LING5702: Lecture Notes 9
Broad-coverage Syntactic Competence

Until now we've only seen processes for recognizing phonemes, words or flat sequences of words.

But that won't help us assemble meanings, which need to distinguish restrictors and nuclear scopes.

This lecture will describe **syntactic** competence (knowledge about **structure** in language).

Specifically, this broad-coverage model is used in van Schijndel & Schuler (2015); Oh et al. (2022).

Contents

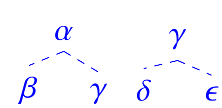
9.1 We compose meanings out of sub-parts (Frege, 1892)

We can assemble meanings by:

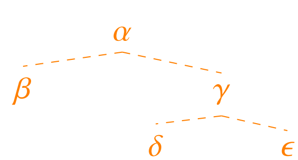
1. composing words into phrases and sentences using **grammar rules**; and then
2. associating these words and rules with meanings via function application in lambda calculus.

Specifically, we use **context-free grammar rules**, which assign categories to sequences of words.

They are 'context-free' (Chomsky, 1956) because they just have to match categories to apply.

These rules can be written like this: $\alpha \rightarrow \beta\ \gamma$, $\gamma \rightarrow \delta\ \epsilon$ or drawn like this:  ,

and they can be used to assemble trees by matching categories (the γ's in both rules match):

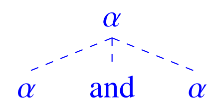


9.2 Coordinating conjunctions: identifying syntactic categories

Human languages seem to have rich category types we can see in conjunctions (Sag et al., 1985).

First, observe that natural languages use different argument structures for different verbs:

- They sleep. – one argument ahead (intransitive)
- They find pets. – one argument ahead and one argument behind (transitive)
- They give people pets. – one argument ahead, two arguments behind (ditransitive)

Next, observe natural languages **coordinate** conjunctions (combine like types): 

.

- [$_{\beta}$ [$_{\beta}$ They sleep] and [$_{\beta}$ they find pets]]. – sounds ok (β is sentence)
- They find [$_{\gamma}$ [$_{\gamma}$ people] and [$_{\gamma}$ pets]]. – sounds ok (γ is noun phrase)
- *They find [$_{\gamma}$ [$_{\beta}$ they sleep] and [$_{\gamma}$ pets]]. – sounds wrong; conjuncts must match

Now, allowable conjunctions give us insight into the category structure of language:

- They [$_{\delta}$ [$_{\delta}$ sleep] and [$_{\delta}$ find pets]]. – sounds ok (δ is verb phrase)
- They [$_{\eta}$ [$_{\eta}$ find] and [$_{\eta}$ give people]] pets. – sounds ok (but what's η?)

Transitive verbs (find) match type with ditransitive verb + indirect object (give people)!

Both lack argument ahead and behind – it seems types are defined by missing arguments!

9.3 Arguments: composed by rules (Ajdukiewicz, 1935; Bar-Hillel, 1953)

We formalize our set of categories C as follows, as clauses with unmet requirements:

1. every U is in C, for some set U of primitive categories;
2. every $C \times O \times C$ is in C, for some set O of type-combining operators;
3. nothing else is in C

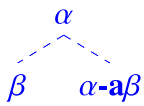
We define **primitive categories** $U = \{\mathbf{N}, \mathbf{V}\}$:

- **N**: noun-headed category with no missing arguments (noun phrase)
- **V**: verb-headed category with no missing arguments (sentence)

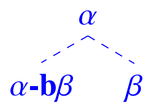
We define **type-combining operators** $O = \{\textbf{-a}, \textbf{-b}\}$:

- α**-a**β: α lacking β argument ahead (e.g. **V-aN** for intransitive δ above)
- α**-b**β: α lacking β argument behind (e.g. **V-aN-bN** for transitive η above)

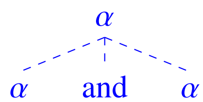
Now we can define **context-free rules** R over these categories:

(Aa) 

: argument attachment/elimination ahead

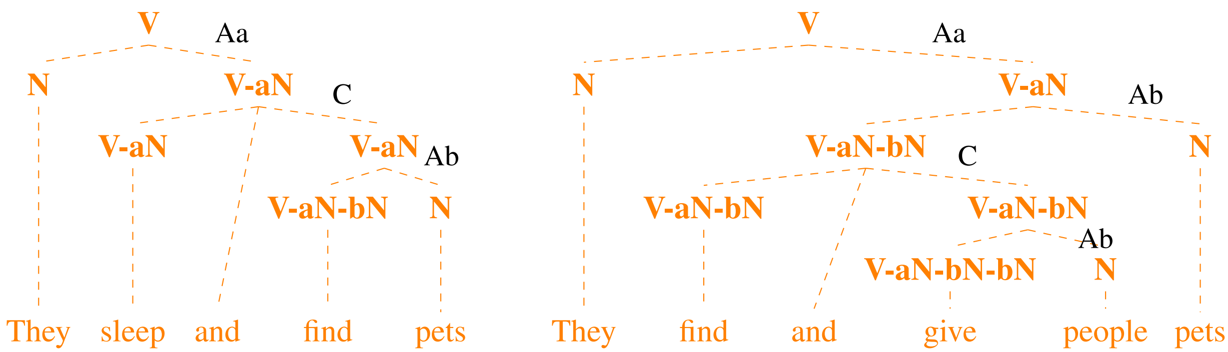
(Ab) 

: argument attachment/elimination behind

(C) 

: conjunction

These three rules model all of the above sentences:



Also note that the parents in these rules all have simpler types than the children.

This means for any lexicon (constraining types at tree leaves), the set of categories C is finite.

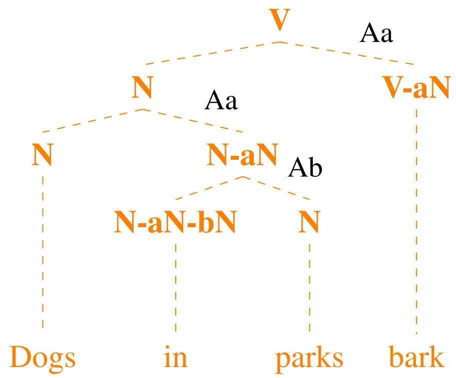
Practice 9.1:

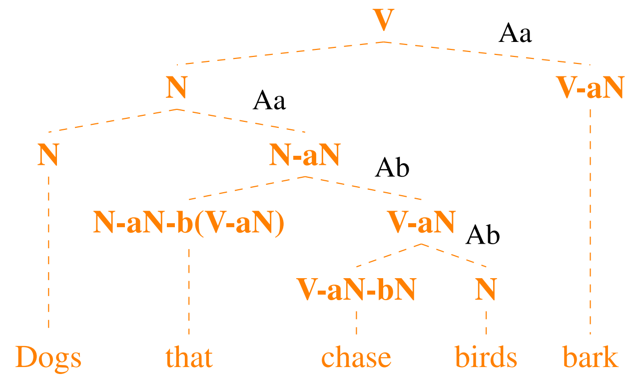
1. Draw an analysis tree for the sentence *Paris is in France* using the following categories:
   - **N** for *Paris* and *France*
   - **A-aN-bN** for *in* (**A-aN** is a predicative phrase that follows a form of *be*)
   - **V-aN-b(A-aN)** for *is* (a form of *be*)
2. Label the rules used in your analysis

9.4 Modifiers: coordination can rule out analyses (Gazdar et al., 1985)

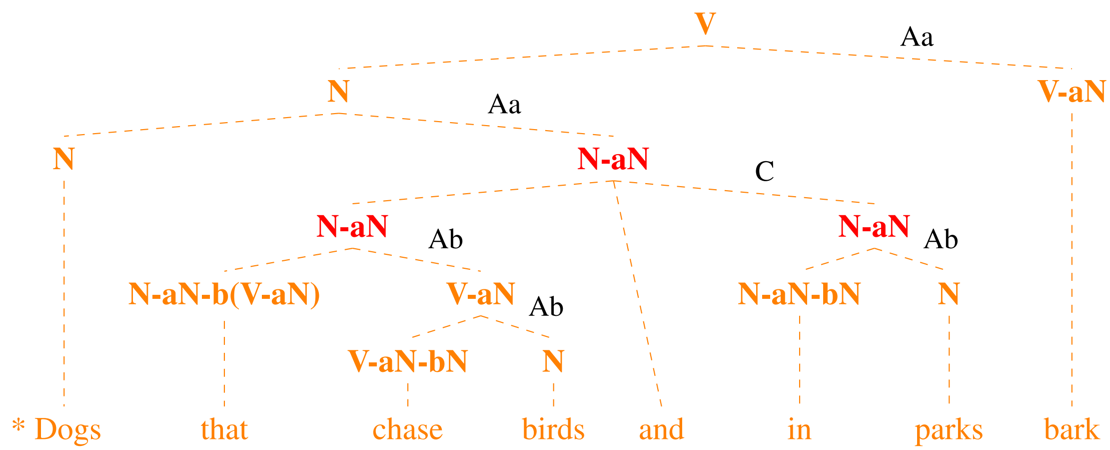
We can also use coordination failure to rule out some possible analyses!

For example, we may be tempted to use argument attachment for post-nominal modifiers:

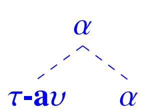




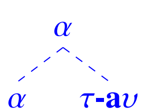
but these modifier categories do not seem to coordinate:



We can distinguish regular modifiers by adding rules that re-use e.g. post-copular category (**A-aN**):

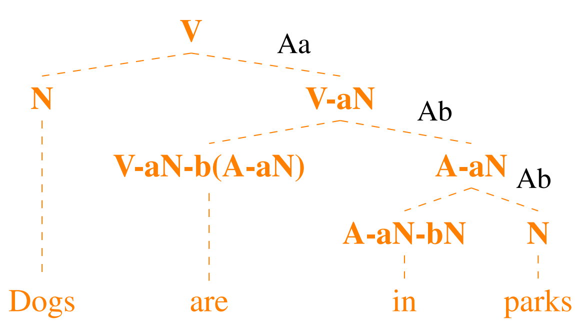
(Ma) 

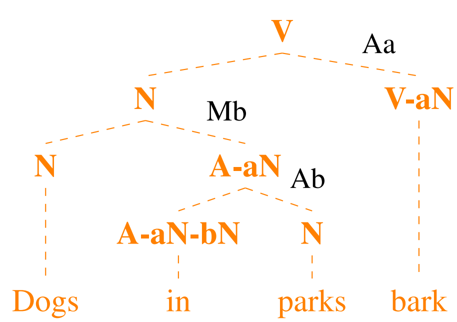
: pre-modifier (τ and υ are primitive categories)

(Mb) 

: post-modifier (τ and υ are primitive categories)

Here's the analysis for a post-copular predicative phrase re-used as a post-nominal modifier:





Note that prepositional phrases coordinate with adjectives (as **A-aN**), and with adverbs (as **R-aN**):

- Your dogs are [$_{\text{A-aN}}$ very noisy] and [$_{\text{A-aN}}$ almost in my yard].
- Your dogs are playing [$_{\text{R-aN}}$ very noisily] and [$_{\text{R-aN}}$ almost in my yard].

They relate as (unpronounced) morphological inflections, like (pronounced) inflections of adverbs.

Practice 9.2:

1. Draw an analysis tree for *cities in France are clean* using the following categories:
   - **N** for *cities* and *France*
   - **A-aN-bN** for *in*
   - **V-aN-b(A-aN)** for *are*
   - **A-aN** for *clean*
2. Label the rules used in your analysis

9.5 Non-local rules: gap fillers (Gazdar et al., 1985; Pollard & Sag, 1994)

Those were local dependencies, but natural languages may also use non-local dependencies.

In English, these show up in **topicalization**, which seem to use a gap '_' at one argument:

- These pets, you say they found _.

(This is a type of **filler-gap construction**, which includes content questions and relative clauses.)

These coordinate as well, but our test shows categories with gaps differ from those without:

- These pets, you [$_\delta$ [$_\delta$ say they found _] and [$_\delta$ think _ gave people joy]]. – sounds ok
- *These pets, you [$_\delta$ [$_{\textbf{V-aN}}$ say they found pets] and [$_\delta$ think _ gave people joy]]. – wrong

We can model this by adding a **new type-combining operator** for non-local dependencies:

- α**-g**β: α lacking non-local β argument (e.g. **V-aN-gN** for intransitive δ above)

adding rules to **introduce** non-local dependencies (propagating other dependencies $\psi_m \in \{\textbf{-g}\} \times C$):

(Ea) $\dfrac{\alpha\textbf{-g}\beta\psi_{1..M}}{\alpha\textbf{-a}\beta\psi_{1..M}}$: introduce non-local dependency to argument ahead

(Eb) $\dfrac{\alpha\textbf{-g}\beta\psi_{1..M}}{\alpha\textbf{-b}\beta\psi_{1..M}}$: introduce non-local dependency to argument behind and adding rules to **attach** non-local dependencies (propagating other dependencies $\psi_m \in \{\textbf{-g}\} \times C$):

(G) $\alpha\psi_{1..M} \rightarrow \beta\psi_{1..m} \quad \alpha\textbf{-g}\beta\psi_{m+1..M}$: non-local dependency attachment/elimination ahead

and modifying existing rules to **propagate** non-local dependencies $\psi_m \in \{\textbf{-g}\} \times C$:

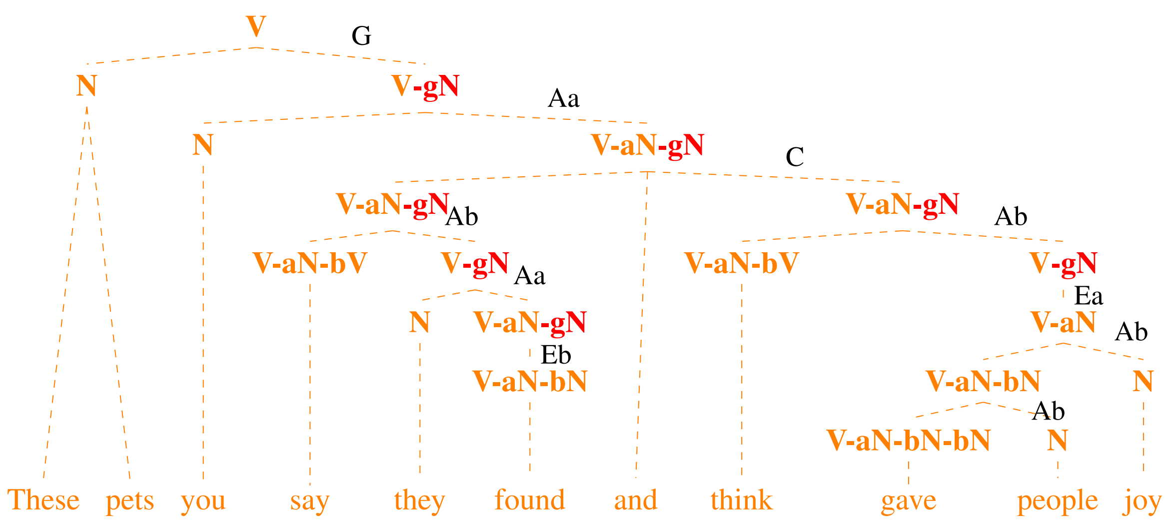
(Aa) $\alpha\psi_{1..M} \rightarrow \beta\psi_{1..m} \quad \alpha\textbf{-a}\beta\psi_{m+1..M}$: argument attachment/elimination ahead, with propagation

(Ab) $\alpha\psi_{1..M} \rightarrow \alpha\textbf{-b}\beta\psi_{1..m} \quad \beta\psi_{m+1..M}$: argument attachment/elimination behind, with propagation

(Ma) $\alpha\psi_{1..M} \rightarrow \tau\textbf{-a}\upsilon\psi_{1..m} \quad \alpha\psi_{m+1..M}$: pre-modifier, with propagation

(Mb) $\alpha\psi_{1..M} \rightarrow \alpha\psi_{1..m} \quad \tau\textbf{-a}\upsilon\psi_{m+1..M}$: post-modifier, with propagation

Here's the analysis:



Note: these schemata may not be instantiated for all combinations of categories:

- *These pets, [$_{\textbf{N-gN}}$ claims that _ are friendly] are false.

Note M above is unbounded, so our rules aren't in a finite set of categories, so aren't context-free.

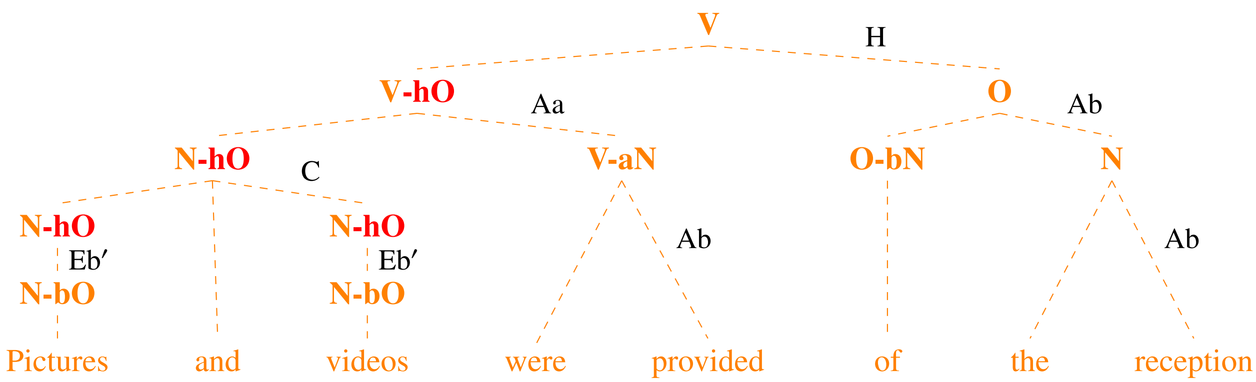
(Any number of arguments may be extracted and propagated up from children.)

Some use evidence like this to argue language isn't context-free but mildly context-sensitive (Shieber, 1985; Joshi, 1985; Steedman, 2000).

In practice, though, we can just constrain category sets to combinations seen in training data.

9.6 Non-local rules: heavy shift and extraposition (Bach, 1981)

We see similar coordination with other kinds of non-local dependencies (**O** is an oblique object):



(This is called **heavy shift**, or **extraposition** if it crosses a clausal boundary.)

Here, the extraposed *of the reception* must apply to both conjuncts:

- *[$_{\text{N}}$ Pictures of the wedding] and [$_{\text{N-hO}}$ videos _] were provided [$_{\text{O}}$ of the reception].

Heavy shift and filler-gap dependencies are distinct because they behave differently:

- subjects can be gap-fillers: [$_{\text{N}}$ These pets] [$_{\text{V-gN}}$ Pat thinks _ give people joy].
- subjects cannot be heavy shifted: *[$_{\text{V-hN}}$ Pat thinks _ give people joy] [$_{\text{N}}$ these pets].

We can model this by adding another type-combining operator:

- α-**h**β: α lacking non-local β argument (e.g. **N-hO** above)

and adding another introduction rule (propagating other non-local dependencies ψ_m):

(Eb′) $\dfrac{\alpha\text{-}\mathbf{h}\beta\psi_{1..M}}{\alpha\text{-}\mathbf{b}\beta\psi_{1..M}}$: introduce non-local dependency to argument behind

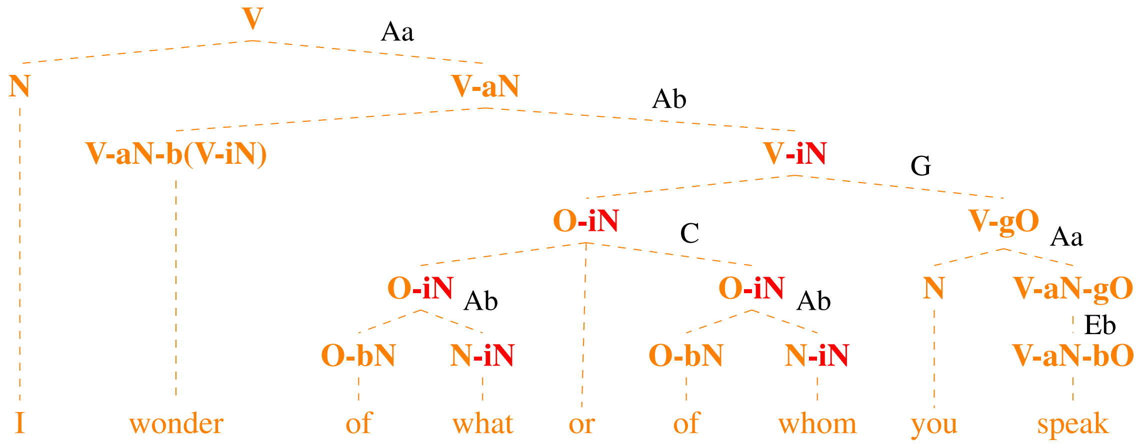
and adding another attachment rule (propagating other non-local dependencies ψ_m):

(H) $\dfrac{\alpha\psi_{m+1..M}}{\alpha\text{-}\mathbf{h}\beta\psi_{1..m} \quad \beta\psi_{m+1..M}}$: non-local dependency attachment/elimination behind

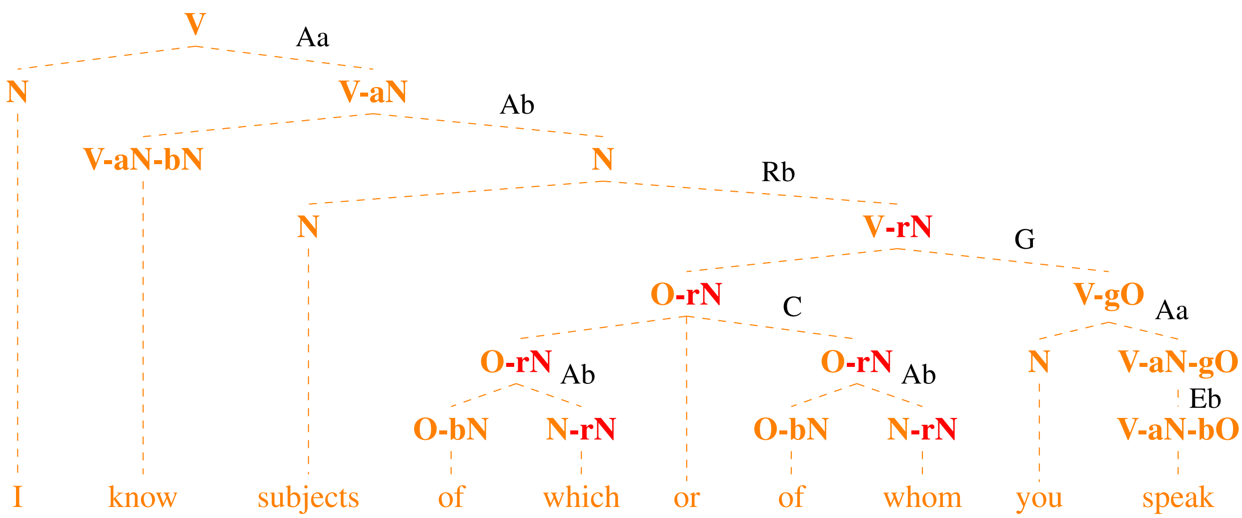
and adding the new operator to the propagation rules: $\psi_m \in \{\mathbf{-g}, \mathbf{-h}\} \times C$.

9.7 Non-local rules: pied piping (Pollard & Sag, 1994)

We also see coordination with interrogative phrases (*of what*):



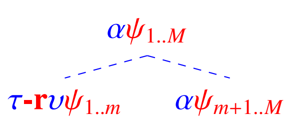
and relative phrases (*of which*):



(This is called **pied piping**, from a folk tale, because the *of*'s are 'led away' with the pronouns.)

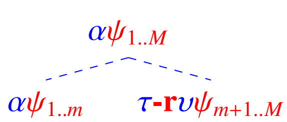
We need a new attachment rule for relative clauses (propagating other non-local dependencies ψ_m):

(Ra)



: non-local modifier attachment/elimination ahead

(Rb)



: non-local modifier attachment/elimination behind

and an additional operator in the propagation rules: $\psi_m \in \{$**-g**, **-h**, **-i**, **-r**$\} \times C$.

Pied-piping dependencies cannot cross clauses, so are distinct from other non-local dependencies:

- These are the subjects [$_{\text{N-rN}}$ a discussion of which] Pat heard _.
- *These are the subjects [$_{\text{N-rN}}$ a brother who discusses which] Pat met _.

and they are distinct from each other because they select different pronouns:

- This is the food [$_{\text{N-rN}}$ which/*what] Pat ate _.
- I wonder [$_{\text{N-iN}}$ what/*which] Pat ate _.

Practice 9.3:

1. Draw an analysis tree for *dogs which are in parks bark* using the following categories:
   - **N** for *dogs* and *parks*
   - **N-rN** for *which*
   - **V-aN-b(A-aN)** for *are*
   - **A-aN-bN** for *in*
   - **V-aN** for *bark*
2. Label the rules used in your analysis

9.8 Non-local rules: passives (Nguyen et al., 2012)

We also see similar coordination with passives (**R-aN** is an adverbial phrase):

- Those bridges were driven [$_{\text{R-aN-vN}}$ under _] and [$_{\text{R-aN-vN}}$ on top of _] by many cars.

Passives are distinct from other non-local dependencies because they cannot cross noun phrases:

- *That person was met [$_{\text{N-vN}}$ the mother of _] by someone.

We need a new attachment rule for passives (**L-aN** is a participial verb phrase):

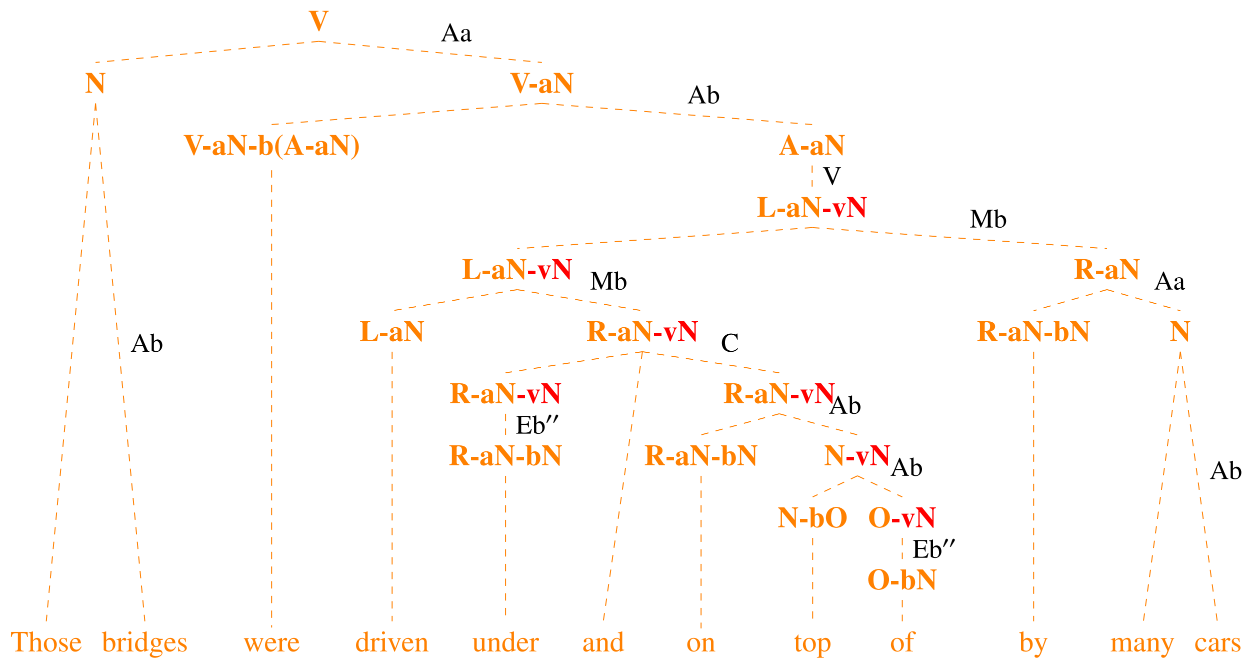
$$\text{(V)}\quad \frac{\textbf{A-aN}}{\textbf{L-aN-vN}} \quad : \text{non-local argument attachment/elimination, unary}$$

and another introduction rule:

$$\text{(Eb}''\text{)}\quad \frac{\alpha\textbf{-v}\beta\psi_{1..M}}{\alpha\textbf{-b}\beta\psi_{1..M}} \quad : \text{introduce non-local dependency to argument behind}$$

and an additional operator in the propagation rules: $\psi_m \in \{\textbf{-g}, \textbf{-h}, \textbf{-i}, \textbf{-r}, \textbf{-v}\} \times C$.

Here's the analysis (**L-aN** is a participial verb phrase):



Note passivized subjects coordinate with prepositional phrases, so aren't treated as oblique objects:

- The charges were placed [$_{\text{R-aN}}$ with every precaution] and [$_{\text{R-aN}}$ by skilled sappers].

This contrasts with prepositional genitives, which are treated as oblique objects:

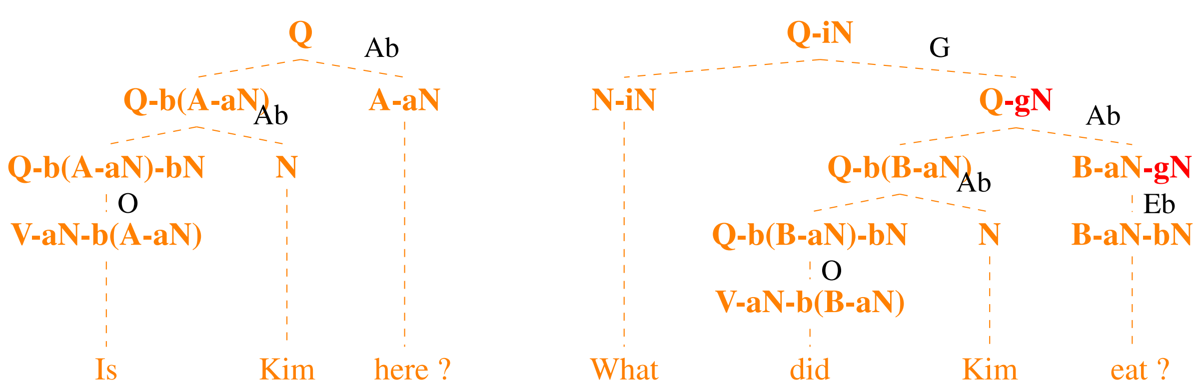
- *Mice are afraid [$_{\text{R-aN}}$ during the day] and [$_{\text{O}}$ of cats].
- *Please include examples [$_{\text{A-aN}}$ with diagrams] and [$_{\text{O}}$ of difficult problems].

That's the last of our non-local dependencies!

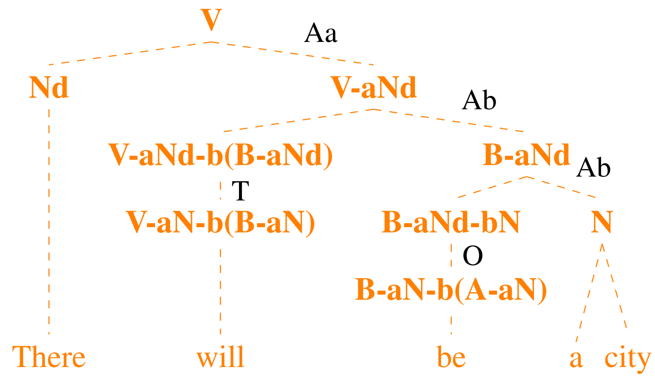
9.9 Extra: argument re-ordering to switch clause types (Nguyen et al., 2012)

Intuitively, we'd like a single cognitive decision to relate declarative and interrogative clauses.

This can be an argument re-ordering rule (**Q** is a subject-auxiliary inverted clause, **B** is base-form):



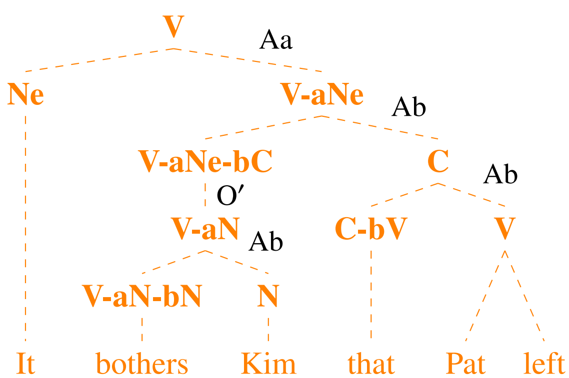
This re-ordering also relates existential clauses (the T rule, type specialization, doesn't re-order):

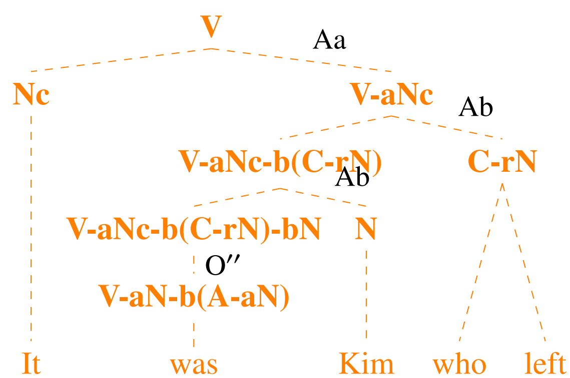


Here's the re-ordering rule, where τ, τ' are primitives and $\varphi, \psi, \varphi', \psi'$ are local dependencies:

(O) $\begin{array}{c}\tau\varphi\psi \\ \vdots \\ \tau'\psi'\varphi'\end{array}$: argument re-ordering

We need different re-ordering for it-extraposition and it-clefts (**C** is a complementized clause):





Here are the rules for that, where τ, τ' are primitive categories, $\varphi, \psi, \varphi', \psi'$ are local dependencies:

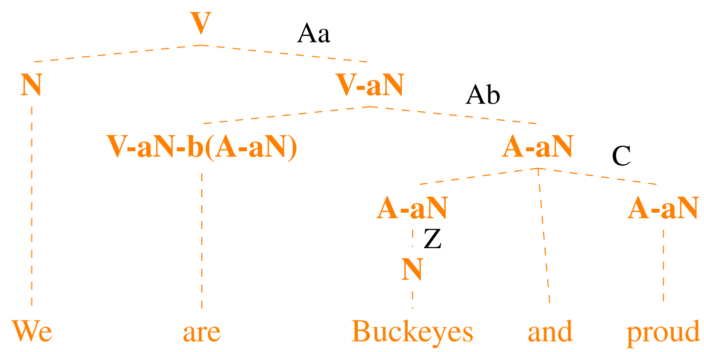
(O′) $\begin{array}{c}\tau\varphi\psi \\ \vdots \\ \tau'\psi'\end{array}$: argument re-ordering

(O″) $\begin{array}{c}\tau\varphi\psi\chi \\ \vdots \\ \tau'\chi'\psi'\end{array}$: argument re-ordering

Again: these (and other) schemata may not be instantiated for all combinations of categories.

9.10 Extra: zero-head rules to coordinate unlike categories (Sag et al., 1985)

We also need rules to handle 'unlike category coordination' of nouns and adjectives:



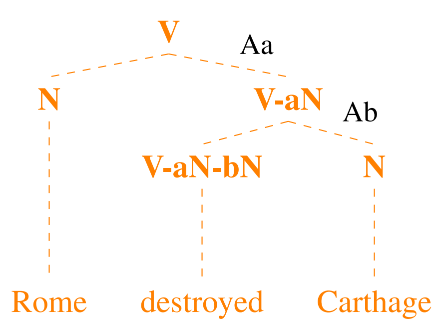
Here's the rule for that:

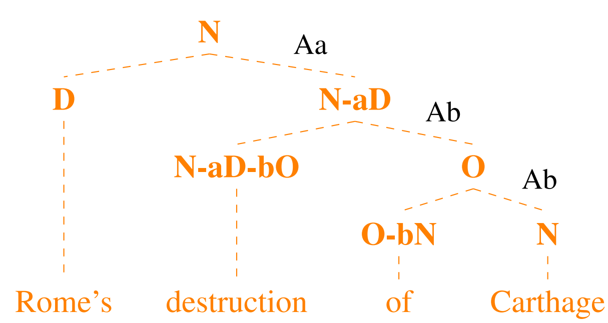
(Z) τ-**a**β / β : zero head rule

9.11 Extra: deverbal nominalizations and argument elision

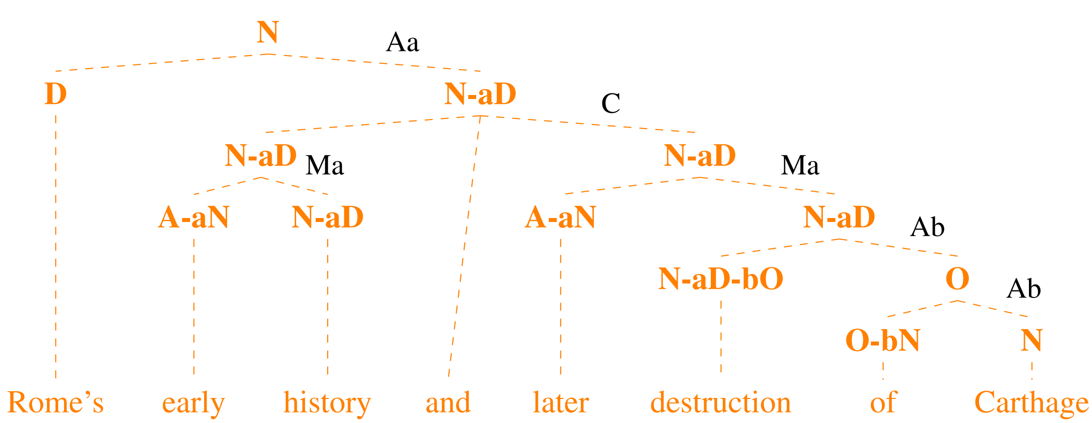
Finally, a note about nominalizations: events can be expressed as nouns as well as verbs.

We can give nouns and verbs a similar analysis (using **D** as a possessive genitive):





This then commits us to coordinate common nouns as **N-aD**:



We can elide arguments like **D** that are optional but ordered (and therefore can't be modifiers):

$$\dfrac{\textbf{N}}{\textbf{N-aD}}\ \text{T}' \qquad \text{dogs}$$

Here's the rule for that:

$$(\text{T}')\quad \dfrac{\gamma}{\gamma\varphi}\ :\ \text{argument elision}$$

9.12 Summary

This lecture notes covered a broad-coverage categorial grammar, motivated by coordination effects:

1. three local attachment rules: C (conjunct), Aa/Ab (argument) and Ma/Mb (modifier);
2. five non-local rules: E (extraction), G (gap), H (heavy shift), R (rel pro), and V (passive);
3. three extra argument re-organization rules: O (re-order), T (sub-type), and Z (zero-head).

Here's a complete list of primitive categories used in this grammar:

top-level discourse	(**T**):	[$_{\text{T}}$ They own a dog. He knows math.]
sentence	(**S**):	They say: [$_{\text{S}}$ Know math]!
finite verb clause	(**V**):	They believe that [$_{\text{V}}$ he knows math].
invinitive verb clause	(**I**):	They allow [$_{\text{I}}$ him to know math].
base-form verb clause	(**B**):	They require that [$_{\text{B}}$ he know math].
participial clause	(**L**):	He has [$_{\text{L-aN}}$ known math].
predicative or small clause	(**A**):	They keep [$_{\text{A}}$ him knowing math].
adverbial form	(**R**):	He adds [$_{\text{R-aN}}$ knowingly].
gerund clause	(**G**):	They get by without [$_{\text{G}}$ him knowing math].
particle	(**P**):	They pick him [$_{\text{Pup}}$ up].
subject-auxiliary inversion	(**Q**):	They ask: [$_{\text{Q}}$ does he know math]?
subordinate finite verb clause	(**C**):	They regret [$_{\text{C}}$ that he knows math].
subordinate infinitive verb clause	(**F**):	They wait [$_{\text{F}}$ for him to know math].
subordinate bare infinitive clause	(**E**):	They stipulate [$_{\text{E}}$ that he know math].
noun phrase or nominal clause	(**N**):	They talk about [$_{\text{N}}$ his knowledge of math].
genitive or possessive determiner	(**D**):	They calculate [$_{\text{D}}$ his knowledge of math's] effect.
oblique argument	(**O**):	They tire [$_{\text{O}}$ of his knowledge of math].

References

Ajdukiewicz, K. (1935). Die syntaktische konnexitat. In S. McCall (Ed.), *Polish Logic 1920-1939* (pp. 207–231). Oxford University Press. Translated from Studia Philosophica 1: 1–27.

Bach, E. (1981). Discontinuous constituents in generalized categorial grammars. *Proceedings of the Annual Meeting of the Northeast Linguistic Society (NELS)*, 11, 1–12.

Bar-Hillel, Y. (1953). A quasi-arithmetical notation for syntactic description. *Language*, 29, 47–58.

Chomsky, N. (1956). Three models for language. *IRE Transactions on Information Theory*, 2, 113–124.

Frege, G. (1892). Uber sinn und bedeutung. *Zeitschrift fur Philosophie und Philosophischekritik*, 100, 25–50.

Gazdar, G., Klein, E., Pullum, G., & Sag, I. (1985). *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Joshi, A. K. (1985). How much context sensitivity is necessary for characterizing structural descriptions: Tree adjoining grammars. In L. K. D. Dowty & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational and theoretical perspectives* (pp. 206–250). Cambridge, U.K.: Cambridge University Press.

Nguyen, L., van Schijndel, M., & Schuler, W. (2012). Accurate unbounded dependency recovery using generalized categorial grammars. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING '12)* (pp. 2125–2140). Mumbai, India.

Oh, B.-D., Clark, C., & Schuler, W. (2022). Comparison of structural parsers and neural language models as surprisal estimators. *Frontiers in Artificial Intelligence*, 5.

Pollard, C. & Sag, I. (1994). *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Sag, I. A., Gazdar, G., Wasow, T., & Weisler, S. (1985). Coordination and how to distinguish categories. *Natural Language & Linguistic Theory*, 3, 117–171.

Shieber, S. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 8, 333–343.

Steedman, M. (2000). *The syntactic process*. Cambridge, MA: MIT Press/Bradford Books.

van Schijndel, M. & Schuler, W. (2015). Hierarchic syntax improves reading time prediction. In *Proceedings of NAACL-HLT 2015*: Association for Computational Linguistics.