements are nonfinite but lack ‘to’ in T) do not license ellipsis.
- Languages with V-to-T movement (Ils réservent déjà leurs vacances) have *temporal affixes in T* (and these do not block movement, unlike Pres in English). Such languages also seem to disallow VPE following auxiliaries.
- Claim: affixal T prevents the ellipsis of post-auxiliary verb phrases? VP-ellipsis *following a modal* in these languages is permitted as simply a case of TP-ellipsis.
- Ellipsis of infinitival VPs (following to) is sensitive to islands, hence the ungrammaticality of the following contexts:
 - ‘You shouldn’t play with rifles because *[to...] is dangerous*’ (subject island)
 - ‘You woke up early in order to study and I woke up in order to [...]’ (adjunct island; any *purpose clause* containing ‘to’, whether or not the whole ‘in order to’ phrase is included, is an *adjunct* situation)
 - ‘We wanted to invite someone, but we couldn’t decide *who to [...]*’ (wh-island)
 - ‘Fred doesn’t want to talk to Mary but Bill *has a plan to [...]*’ (complex NP; note that ‘plans to’ is ok, but since ‘a plan’ is nominal, this switch to a sentential complement renders the sentence ungrammatical) (recall that complex NPs are {relatives, sentential complements})
- **If island constraints apply, movement must be occurring**
- Finite VPE is not island-sensitive (‘I wonder who did’ etc) so it seems that the movement is related specifically to ‘to’
- Note the grammaticality difference of ellipsis in the cases of ‘I am [VP] too’ and ‘I’m too’ - it *seems that non-affixal T, which permits VPE, excludes contracted auxiliaries but includes any form of ‘to’* (contracted ‘to’ is fine provided it doesn’t give rise to incorrect syntax)

12.2 Indexation



- Coindexation interpretations seem to be more liberal in the case of ellipsis, it is often possible to interpret indexation as strict:
 - ‘John loves his mother and Bill does too’ - the elided pronoun (‘his mother’ is elided on the second occasion) could have the same antecedent as the overt pronoun (John) or a different one (Bill)
 - ‘Mary loves John and he thinks Susan does [love John] too’ - this should be a Principle C violation since ‘John’ is an R-expression and it is c-commanded in its binding domain by ‘he’, which coindexes it, but somehow it is still grammatical. It seems that the elided copy of ‘John’ is changed to a pronoun and its binding is therefore grammatical.
 - This phenomenon is known as **vehicle change**. It applies to a *lot* of mismatches between elided copies and grammatically correct forms, including voice mismatch (which only works when identical Voice heads are being elided), negative polarity mismatch, agreement mismatch (with ‘be’).
 - The most liberal coindexation interpretations are available when the pronouns don’t c-command each other and are not potential antecedents for each other.

Rule H (yeah this is a dumbass name for a rule) A pronoun P can only be bound by an antecedent A when there is no closer antecedent B where binding P to B could give the same interpretation.

- There is some open problem concerning why sentences with ellipsis through islands are ok (‘We want to hire a linguist but we can’t decide who (we want to hire)’). Eliding an entire sentential complement (?? as in ‘who (we want to hire)’ is known as **slicing**.
- **Deep anaphora** involves lexically-inserted pronouns whose antecedents may be pragmatically (rather than syntactically/formally) determined (e.g. pronouns such as ‘John threw out his letters and Mary threw *them* out too’, null complements such as ‘I don’t see why you even try [...]’, and ‘it’ (‘she finally did it’))
- **Surface anaphora** involves syntactically-controlled deletion (VP-ellipsis, slicing)
- Deep anaphora does not allow for strict/sloppy ambiguity (i.e. for agreeing pronouns, elided complements and ‘it’, there is only ever exactly one possible referent).

12.3 Null pronouns

- Are there null objects or pronominals (i.e. distinct from the rest of the NP) or is there NP-ellipsis?

Tomioka's Generalisation Null pronouns in discourse pro-drop languages are a result of N'-deletion/NP-ellipsis without determiner isolation.

- **If only a strict reading is possible/no reference ambiguity exists, argument ellipsis - rather than pronoun ellipsis - is occurring.**
- Null pronoun languages (like Spanish) occasionally permit alternation between null forms and pronominal forms in embedded clauses. When pronominal forms are used, they are bound in their binding domain by the null referent in the main clause.
- If the embedded pronoun is null, it only has one possible referent: the aforementioned subject ('mislim da je očito' - the subject of the copula can only be impersonal/expletive). This means null pronouns have to be an instance of deep anaphor, similar to overt pronouns in non-null pronoun languages.
- In expletive pronoun situations, the only way to fill the subject/specTP position (by EPP) is to postulate that some dummy pronoun exists in the gap (to which deep anaphor applies).
- Properties of consistent null pronoun languages (Spanish, Croatian etc):
 - The subject of a finite clause can be silent and still have all other properties expected of a subject (such as providing a referent)
 - Free subject inversion (the subject can easily follow the verb)
 - Wh-movement is permitted across an *overt, non-null complementiser*
 - 3sg null subjects are only ever definite; arbitrary or non-definite null subjects need to be marked specifically (e.g. 'ne može se...' - the reflexive marker acts as an indefiniteness marker)
 - Rich agreement inflection on finite verbs (it is unclear how rich; German has non-null subjects and Chinese has null subjects)
- East Asian languages have pro-drop without agreement cues; this gives rise to more ambiguity than exists in European languages with richer agreement, hence we consider this to be argument ellipsis rather than pronoun ellipsis.
- Properties of these languages (known as 'radical null0-subject languages'):
 - No restrictions on pronoun omission
 - No agreement inflection on the verb

- No restriction on the value of the definiteness feature
- No free inversion
- Argument ellipsis, i.e. surface anaphora
- Properties of partial null-subject languages:
 - Person-based restrictions on pronoun omission
 - Agreement inflection not necessary
 - Indefiniteness is available to 3sg null pronouns without an indefiniteness marker
 - No free inversion; subjects raise to specTP or higher
 - No semantic difference from overt pronouns
- The distinction between consistent, radical and partial null subject languages is expressible in terms of richness of agreement and marking of definiteness (the latter is exclusive to consistent NSLs; the former is common to consistent and partial NSLs).
- Consistent NSLs have definite/indefinite articles or rich clitic systems; partial NSLs have restricted or no articles as well as rich inflection; radical NSLs completely lack articles and agreements
- In summary (ϕ = agreement):
 - RNSL: $[_{DP} \emptyset_{[-\phi]} \dots [_{NP} \text{surface}]]$
 - PNSL: $[_{DP} D_{-[\text{def}]} \dots [_{NP} \text{deep}]]$
 - CNSL: $[_{DP} D_{[-\text{def}, +\phi]} \dots [_{NP} \text{deep}]]$
- In non-NSLs, are overt pronouns simply definiteness markers (with agreement, similar to the properties of CNSL Ds)?
- **The expression of definiteness in T is linked to agreement richness** (and hence to subject nullity, at least for CNSLs and PNSLs)

13 Parameters

- How to adequately describe and account for language variation has been a central issue in linguistic research. In the framework of Principles and Parameters (Chomsky 1981), efforts were focused on identifying parameters accommodating clusters of differences across languages. The null subject parameter, connecting the availability of null subjects and others such as extraction possibilities, agreement, etc., was an influential example (Rizzi 1982; Jaeggli & Safir 1989; among many others). Unfortunately, such large-scale clustering of cross-linguistic variation turned out to be more ideal than reality (Gilligan 1987; Newmeyer 2004; Boeckx 2014; Paul

2015). Accordingly, attention has been shifted to smaller-scale differences, i.e. moving from the search for macroparameters to the identification of microparameters. For microparameters, triggers have generally been attributed to feature specifications of lexical items, specifically functional heads. This is the line of research led by Borer (1984) and Chomsky (1995), basing on the “Borer-Chomsky conjecture” as termed by Baker (2008).

- Nonetheless, the shift to identifying microparameters raised questions about the status of macroparameters, as noted in Kayne (2005; 2013), Baker (2008), Gianollo et al. (2008), Holmberg (2010), Roberts & Holmberg (2010), Huang (2014), Huang & Roberts (2017), Roberts (2017), among others. The special “Parameter” issue of *Linguistic Analysis* (41: 3–4) raised fundamental issues on the role of parameters in capturing cross-linguistic variation. There have been different views on the status of macro and microparameters (grammatical and lexical).
- One among them is to take microparameters as the core and to search for as-many-as-possible microparameters that seem to converge to an overarching macroparameter. That is, a macroparameter is the converging product of the clustering of properties describable in terms of microparameters. Much effort in this line of research is on the comparison of dialects within a language family or among the varieties of closely related language families.
- A different view is articulated in Baker (2008), which compares microparametric and macroparametric syntax and shows that there can be variation in the grammar proper (macroparameter) (as opposed to the lexicon). That is, there are some parameters within the statements of the general principles that shape natural language syntax—the so-called “grammatical parameters”, because they concern principles of grammar that cannot be localized in the lexicon per se, in contrast to “lexical parameters” or microparameters, presupposed by the Borer-Chomsky conjecture: **all parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon.** Baker notes that the parameters of a more or less macroparametric sort he presented in the paper were discovered by comparing Bantu languages with Indo-European languages, which probably would not have been achieved if we focus on microcomparative syntax.
- The original notion of parameters was that they are ‘options’ associated with principles.
 - E.g. the formation of X’, via the merging of the head and the complement, is a principle
 - The accompanying parameter is that complements can either precede or follow their heads.

- Language acquisition is the gradual adaptation of parameters to their adult form.
- The reduction of parameters to differences in formal features (i.e. the adjustment of the feature value to 1 or 0) (Borer-Chomsky) is a more recent approach. It reduces parameter values to the lexicon (the consciously-acquired part of language)
- Microparameters refer to classes of functional heads, e.g. modal auxiliaries, subject clitics. Microparametric views predict that there are more languages that look like roughly equal mixtures of two properties than are pure (pure is the macro view)
- NB. Typical FFs stand in an arbitrary relation with semantic features, just as phonological and semantic features are in an arbitrary relation with each other. In this way, the ‘substantive’ FFs ‘piggy-back’ on semantic features (e.g. Person in the hierarchy above).
- All UG contributes to substantive FFs is just the template that makes them able to participate in Agree relations, which Biberauer terms the ‘[uF]/[iF] template’. The actual list of substantive FFs is not prespecified.