XIV UNL School

Patras, March 10-14, 2014

Day #3

Morning

- Theory of Syntax
- Subcategorization
- Composition

Lunch break

Afternoon

- Analysis Dictionaries

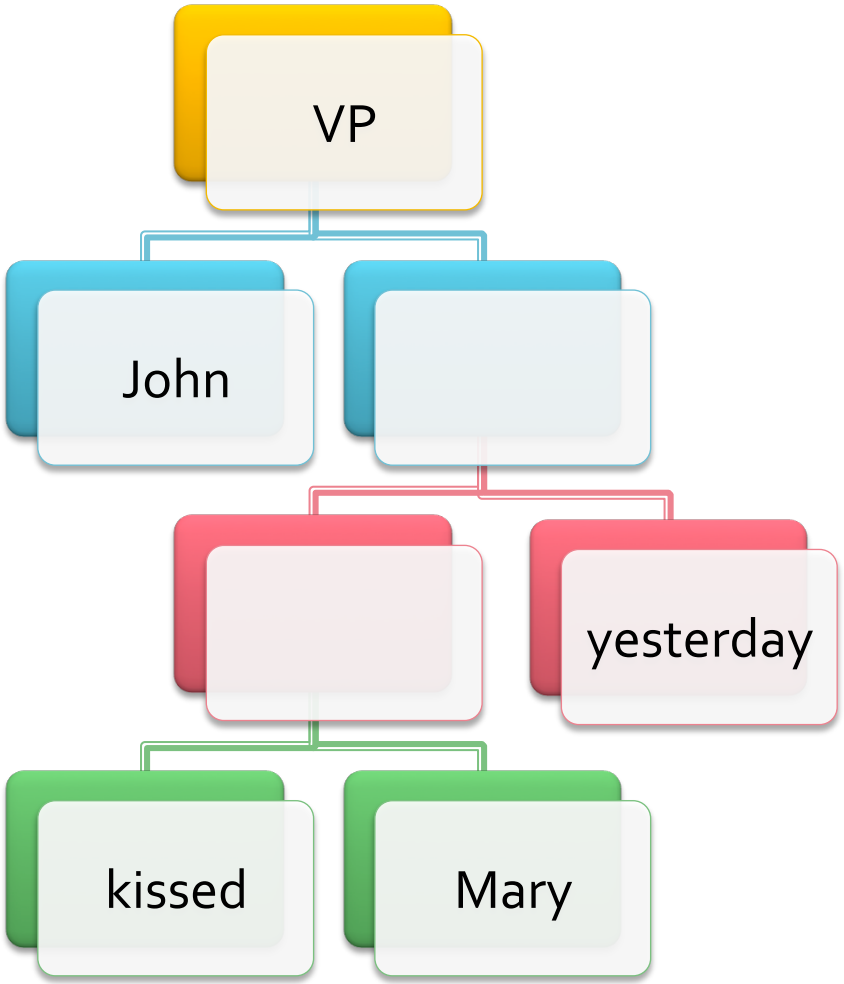
Theory of Syntax

Grammar frameworks

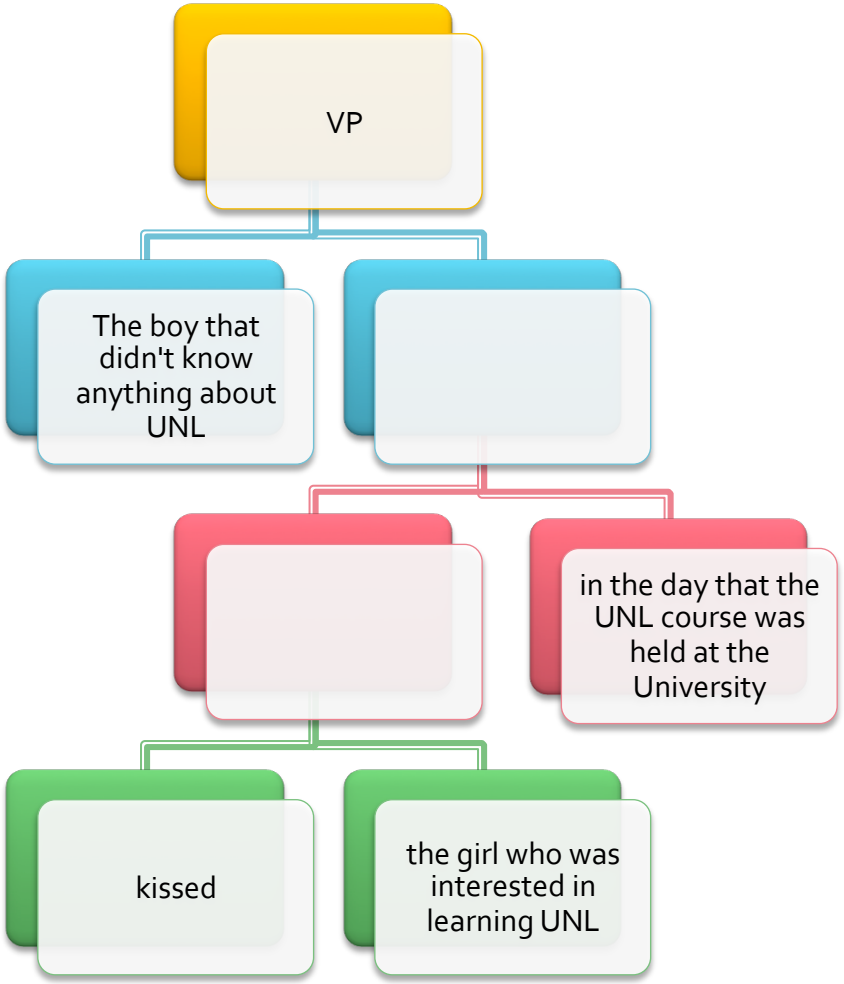
- DCG (Definite Clause Grammar)
- GPSG (Generalized Phrase Structure Grammar)
- HPSG (Head-driven Phrase-Structure Grammar)
- TAG (Tree Adjoining Grammar)
- UG (Unification Grammar)
- CG (Categorial Grammar)
- FUG (Functional Unification Grammar)
- SFG (Systemic functional grammar (SFG)
- LFG (Lexical-functional Grammar)
- Generative Grammar
  - ST (Standard Theory)
  - EST (Extended Standard Theory)
  - X-bar
  - GB (Government and Binding)
  - PP (Principles and Parameters)
  - Minimalist Theory

X-bar

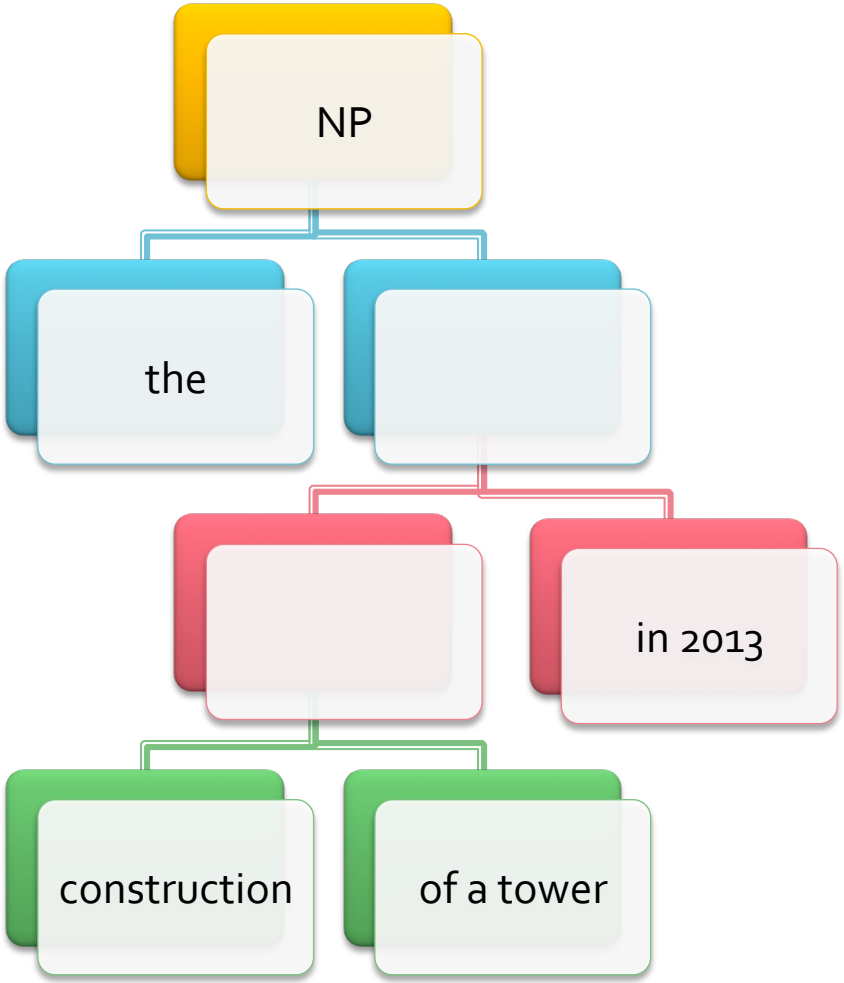
John kissed Mary yesterday



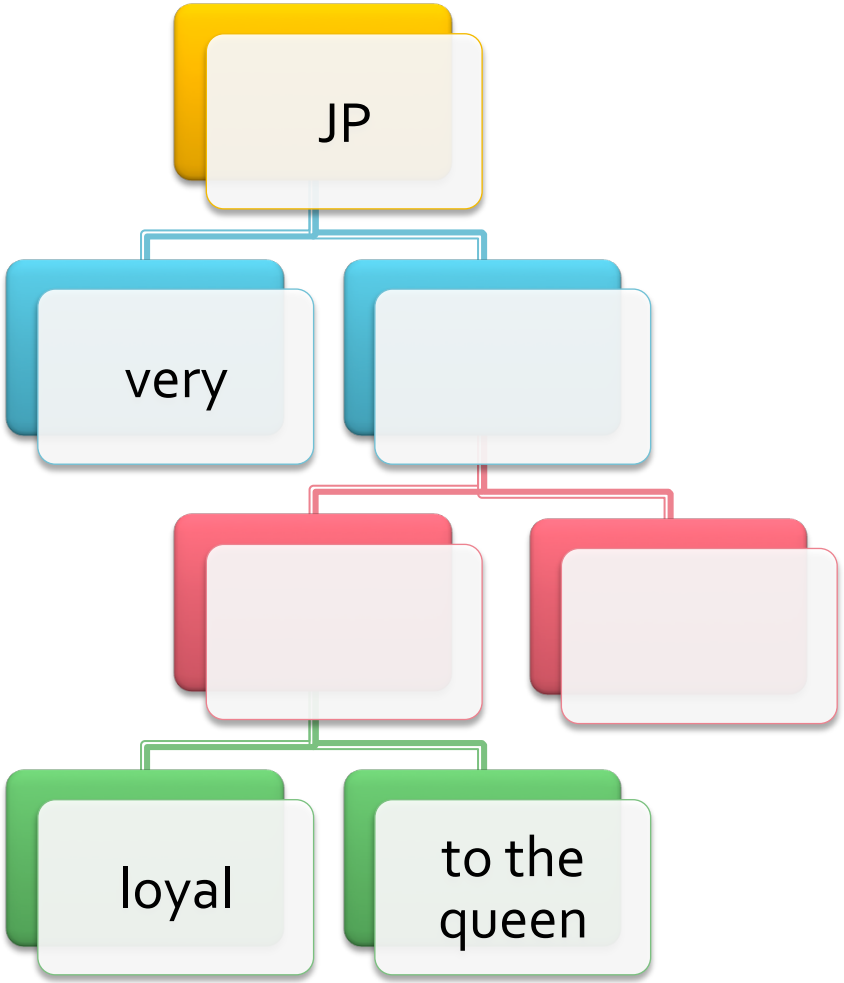
The boy that didn't know anything about UNL kissed the girl who was interested in learning UNL in the day that the UNL course was held at the University.



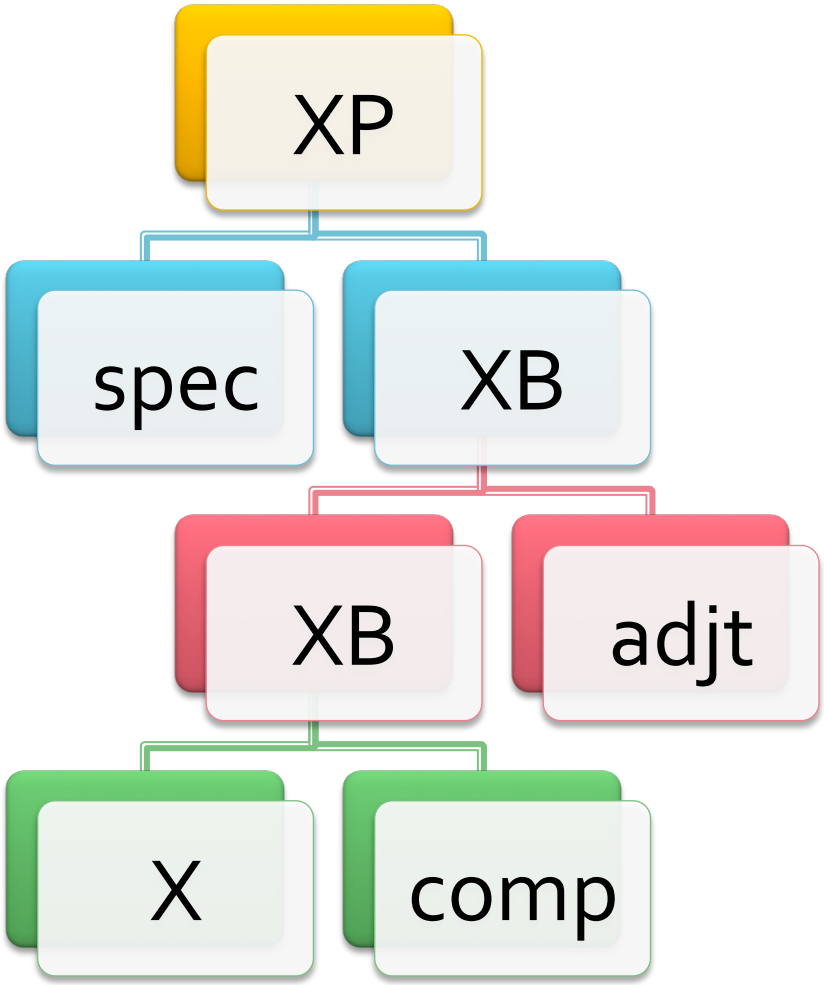
The construction of a tower in 2014



very loyal to the queen

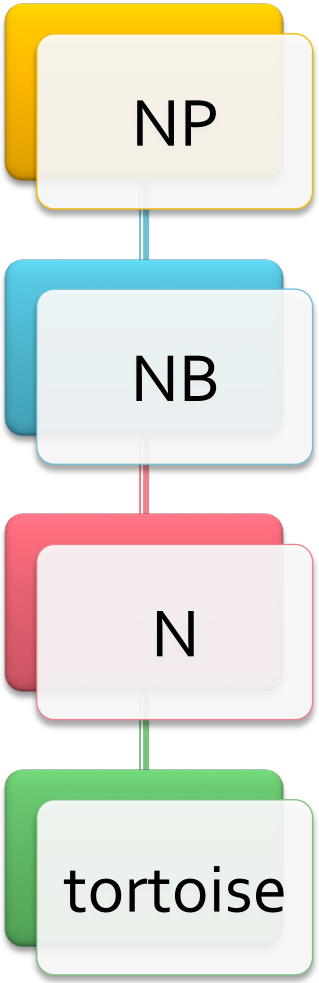


X-bar structure

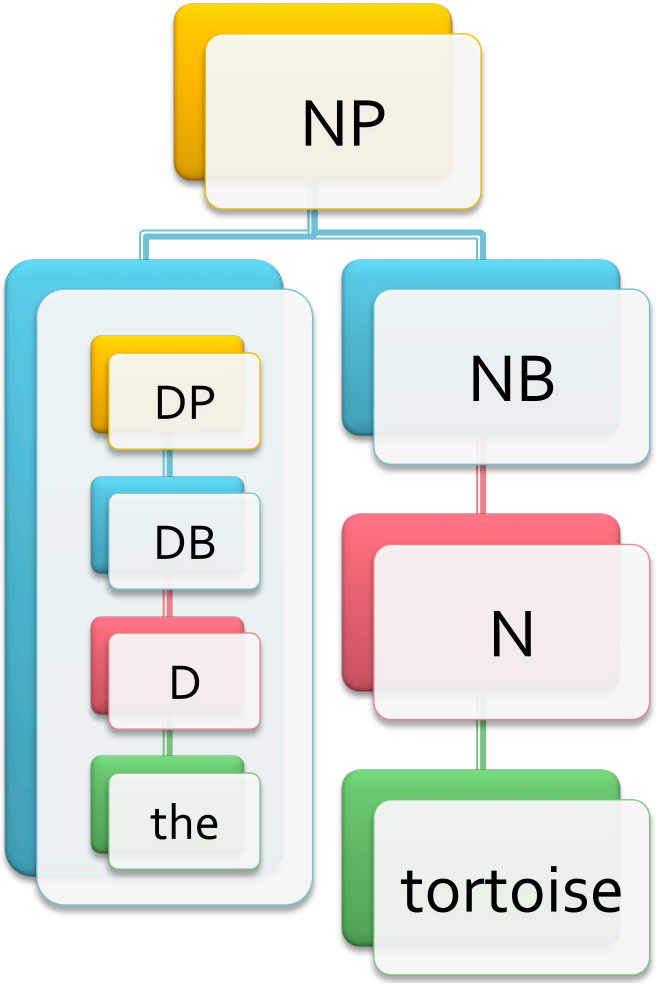


- Where:
  - XP = maximal projection
  - XB = intermediate projections
  - spec = specifier
  - adjt = adjunct
  - comp = complement
  - X = head
    - N (noun)
    - V (verb)
    - J (adjective)
    - A (adverb)
    - D (determiner)
    - P (preposition)
    - C (conjunction)

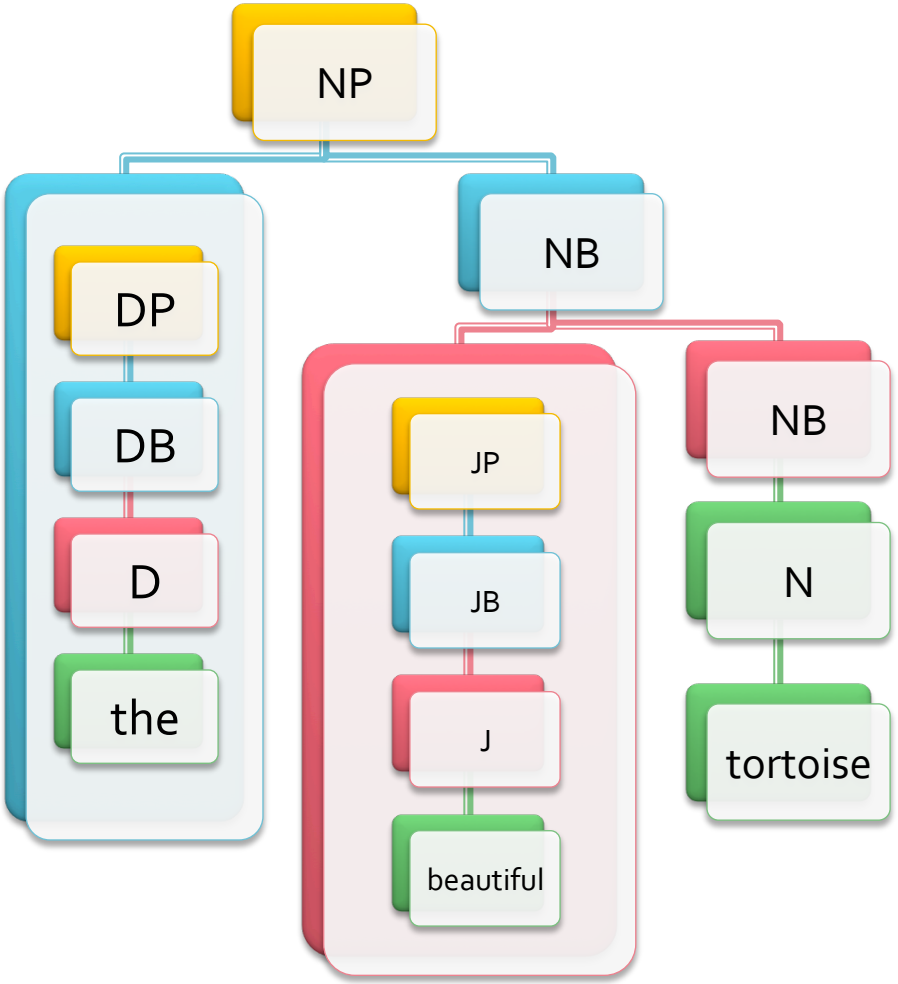
Tortoise



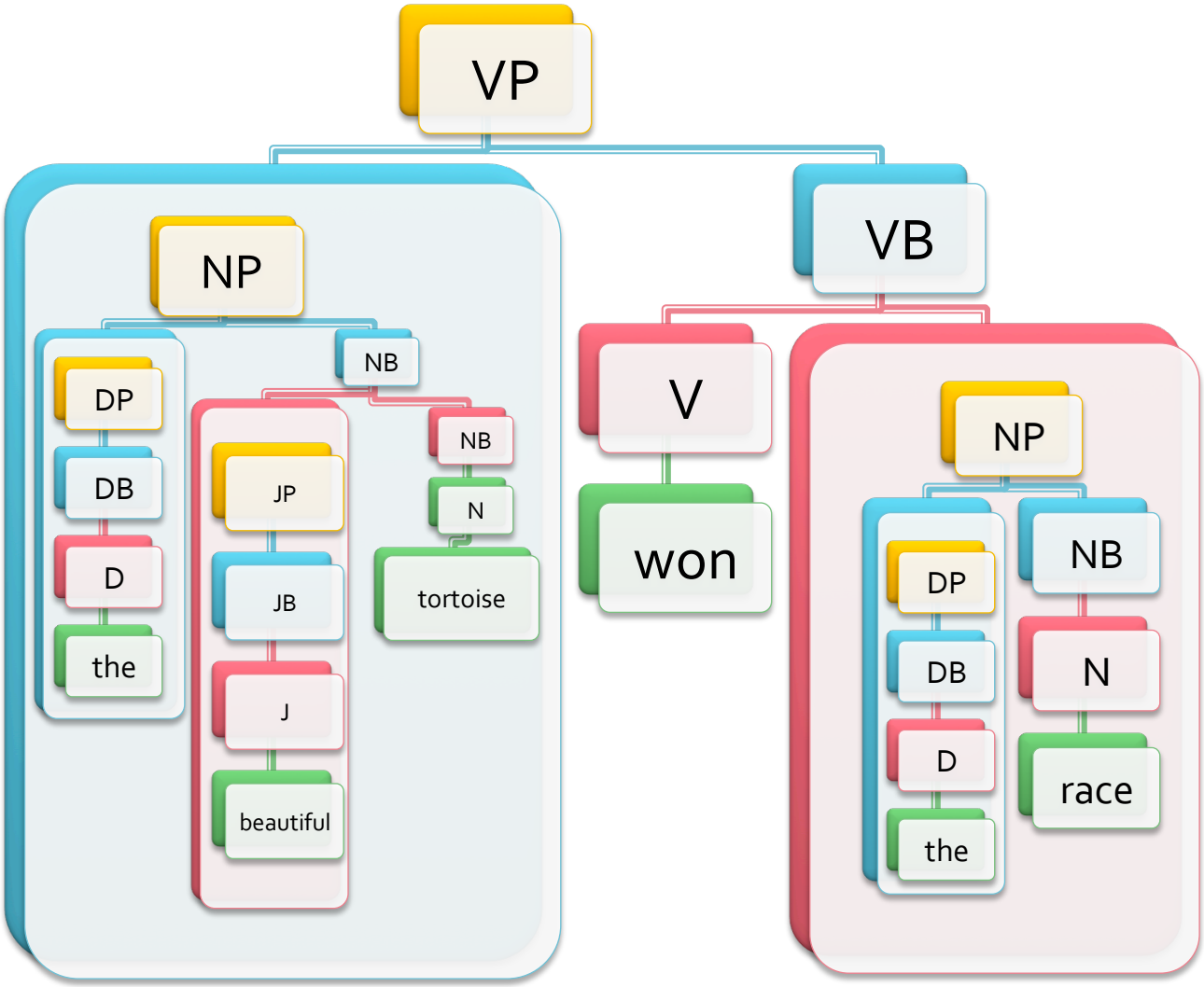
The tortoise



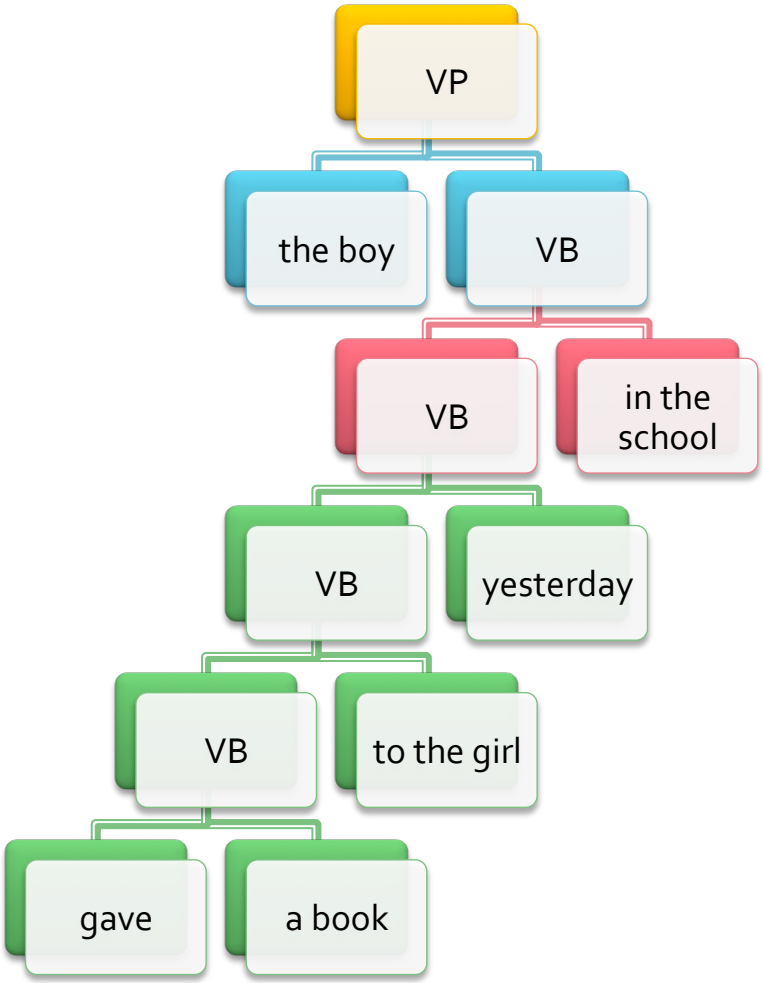
The beautiful tortoise



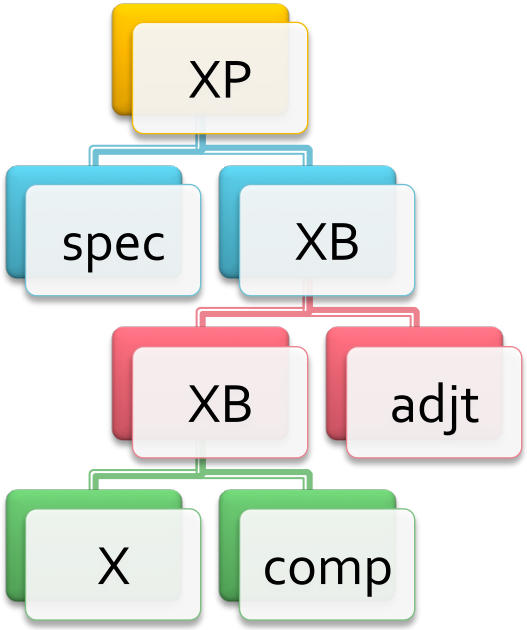
The beautiful tortoise won the race



The boy gave a book to the girl yesterday in the school



X-bar structure



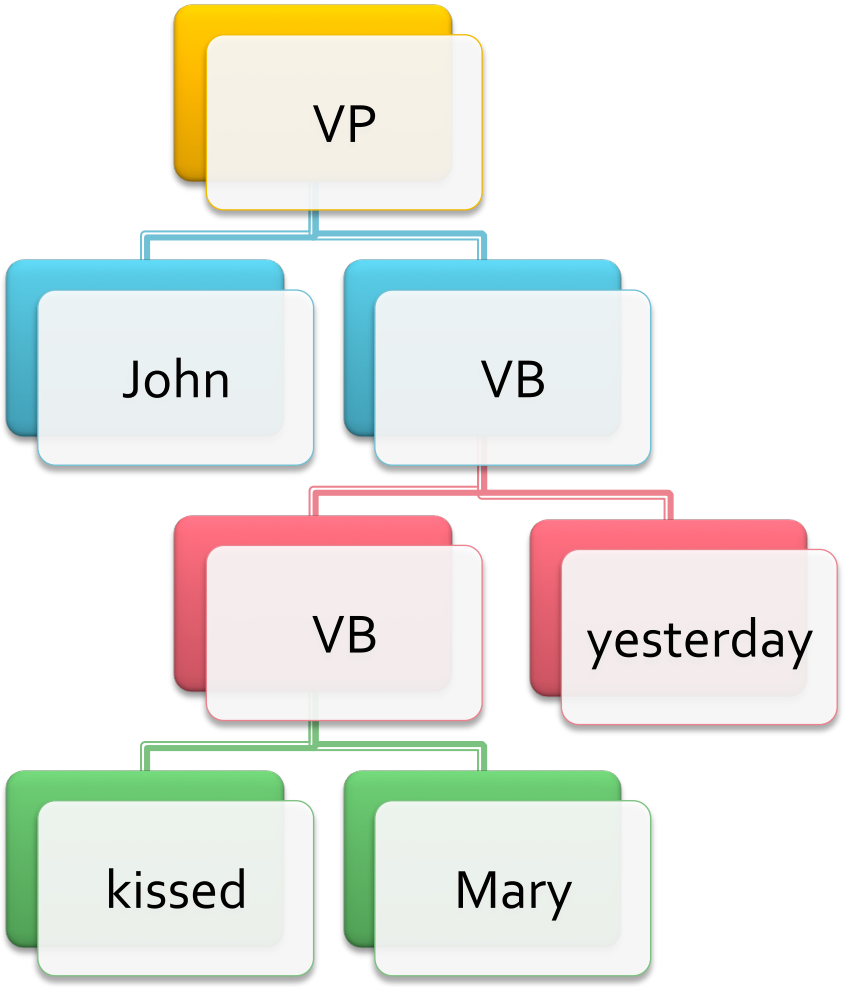
- Phrase-driven

XP(XB(XB(X;comp);adjt);spec)

- Head-driven

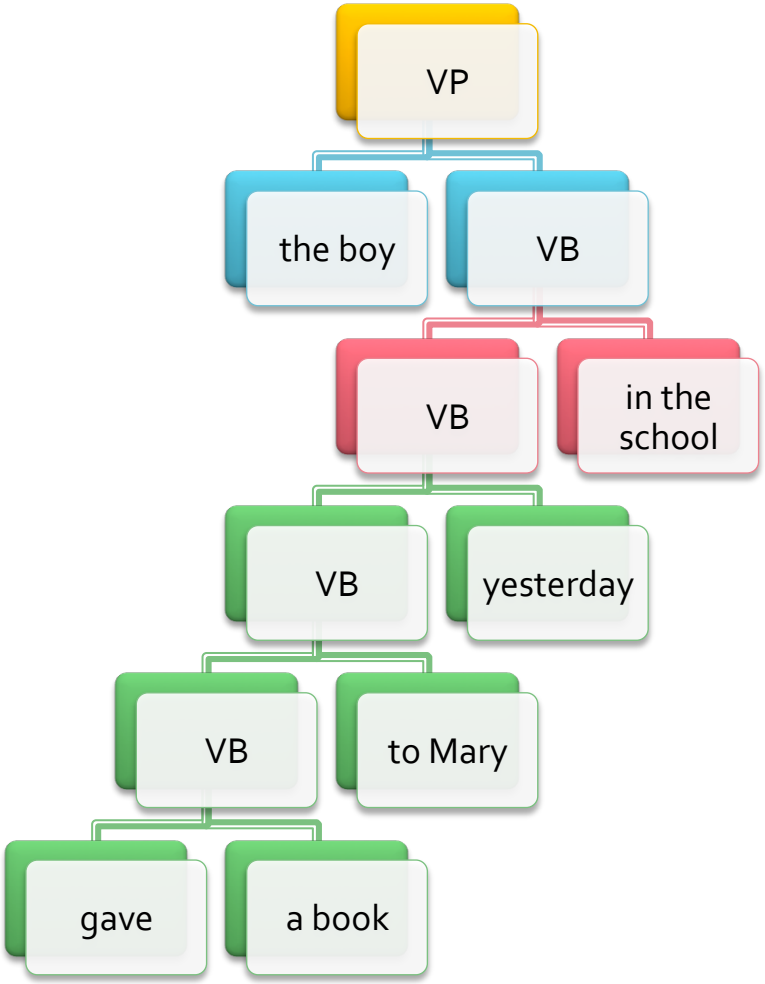
- XC(X;comp)
- XA(X;adjt)
- XS(X;spec)

John kissed Mary yesterday



VC(kissed;Mary)
VA(kissed;yesterday)
VS(kissed;John)

John gave a book to Mary yesterday in the school



VC(gave;a book)
VC(gave;to Mary)
VA(gave;yesterday)
VA(kissed;in the school)
VS(kissed;John)

Exercise #1

- Identify the syntactic roles in the sentence:
  - The Hare one day ridiculed the short feet and slow pace of the Tortoise

1. ??(hare;the)
2. ??(feet;short)
3. ??(slow;pace)
4. ??(pace;of the Tortoise)
5. ??(ridiculed;the hare)
6. ??(ridiculed;one day)
7. ??(ridiculed;the short feet and slow pace of the Tortoise)

Grammar

Subcategorization

Valence

AVALENT (VALENCE = 0)

Peter
Mary
Peter killed Mary
Peter went to NY
Mary is beautiful

VALENT (VALENCE > 0)

*killed Mary
*Peter killed
*Peter killed Mary with
*Peter went
*is beautiful

Subcategorization

VALENCY	EXAMPLES
0	book, computer, table New York, Greece beautiful, ugly, yesterday, happily
1	arrival (of John) (the) United States, (the) United Kingdom full (of students) with (Mary) differently (from them) (Mary) died (it) rains
2	between (Mary) and (John) (Peter) killed (Mary)
3	(Peter) gave (the book) (to Mary)

Subcategorization frames

- Y0 = AVALENT (no necessary argument)
  - book, computer, beautiful, …
- Y1 = IRREGULAR (very specific subcategorization, to be provided in the dictionary, through subcategorization rules)
- Y2-… REGULAR FRAMES, to be provided in the grammar (they will be available in the dictionary, through a drop-down menu)
  - intransitive verbs
  - direct transitive verbs
  - indirect transitive verbs that select the preposition X
  - indirect transitive verbs that select the preposition Y
  - …

S-rules (I)

SYNTACTIC ROLE(ARGUMENT);

- VS(NP);
  - requires a verb specifier which is a NP (die, run, dance)
- VS(NP,NOM);
  - requires a verb specifier which is a NP in the nominative case (die, run, dance)
- VS([it]);
  - requires a verb specifier which is equal to the headword [it] (rain, snow)
- NS([the]);
  - requires a noun specifier which is equal to the headword [the] (United States, United Kingdom)

S-rules (II)

- VS(NP)VC(NP);
  - requires a verb specifier (which is a NP) and verb complement (which is also a NP) (kill, know, contain)
- VS(NP,NOM)VC(NP,ACC);
  - requires a verb specifier (which is a NP in the nominative case) and a verb complement (which is a NP in the accusative case); (kill, know, contain)
- VS(NP,NOM)VC(NP,ACC)VC(NP,DAT);
  - requires a verb specifier (which is a NP in the nominative case), a first verb complement (which is a NP in the accusative case) and a second verb complement (which is a NP in the dative case)

S-rules (III)

- VS(NP)VC(PH([on]));
  - requires a verb specifier (which is a NP) and a verb complement (which is a prepositional phrase headed by the preposition [on])
    - depend

Exercise #2

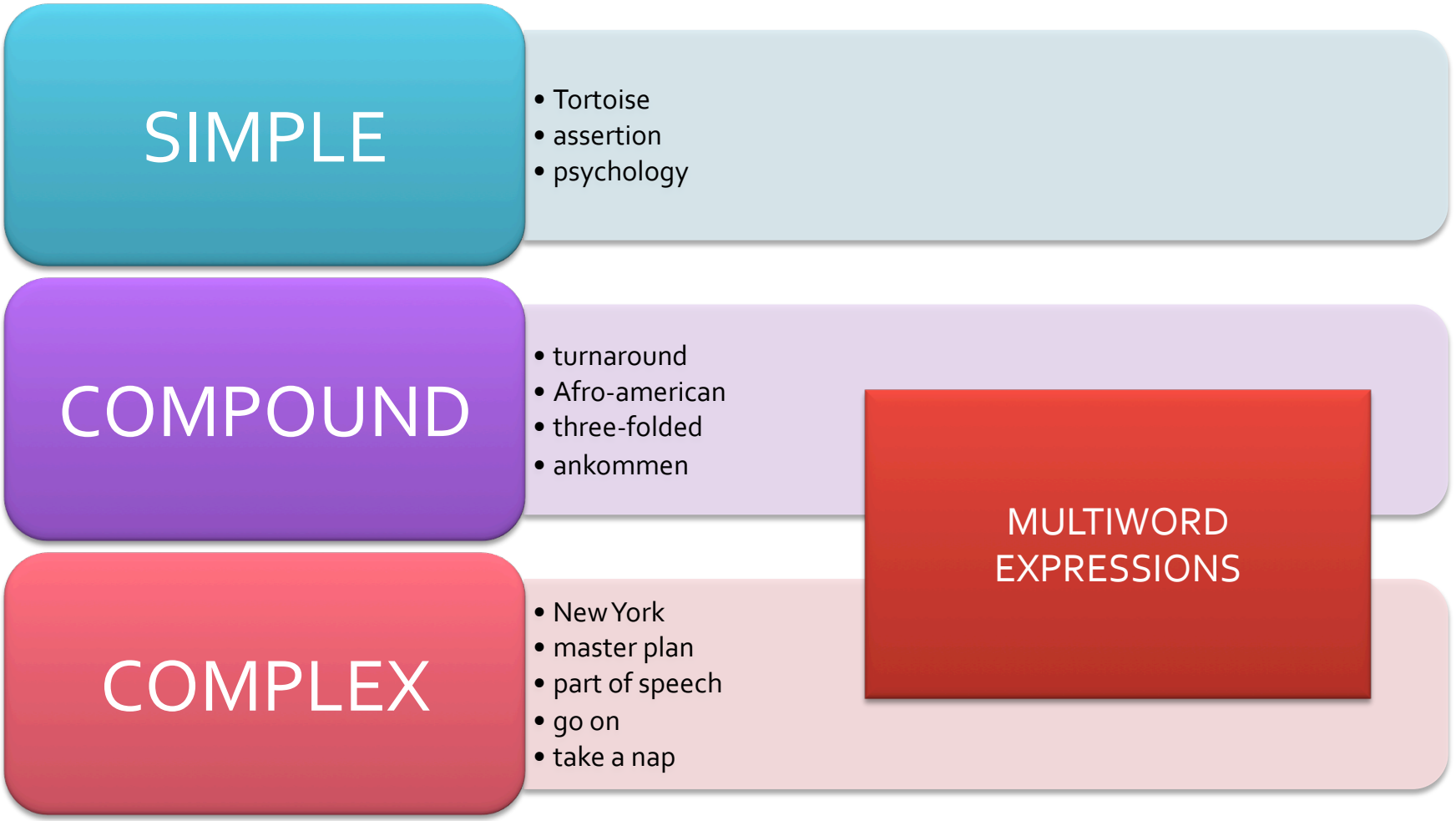
- Describe the subcategorization frames of the following English words
  1. tortoise
  2. comparison
  3. slow
  4. capable
  5. to kiss
  6. to win
  7. to reply
  8. to persuade

Exercise #3

- Create, in the UNLsandbox, the subcategorization frames necessary to describe the dependencies of the VALENT words of the dictionary generated by the project AESOP-3

Composition

Words


SIMPLE
• Tortoise
• assertion
• psychology
COMPOUND
• turnaround
• Afro-american
• three-folded
• ankommen
COMPLEX
• New York
• master plan
• part of speech
• go on
• take a nap
MULTIWORD EXPRESSIONS

Multiword expressions

INVARIANT	REGULAR	IRREGULAR (infixation)
• Afro-american • New York • three-folded	• turnaround • master plan	• part of speech • go on • take a nap • ankommen

C-Rules

<SYNTACTIC ROLE>(<ADDED>,<FEATURES>)

- <SYNTACTIC ROLE>
- <ADDED>
- <FEATURES>
  - [LEMMA] AND LEX (A,J,N,V,C,P,D) OR
  - "PHRASE" AND XP (AP,JP,NP,VP,CP,PP,DP(
  - PARADIGM
  - DISTRIBUTION (AFT, BEF, IAFT, IABEF, …)
  - ADJACENCY (AJ0, AJ1, AJ2, …)

Examples (I)

- master plan
  - M2: SNG=o>""; PLR:=o>"s";
  - master plan; master plans
  - BASE FORM = LEMMA
- part of speech
  - M2: SNG=o>"";PLR:=o>"s";
  - part of speech; part of speechs
  - BASE FORM ≠ LEMMA
    - LEMMA = part of speech
    - BASE FORM = part (to be associated to M2)
    - COMPOSITION RULE = NA("of speech",PP);

Examples (II)

- non-governmental organization
  - M2: SNG=o>""; PLR:=o>"s";
  - non-governmental organization, non-governmental organizations
  - BASE FORM = LEMMA
- take into account
  - M5: INF:=o>""; PAS:=3>"ook"; PTP:= o>"n":
  - take into account, take int accoook, take into acountn
  - BASE FORM ≠ LEMMA
    - LEMMA = take into account
    - BASE FORM = take (to be associated to M2)
    - COMPOSITION RULE = VA("into account",PP);

Examples (III)

- make sense
  - M125: INF:=o>""; PAS:=2>"de"; PTP:=2>"de":
  - make sense, make sende, make sende
  - BASE FORM ≠ LEMMA
    - LEMMA = make sense
    - BASE FORM = make (to be associated to M125)
    - COMPOSITION RULE = VC([sense],N,Mo);
- coffee shop
  - M2: SNG=o>""; PLR:=o>"s";
  - coffee shop, coffee shop
  - BASE FORM = LEMMA

Examples (IV)

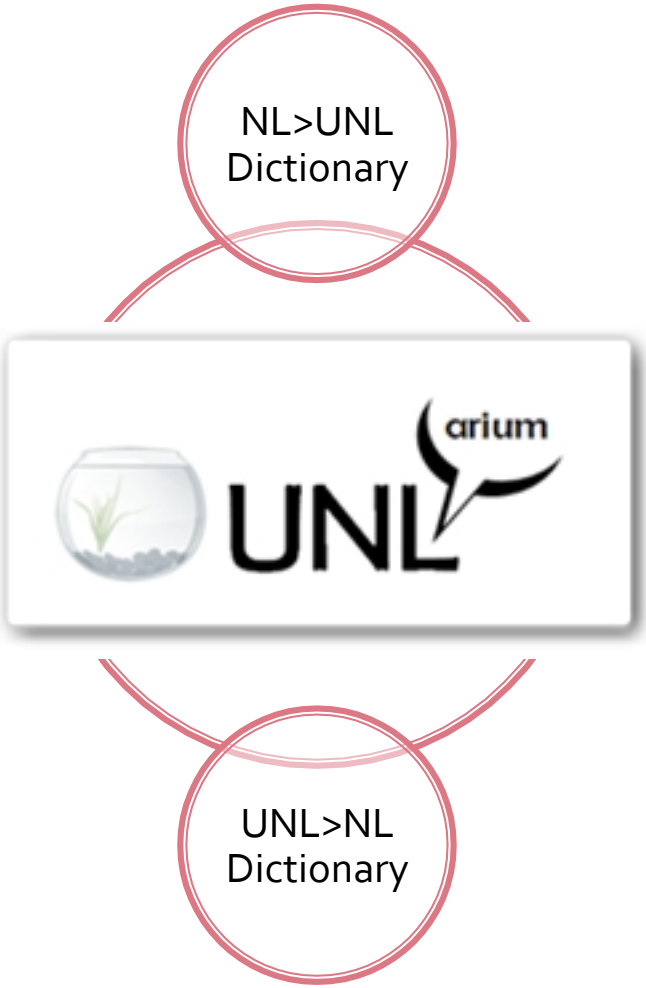
- lingua franca
  - M2: NOM&SNG:=o>"";NOM&PLR:=o>"e";…
  - lingua franca, lingua francae, …
  - BASE FORM ≠ LEMMA
    - LEMMA = lingua franca
    - BASE FORM = lingua (to be associated to M2)
    - COMPOSITION RULE = NA([franca],J,M123);
- ankommen
  - M3: 3PS&PRS&IND = "en">"t";
  - ankommt
  - BASE FORM ≠ LEMMA
    - LEMMA = ankommen
    - BASE FORM = kommen (to be associated to M3)
    - COMPOSITION RULE = VH([an],F);

Exercise #3

- Create the composition rules for the following English verbs:
  1. to go on
  2. to take a nap
  3. to do the cleaning
  4. to give a presentation
  5. to make a mistake
  6. to have a cigarette
  7. to take advantage of

Dictionaries #2

Dictionaries



AESOP-4

- Create the analysis dictionary corresponding to the project AESOP-4, available in the UNLsandbox.
  - Link the entry to as many UW's as possible
  - Provide the corresponding composition rules, whenever necessary
  - Link them to the corresponding paradigms (and create them, if they do not exist, or define the inflectional rules, if they are irregular)
  - Link them to the corresponding subcategorization frames (and create them, if they do not exist)