Parts of Speech
Part 2

ICS 482 Natural Language Processing

Lecture 10:

Husni Al-Muhtaseb

بسم الله الرحمن الرحيم
ICS 482 Natural Language Processing

Lecture 10: Parts of Speech
Part 2

Husni Al-Muhtaseb

NLP Credits and Acknowledgment

These slides were adapted from presentations of the Authors of the book

SPEECH and LANGUAGE PROCESSING:
An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition

and some modifications from presentations found in the WEB by several scholars including the following

NLP Credits and Acknowledgment

If your name is missing please contact me
muhtaseb
At
Kfupm.
Edu.
sa

NLP Credits and Acknowledgment

Husni Al-Muhtaseb
James Martin
Jim Martin
Dan Jurafsky
Sandiway Fong
Song young in
Paula Matuszek
Mary-Angela Papalaskari
Dick Crouch
Tracy Kin
L. Venkata Subramaniam
Martin Volk
Bruce R. Maxim
Jan Hajič
Srinath Srinivasa
Simeon Ntafos
Paolo Pirjanian
Ricardo Vilalta
Tom Lenaerts

Heshaam Feili
Björn Gambäck
Christian Korthals
Thomas G. Dietterich
Devika Subramanian
Duminda Wijesekera
Lee McCluskey
David J. Kriegman
Kathleen McKeown
Michael J. Ciaraldi
David Finkel
Min-Yen Kan
Andreas Geyer-Schulz
Franz J. Kurfess
Tim Finin
Nadjet Bouayad
Kathy McCoy
Hans Uszkoreit
Azadeh Maghsoodi

Khurshid Ahmad
Staffan Larsson
Robert Wilensky
Feiyu Xu
Jakub Piskorski
Rohini Srihari
Mark Sanderson
Andrew Elks
Marc Davis
Ray Larson
Jimmy Lin
Marti Hearst
Andrew McCallum
Nick Kushmerick
Mark Craven
Chia-Hui Chang
Diana Maynard
James Allan

Martha Palmer
julia hirschberg
Elaine Rich
Christof Monz
Bonnie J. Dorr
Nizar Habash
Massimo Poesio
David Goss-Grubbs
Thomas K Harris
John Hutchins
Alexandros Potamianos
Mike Rosner
Latifa Al-Sulaiti
Giorgio Satta
Jerry R. Hobbs
Christopher Manning
Hinrich Schütze
Alexander Gelbukh
Gina-Anne Levow
Guitao Gao
Qing Ma
Zeynep Altan

Previous Lectures

- Pre-start questionnaire
- Introduction and Phases of an NLP system
- NLP Applications - Chatting with Alice
- Finite State Automata & Regular Expressions & languages
- Deterministic & Non-deterministic FSAs
- Morphology: Inflectional & Derivational
- Parsing and Finite State Transducers
- Stemming & Porter Stemmer
- 20 Minute Quiz
- Statistical NLP – Language Modeling
- N-Grams
- Smoothing and N-Gram: Add-one & Witten-Bell
- Return Quiz 1
- Parts of Speech

Today's Lecture

- Continue with Parts of Speech

- Arabic Parts of Speech

Parts of Speech

Start with eight basic categories

- Noun   اسم
- Verb فعل
- preposition حرف جر
- Pronoun   ضمير

- adjective صفة
- adverb ظرف
- Article أداة
- Conjunction حرف عطف

These categories are based on morphological and distributional properties (not semantics)
Some cases are easy, others are not

Parts of Speech

- Closed classes
  - Prepositions: on, under, over, near, by, at, from, to, with, etc.
  - Determiners: a, an, the, etc.
  - Pronouns: she, who, I, others, etc.
  - Conjunctions: and, but, or, as, if, when, etc.
  - Auxiliary verbs: can, may, should, are, etc.
  - Particles: up, down, on, off, in, out, at, by, etc.
- Open classes:
  - Nouns:
  - Verbs:
  - Adjectives:
  - Adverbs:

Sets of Parts of Speech: Tagsets

- There are various standard tagsets to choose from; some have a lot more tags than others

- The choice of tagset is based on the application

- Accurate tagging can be done with even large tagsets

Some of the known Tagsets (English)

- Brown corpus: 87 tags

- Penn Treebank: 45 tags

- Lancaster UCREL C5: 61 tags

- Lancaster C7: 145 tags

Some of Penn Treebank tags

NN	noun	**JJ**	adjective
NNP	proper noun	**CC**	coord conj
DT	determiner	**CD**	cardinal number
IN	preposition	**PRP**	personal pronoun
VB	verb	**RB**	adverb
-R	comparative		
-S	superlative or plural		
-$	possessive		

Verb inflection tags

VBP	base present	*take*
VB	infinitive	*take*
VBD	past	*took*
VBG	present participle	*taking*
VBN	past participle	*taken*
VBZ	present 3sg	*takes*
MD	modal	*can, would*

The entire Penn Treebank tagset

Tag	Description	Example	Tag	Description	Example
CC	Coordin. Conjunction	*and, but, or*	SYM	Symbol	*+,%, &*
CD	Cardinal number	*one, two, three*	TO	"to"	*to*
DT	Determiner	*a, the*	UH	Interjection	*ah, oops*
EX	Existential 'there'	*there*	VB	Verb, base form	*eat*
FW	Foreign word	*mea culpa*	VBD	Verb, past tense	*ate*
IN	Preposition/sub-conj	*of, in, by*	VBG	Verb, gerund	*eating*
JJ	Adjective	*yellow*	VBN	Verb, past participle	*eaten*
JJR	Adj., comparative	*bigger*	VBP	Verb, non-3sg pres	*eat*
JJS	Adj., superlative	*wildest*	VBZ	Verb, 3sg pres	*eats*
LS	List item marker	*1, 2, One*	WDT	Wh-determiner	*which, that*
MD	Modal	*can, should*	WP	Wh-pronoun	*what, who*
NN	Noun, sing. or mass	*llama*	WP$	Possessive wh-	*whose*
NNS	Noun, plural	*llamas*	WRB	Wh-adverb	*how, where*
NNP	Proper noun, singular	*IBM*	$	Dollar sign	*$*
NNPS	Proper noun, plural	*Carolinas*	#	Pound sign	*#*
PDT	Predeterminer	*all, both*	"	Left quote	*(' or ")*
POS	Possessive ending	*'s*	"	Right quote	*(' or ")*
PP	Personal pronoun	*I, you, he*	(	Left parenthesis	*([, (, {, <)*
PP$	Possessive pronoun	*your, one's*	)	Right parenthesis	*( ],), }, >)*
RB	Adverb	*quickly, never*	,	Comma	*,*
RBR	Adverb, comparative	*faster*	.	Sentence-final punc	*(. ! ?)*
RBS	Adverb, superlative	*fastest*	:	Mid-sentence punc	*(: ; ... – -)*
RP	Particle	*up, off*			

UCREL C5

| Tag	Description	Example
PNX	reflexive pronoun	*itself, ourselves*
POS	possessive *'s* or *'*	
PRF	the preposition *of*	
PRP	preposition (except *of*)	*for; above, to*
PUL	punctuation – left bracket	(or [
PUN	punctuation – general mark	. ! , : ; - ? ...
PUQ	punctuation – quotation mark	' ' "
PUR	punctuation – right bracket	) or]
TO0	infinitive marker *to*	
UNC	unclassified items (not English)	
VBB	base forms of *be* (except infinitive)	*am, are*
VBD	past form of *be*	*was, were*
VBG	-ing form of *be*	*being*
VBI	infinitive of *be*	
VBN	past participle of *be*	*been*
VBZ	-s form of *be*	*is, 's*
VDB	base form of *do* (except infinitive)	*does*
VDD	past form of *do*	*did*
VDG	-ing form of *do*	*doing*
VDI	infinitive of *do*	*to do*
VDN	past participle of *do*	*done*
VDZ	-s form of *do*	*does*
VHB	base form of *have* (except infinitive)	*have*
VHD	past tense form of *have*	*had, 'd*
VHG	-ing form of *have*	*having*
VHI	infinitive of *have*	
VHN	past participle of *have*	*had*
VHZ	-s form of *have*	*has, 's*
VM0	modal auxiliary verb	*can, could, will, 'll*
VVB	base form of lexical verb (except infin.)	*take, live*
VVD	past tense form of lexical verb	*took, lived*
VVG	-ing form of lexical verb	*taking, living*
VVI	infinitive of lexical verb	*take, live*
VVN	past participle form of lex. verb	*taken, lived*
VVZ	-s form of lexical verb	*takes, lives*
XX0	the negative *not* or *n't*	
ZZ0	alphabetical symbol	*A, B, c, d*

Tagging

- Part of speech tagging is the process of assigning parts of speech to each word in a sentence… Assume we have
  - A tagset
  - A dictionary that gives you the possible set of tags for each entry
  - A text to be tagged
  - A reason?

POS Tagging: Definition

- The process of assigning a part-of-speech or lexical class marker to each word in a corpus:

WORDS

TAGS

the
driver
put
the
keys
on
the
table

N
V
P
DET
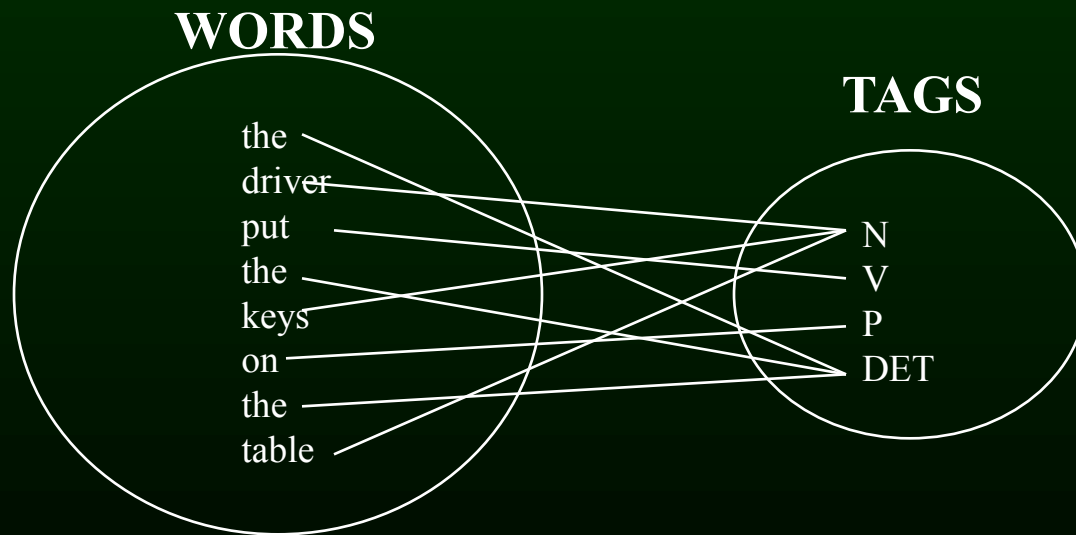
Tag Ambiguity (updated)

	87-tagset	45-tagset
Unambiguous (1 tag)	44,019	38,857
Ambiguous (2-7 tags)	5,490	8,844
2 tags	4,967	6,731
3 tags	411	1621
4 tags	91	357
5 tags	17	90
6 tags	2 (well, beat)	32
7 tags	2 (still, down)	6 (well, set, round, open, fit, down)
8 tags		4 ('s, half, back, a)
9 tags		3 (that, more, in)

- Most words are unambiguous
- Many of the most common English words are ambiguous

Tagging: Three Methods

- Rules

- Probabilities (Stochastic)

- Transformation-Based: Sort of both

Rule-based Tagging

- Use dictionary (lexicon) to assign each word a list of potential POS

- Use large lists of hand-written disambiguation rules to identify a single POS for each word.

- Example of rules: NP $\rightarrow$ Det (Adj*) N
  - For example: *the clever student*

Probabilities: Tagging with lexical frequencies

- Sami is expected to race tomorrow.
- Sami/NNP is/VBZ expected/VBN to/TO **race**/VB tomorrow/NN
- People continue to inquire the reason for the race for outer space.
- People/NNS continue/VBP to/TO inquire/VB the/DT reason/NN for/IN the/DT **race**/NN for/IN outer/JJ space/NN
- Problem: assign a tag to race given its lexical frequency
- Solution: we choose the tag that has the greater
  - P(race|VB)
  - P(race|NN)
- Actual estimate from the Switchboard corpus:
  - P(race|NN) = .00041
  - P(race|VB) = .00003

Transformation-based: The Brill Tagger

- An example of Transformation-based Learning
- Very popular (freely available, works fairly well)
- A SUPERVISED method: requires a tagged corpus
- Basic idea: do a quick job first (using frequency), then revise it using contextual rules
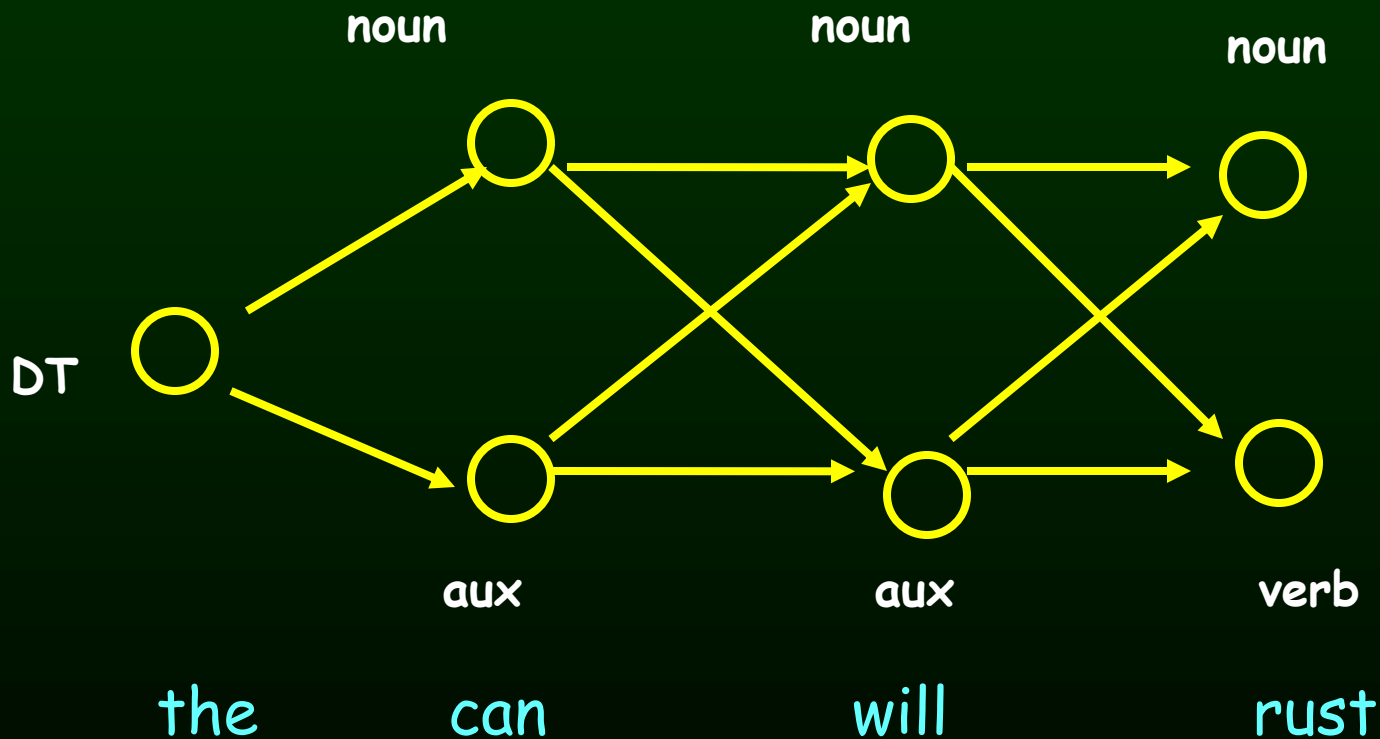
An example

- Examples:
  - It is expected to race tomorrow.
  - The race for outer space.
- Tagging algorithm:
  1. Tag all uses of "race" as NN (most likely tag in the Brown corpus)
     - It is expected to race/NN tomorrow
     - the race/NN for outer space
  2. Use a transformation rule to replace the tag NN with VB for all uses of "race" preceded by the tag TO:
     - It is expected to race/VB tomorrow
     - the race/NN for outer space

Stochastic (Probabilities)

- Simple approach
  - Disambiguate words based on the probability that a word occurs with a particular tag

- N-gram approach
  - The best tag for given words is determined by the probability that it occurs with the n previous tags

- Viterbi Algorithm
  - Trim the search for the most probable tag using the best N Maximum Likelihood Estimates (N is the number of tags of the following word)

- Hidden Markov Model combines the above two approaches

Viterbi Maximum Likelihood Estimates

Want the most likely path through this graph.

Viterbi Maximum Likelihood Estimates
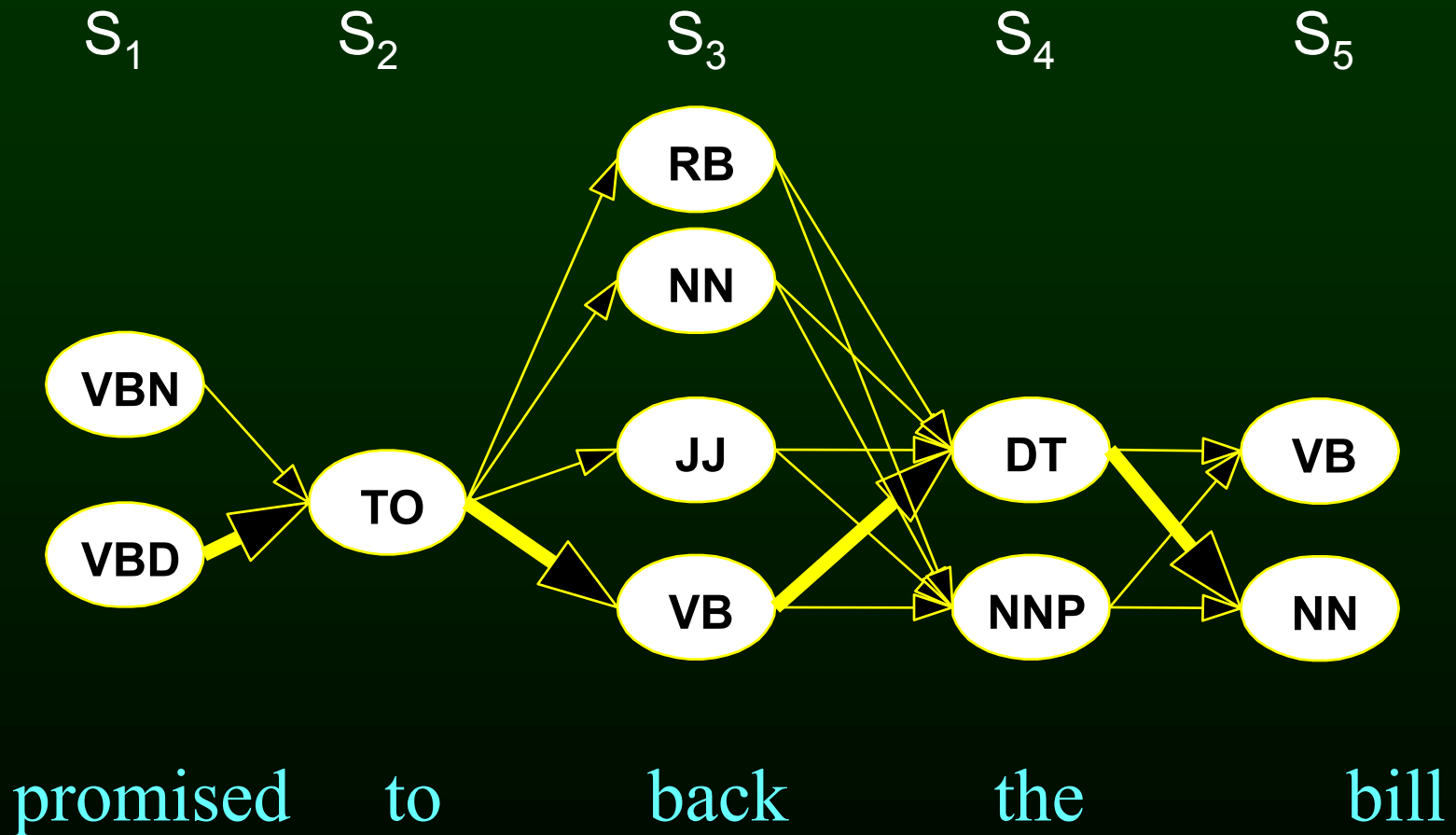
S_1 S_2 S_3 S_4 S_5

RB

NN

VBN

JJ   DT   VB

TO

VBD   VB   NNP   NN

promised   to   back   the   bill

Viterbi Maximum Likelihood Estimates

- We want the best set of tags for a sequence of words (a sentence)

- W is a sequence of words
- W= $w_1 w_2 w_3 .. w_n$
- T is a sequence of tags
- T= $t_1 t_2 t_3 .. t_n$

$$\arg\max P(T \mid W) = \frac{P(W \mid T)P(T)}{P(W)}$$

- P(w) is common

Viterbi Maximum Likelihood Estimates

- We want the best set of tags for a sequence of words (a sentence)

- W is a sequence of words
- W= $w_1 w_2 w_3 .. w_n$
- T is a sequence of tags
- T= $t_1 t_2 t_3 .. t_n$

$$\arg \max P(T \mid W) = P(W \mid T) P(T)$$

- P(w) is common

Stochastic POS Tagging: Example

1) Sami is expected to **race** tomorrow.

2) People continue to inquire the reason for the **race** for outer space.

Stochastic POS Tagging: Example

Example: suppose w_i = race, a verb (VB) or a noun (NN)?

Assume that other mechanism has already done the best tagging to the surrounding words, leaving only **race** untagged

1) Sami/NNP is/VBZ expected/VBN to/TO **race/?** tomorrow/NN

2) People/NNS continue/VBP to/TO inquire/VB the/DT reason/NN for/IN the/DT **race/?** For/IN outer/JJ space/NN

$$t_i = \arg\max{}_j P(t_j \mid t_{i-1}) P(w_i \mid t_j)$$

Bigram

Simplify the problem:
to/To   race/???
the/DT   race/???

P(VB|TO) P(race | VB)

P(NN|TO) P(race | NN)

Where is the data?

Look at the Brown and Switchboard corpora

$P(NN \mid TO) = 0.021$
$P(VB \mid TO) = 0.34$

If we are expecting a verb, how likely it would be "race"

$P(\text{race} \mid NN) = 0.00041$
$P(\text{race} \mid VB) = 0.00003$

Finally:

$P(NN \mid TO) \, P(\text{race} \mid NN) = 0.000007$
$P(VB \mid TO) \, P(\text{race} \mid VB) = 0.00001$

Example: Bigram of Tags from a Corpus

Cat	# at i	Pair	# at i, i+1	Bigram	Estimate
0	300	0, ART	213	Prob(ART|0)	0.71
0	300	0, N	87	Prob(N|0)	0.29
ART	558	ART, N	558	Prob(N | ART)	1
N	833	N, V	358	Prob(V | N)	0.43
N	833	N, N	108	Prob(N | N)	0.13
N	833	N, P	366	Prob(P | N)	0.44
V	300	V, N	75	Porb(N | V)	0.35
V	300	V, ART	194	Prob(ART | V)	0.65
P	307	P, ART	226	Prob (ART | P)	0.74
P	307	P, N	81	Prob (N | P)	0.26

A Markov Chain
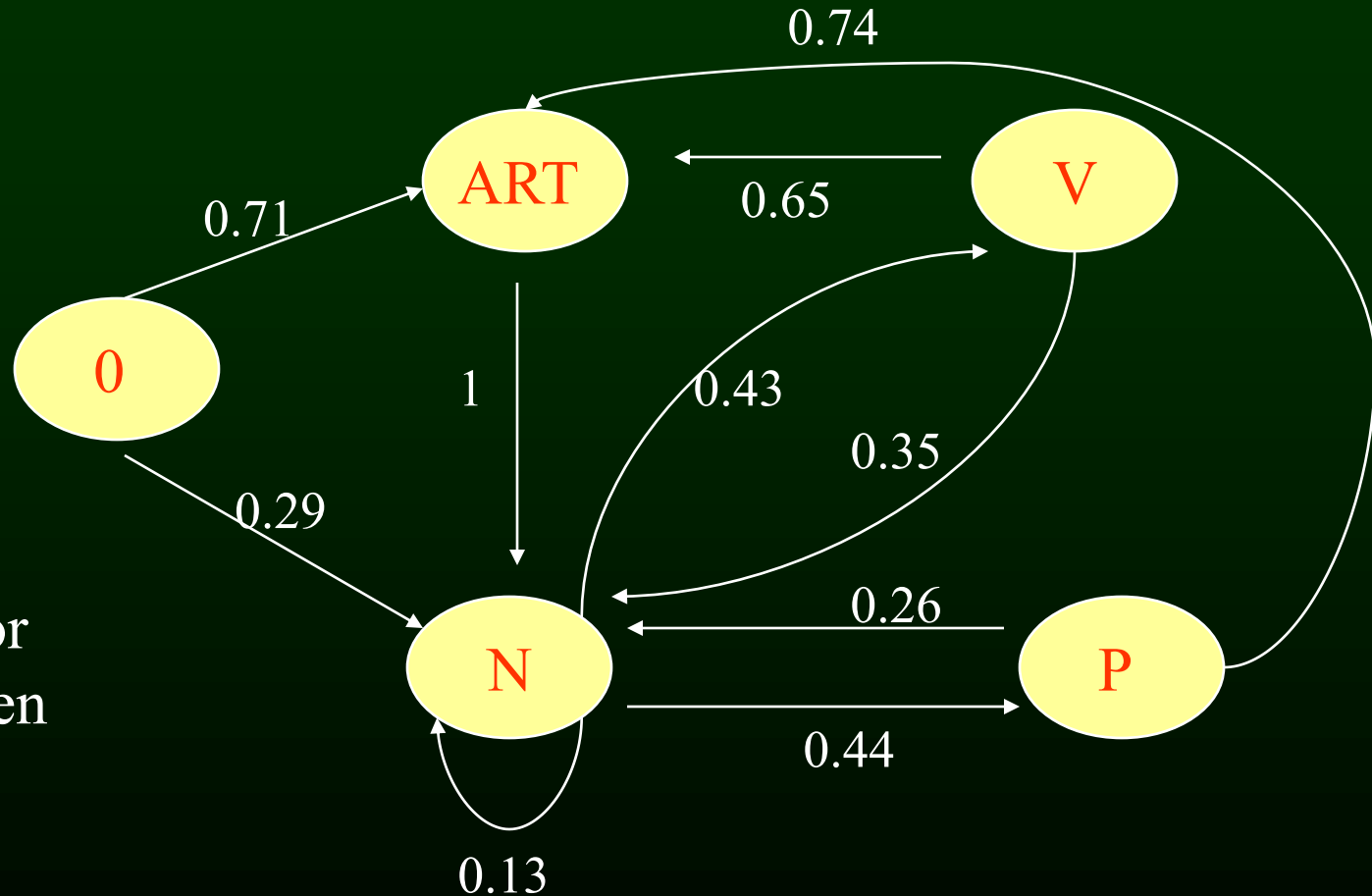
assume
0.0001 for
any unseen
bigram

Word Counts

	N	V	ART	P	Total
flies	21	23	0	0	44
fruit	49	5	1	0	55
like	10	30	0	21	61
a	1	0	201	0	202
the	1	0	300	2	303
flower	53	15	0	0	68
flowers	42	16	0	0	58
birds	64	1	0	0	65
others	592	210	56	284	1142
Total	833	300	558	307	1998

Computing Probabilities using previous Tables

P(the | ART) = 300/558 = 0.54
P(flies | N) = 0.025
P(flies | V) = 0.076
P(like | V) = 0.1
P(like | P) = 0.068
P(like | N) = 0.012

P(a | ART) = 0.360
P(a | N) = 0.001
P(flower | N) = 0.063
P(flower | V) = 0.05
P(birds | N) = 0.076

Viterbi Algorithm - Example

assume 0.0001 for any unseen bigram
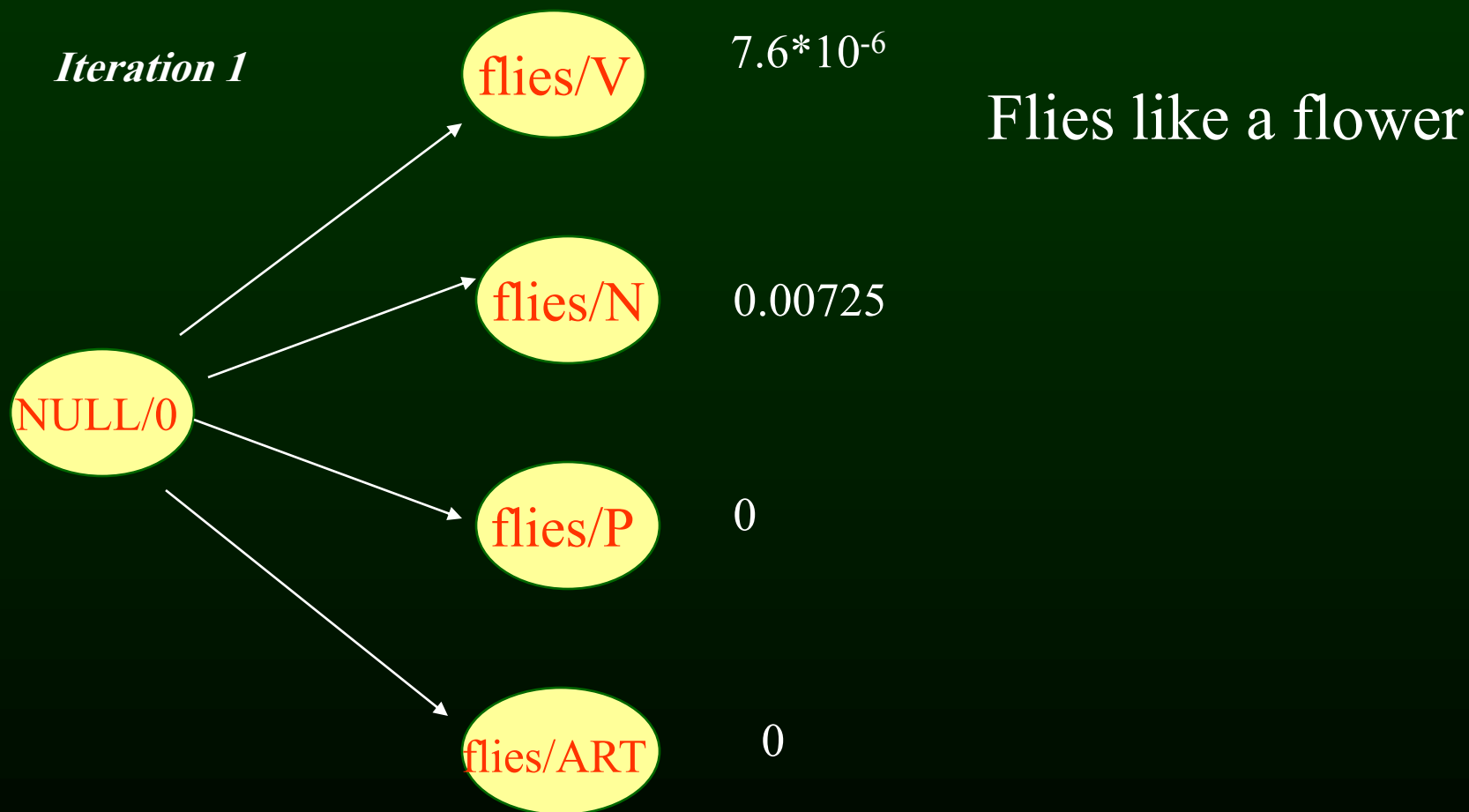
Iteration 1

flies/V — $7.6*10^{-6}$

Flies like a flower

flies/N — 0.00725

NULL/0

flies/P — 0

flies/ART — 0

Viterbi Algorithm - Example

Iteration 2

flies/V → like/V

flies/N

like/N

like/P

like/ART

Flies like a flower

Viterbi Algorithm - Example

Iteration 2

flies/V → like/V   0.00031

flies/N   like/N $1.3*10^{-5}$
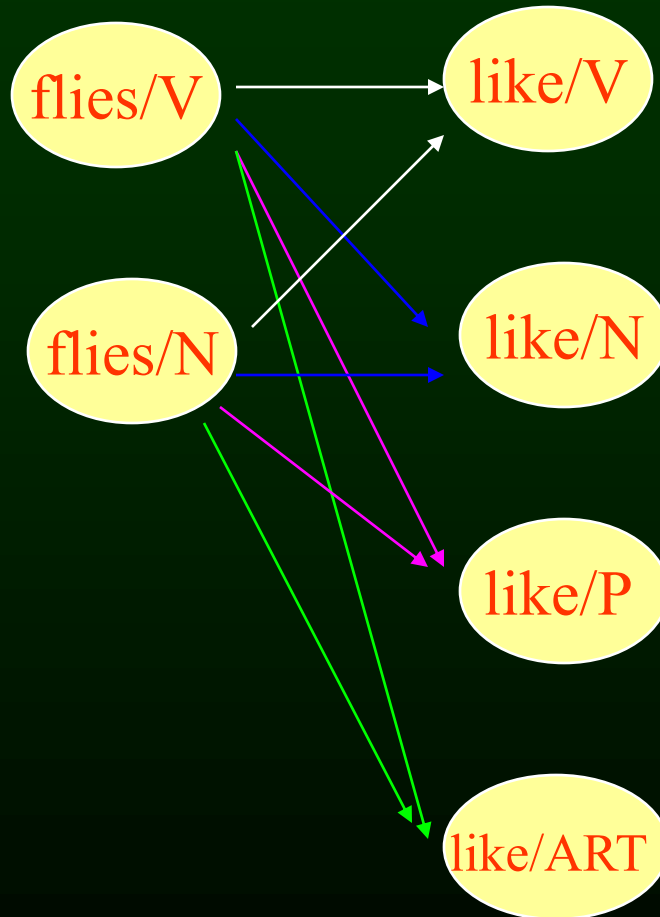
like/P   0.00022

Flies like a flower

like/ART   0
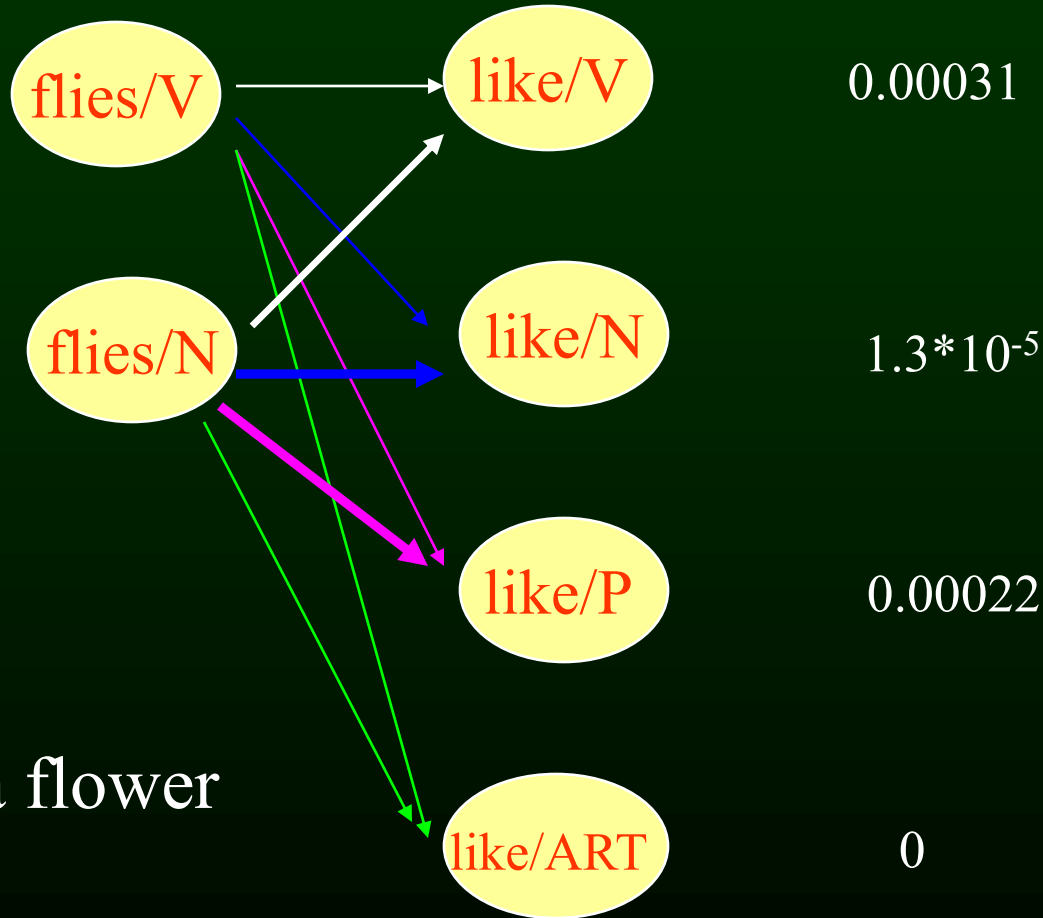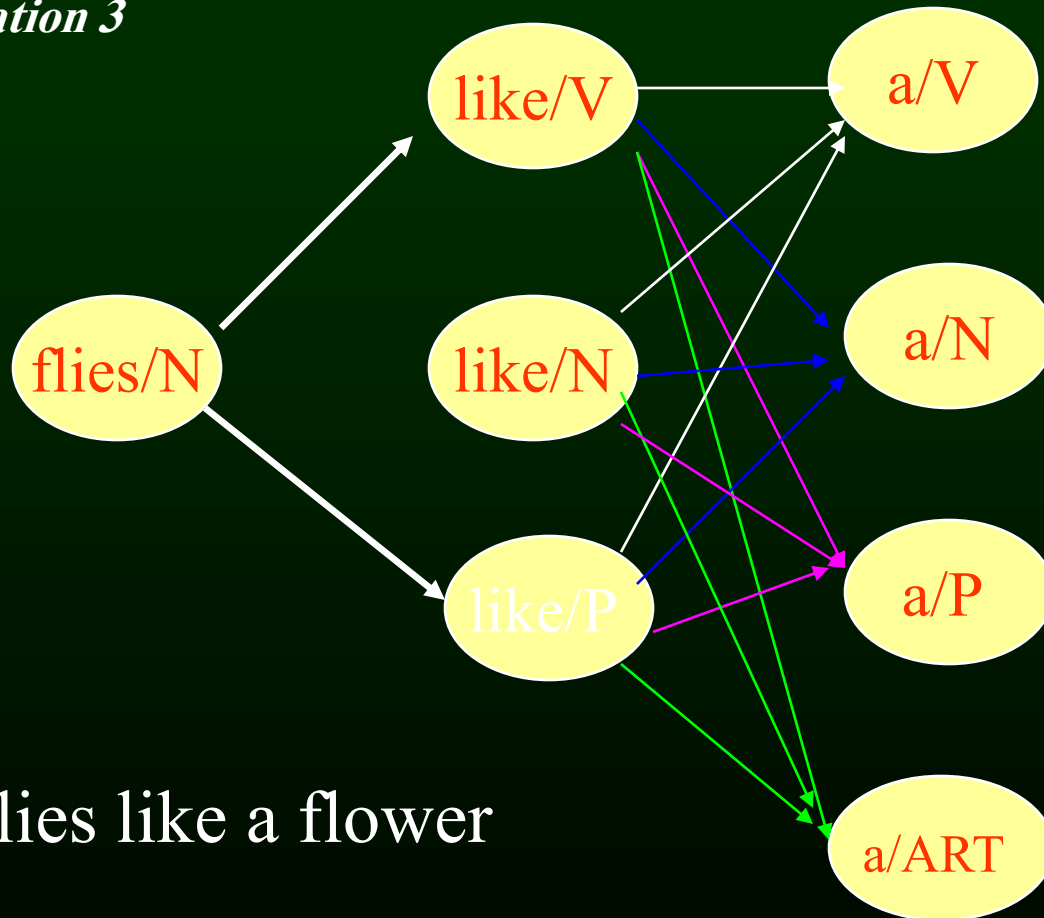
Viterbi Algorithm - Example

Iteration 3

Flies like a flower
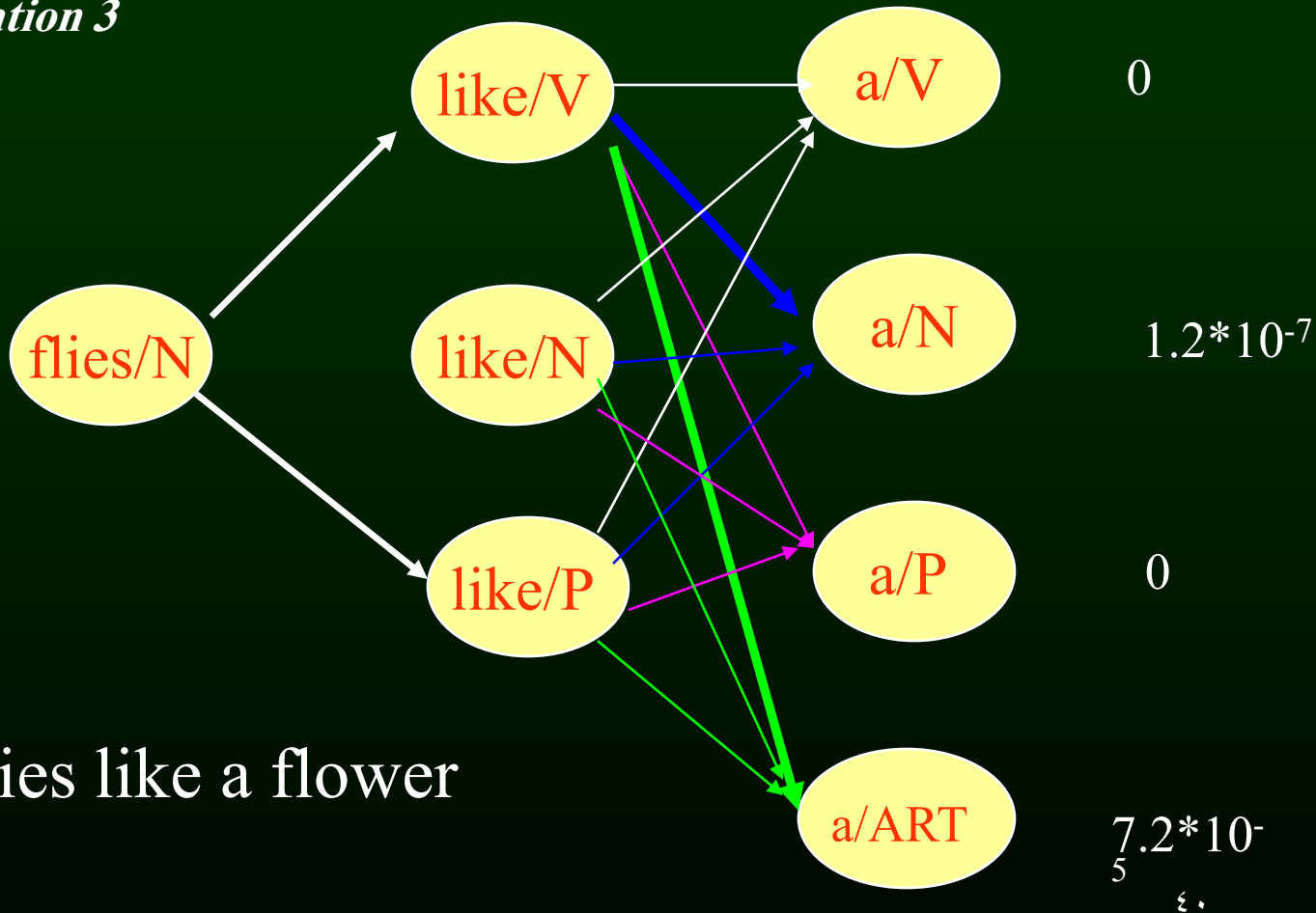
Viterbi Algorithm - Example

Iteration 3

0

$1.2*10^{-7}$

0

$7.2*10^{-5}$

Flies like a flower

Viterbi Algorithm - Example

Iteration 4

like/V

flower/V

flies/N

a/N

flower/N

like/P

flower/P

a/ART

flower/ART

Flies like a flower

Viterbi Algorithm - Example

Iteration 4

flower/V $2.6*10^{-9}$

like/V

flower/N $4.3*10^{-6}$

flies/N   a/N

flower/P   0

like/P

a/ART   flower/ART   0

Flies like a flower

٤٢
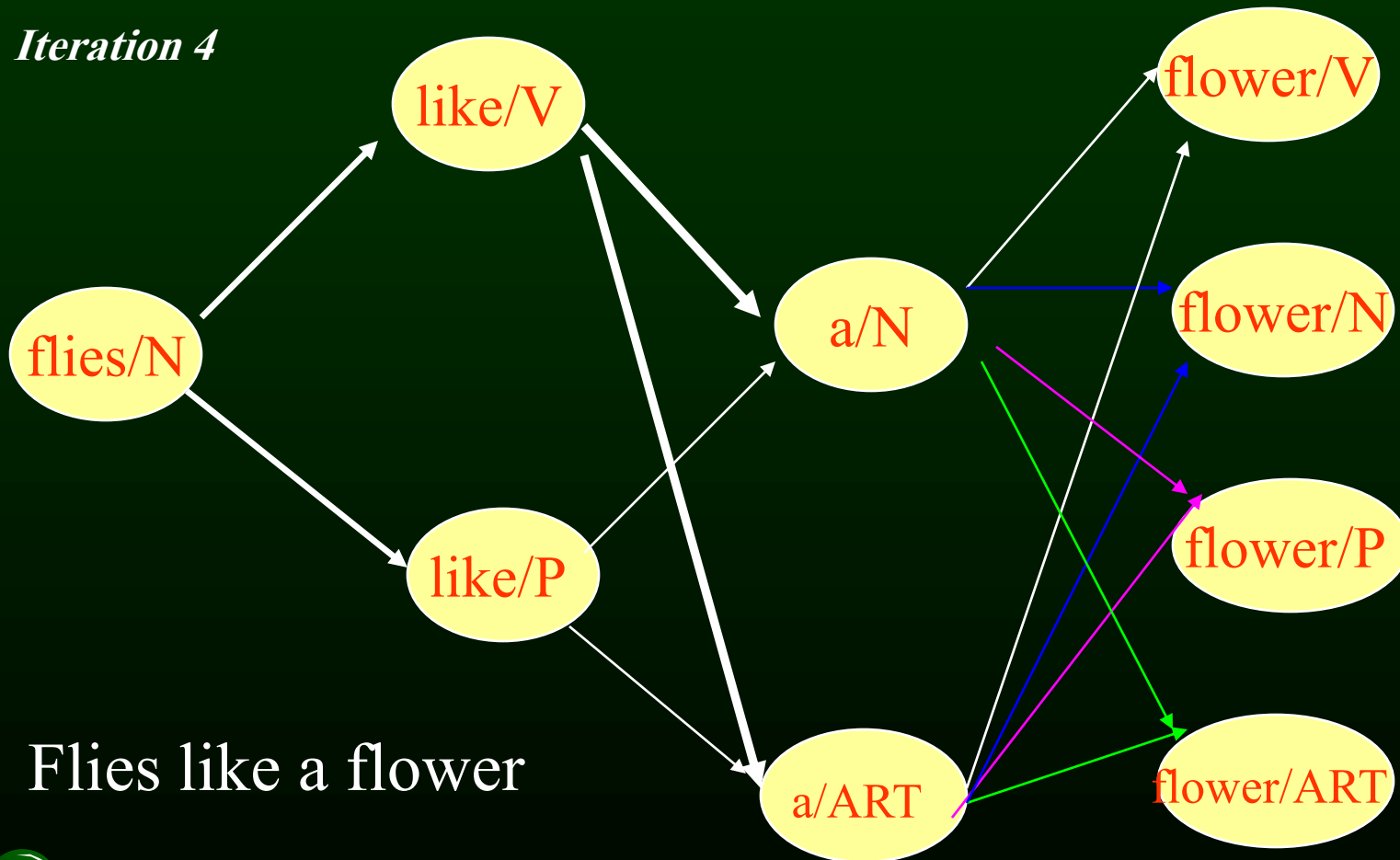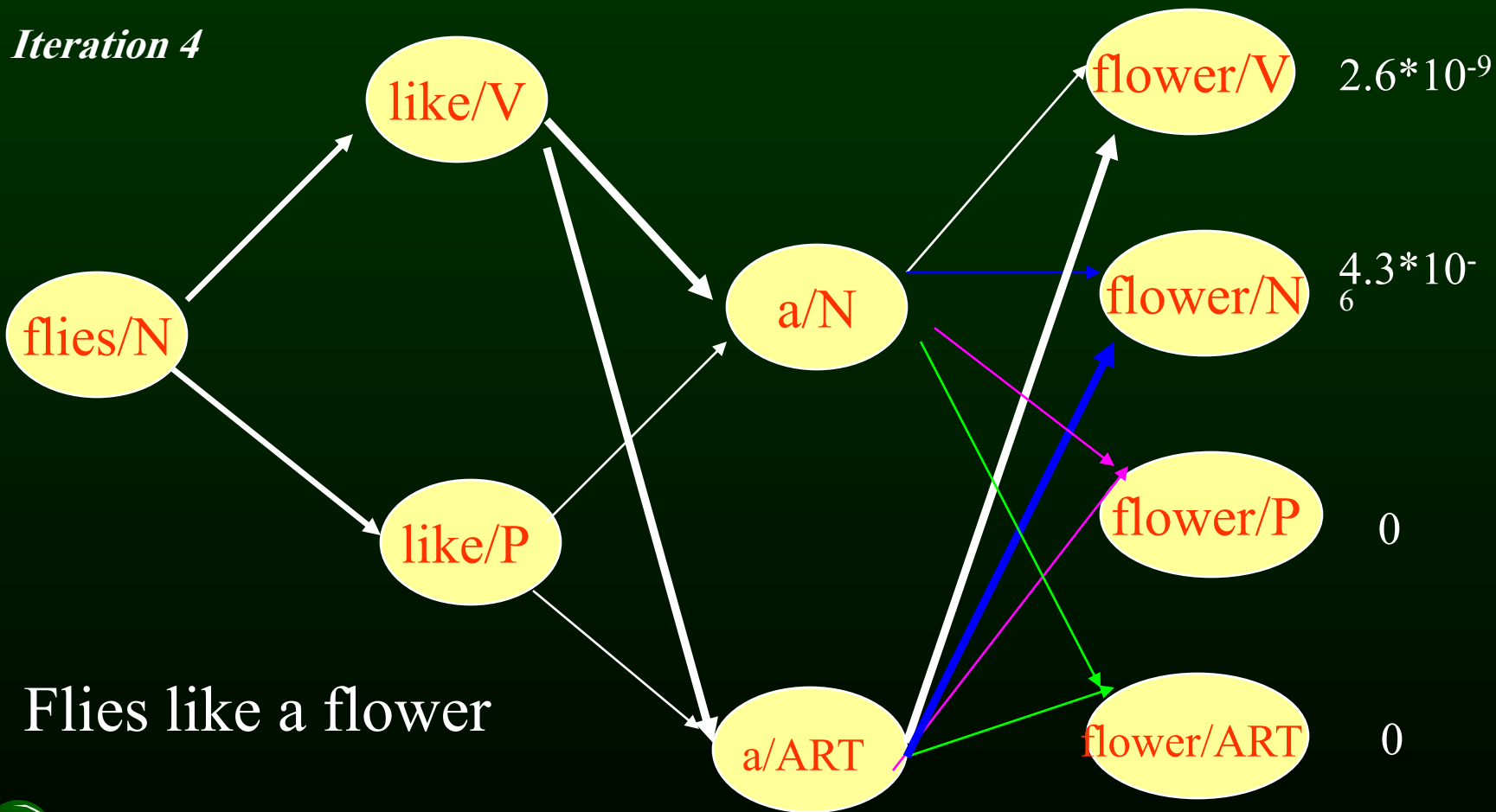
Performance

- This method has achieved 95-96% correct with reasonably complex English tagsets and reasonable amounts of hand-tagged training data.

- Forward pointer… its also possible to train a system without hand-labeled training data

How accurate are they?

- POS Taggers boast accuracy rates of 95-99%
  - Vary according to text/type/genre
    - Of pre-tagged corpus
    - Of text to be tagged
- Worst case scenario: assume success rate of 95%
  - Prob(one-word sentence) = .95
  - Prob(two-word sentence) = .95 * .95 = 90.25%
  - Prob(ten-word sentence) = 59% approx

- End of Part 1

بسم الله الرحمن الرحيم
معالجة اللغات الطبيعية
Natural Language Processing

Lecture 10: Parts of Speech 2-2
Morphosyntactic Tagset Of Arabic
أوسام صرفية ــ نحوية للعربية
Husni Al-Muhtaseb

أوسام صرفية – نحوية للعربية

- شيرين خوجة Shereen Khoja
- ١٧٧ وسما 177 tags
- ١٠٣ للأسماء 103 Nouns
- ٥٧ للأفعال 57 Verbs
- ٩ للأدوات 9 Particles
- ٧ للفضلة 7 Residual
- ١ لعلامات الترقيم 1 Punctuation

اللغة العربية

- اللغة العربية من اللغات السامية
- تبنى الكلمات بإضافة زوائد إلى الجذر واستخدام أوزان محددة
  - درس مدرس
- ثلاثة أجناس three genders:
  - مذكر Masculine
  - مؤنث Feminine
  - حيادي Neuter

اللغة العربية

- الخطاب Three persons
  - المتكلم The speaker
  - المخاطب The person being addressed
  - الغائب The person that is not present
- العدد Three numbers
  - مفرد Singular
  - مثنى Dual
  - جمع Plural

اللغة العربية

- three moods of the verb حالات الفعل
  - الرفع Indicative
  - النصب Subjunctive
  - الجزم Jussive
- three case forms of the noun حالات الاسم
  - الرفع Nominative
  - النصب Accusative
  - الجر Genitive
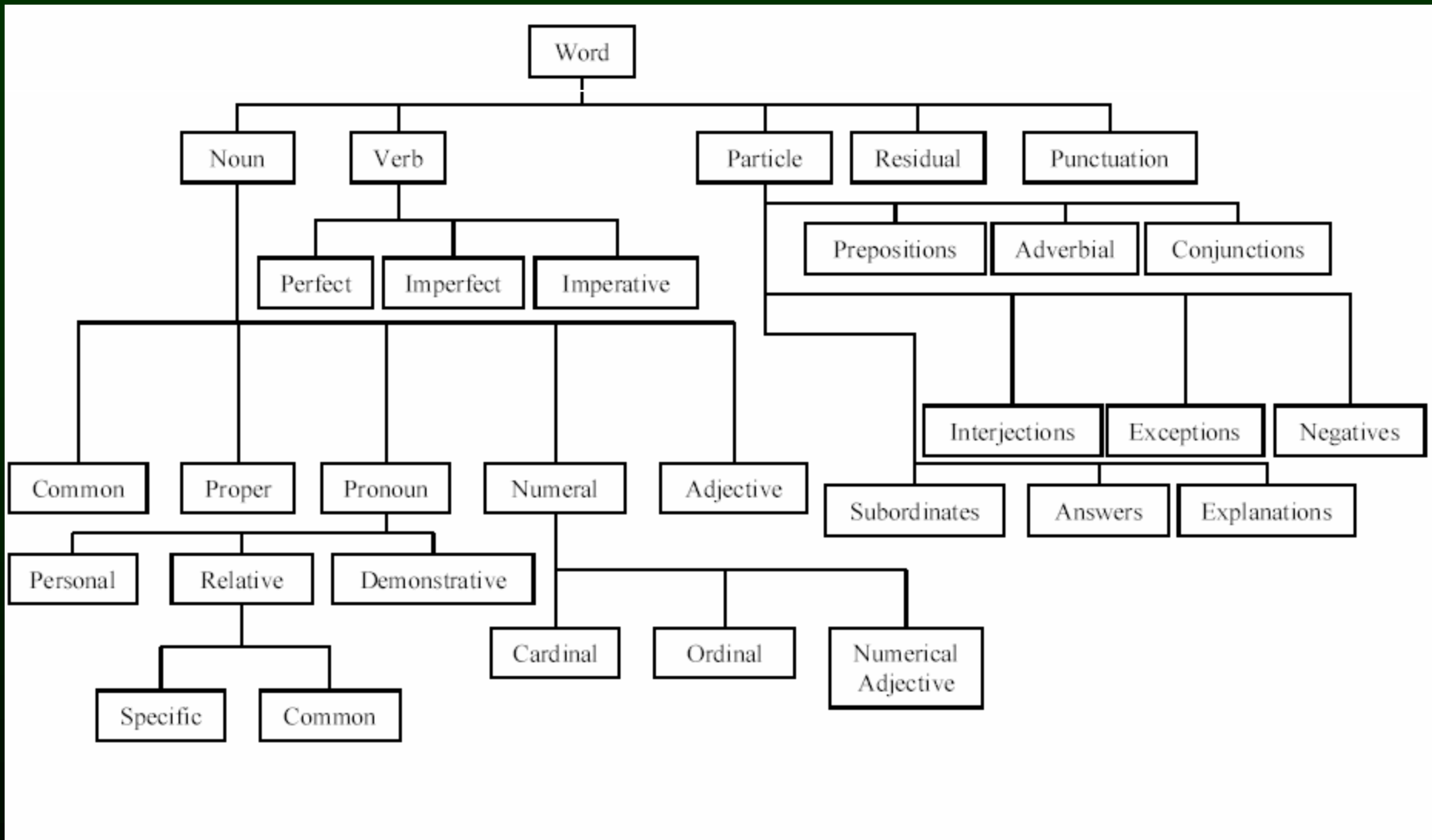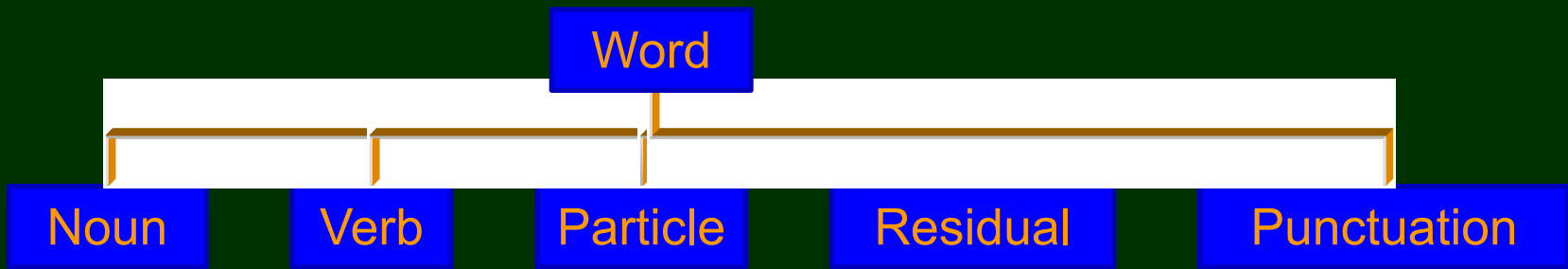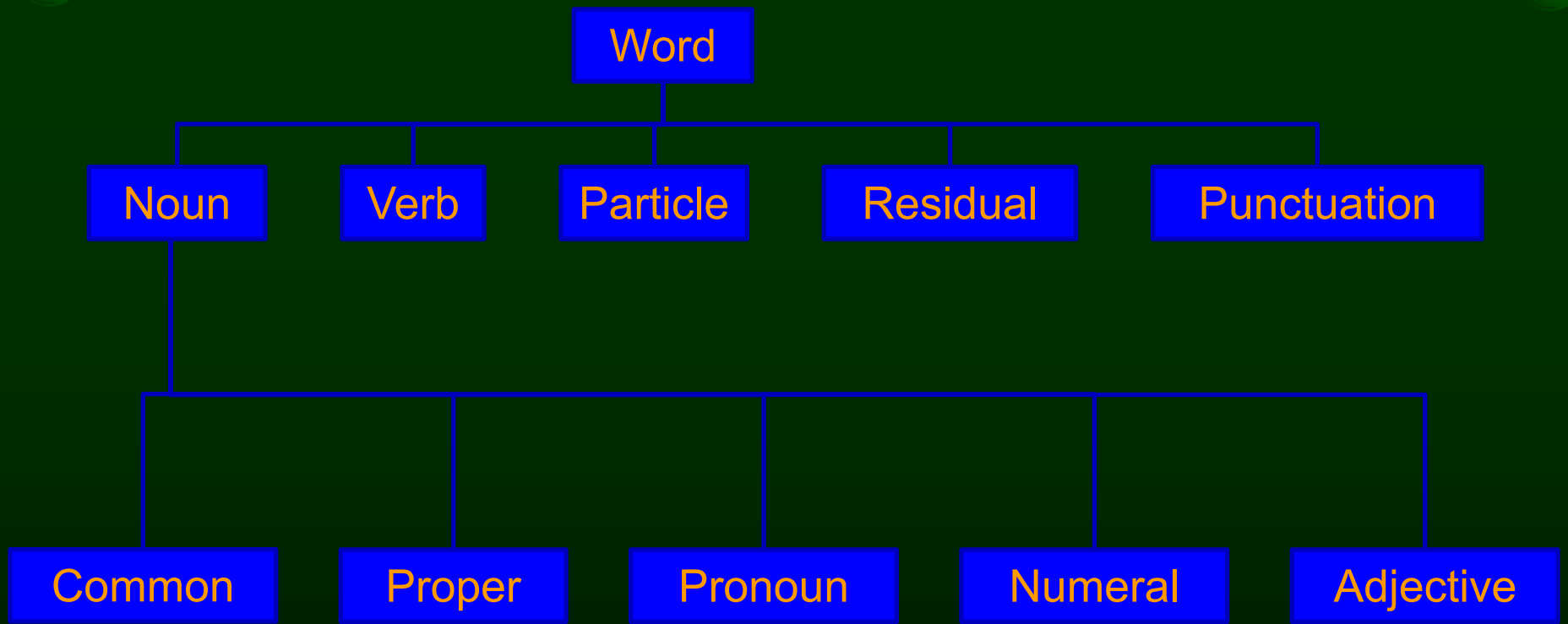
```
                          ┌──────────┐
                          │   Word   │
                          └────┬─────┘
     ┌──────────┬─────────────┼─────────────┬──────────────────┐
┌────┴───┐ ┌────┴───┐ ┌───────┴──┐ ┌─────────┴──┐ ┌─────────────┴────┐
│  Noun  │ │  Verb  │ │ Particle │ │  Residual  │ │   Punctuation    │
└────────┘ └────────┘ └──────────┘ └────────────┘ └──────────────────┘
```

```
                              ┌──────────┐
                              │   Word   │
                              └──────────┘
          ┌──────────┬────────────┼────────────┬──────────────┐
     ┌────────┐ ┌────────┐  ┌───────────┐ ┌───────────┐ ┌──────────────┐
     │  Noun  │ │  Verb  │  │ Particle  │ │ Residual  │ │ Punctuation  │
     └────────┘ └────────┘  └───────────┘ └───────────┘ └──────────────┘
               ┌──────┼─────────────┐
          ┌─────────┐ ┌───────────┐ ┌────────────┐
          │ Perfect │ │ Imperfect │ │ Imperative │
          └─────────┘ └───────────┘ └────────────┘
```

Arabic POS Tagger

Plain Arabic Text → ManTag → Training Corpus → DataExtract

DataExtract → Probability Matrix

DataExtract → Lexicons

Probability Matrix → APT

Lexicons → APT

Tagged Corpus ← APT ← Untagged Arabic Corpus
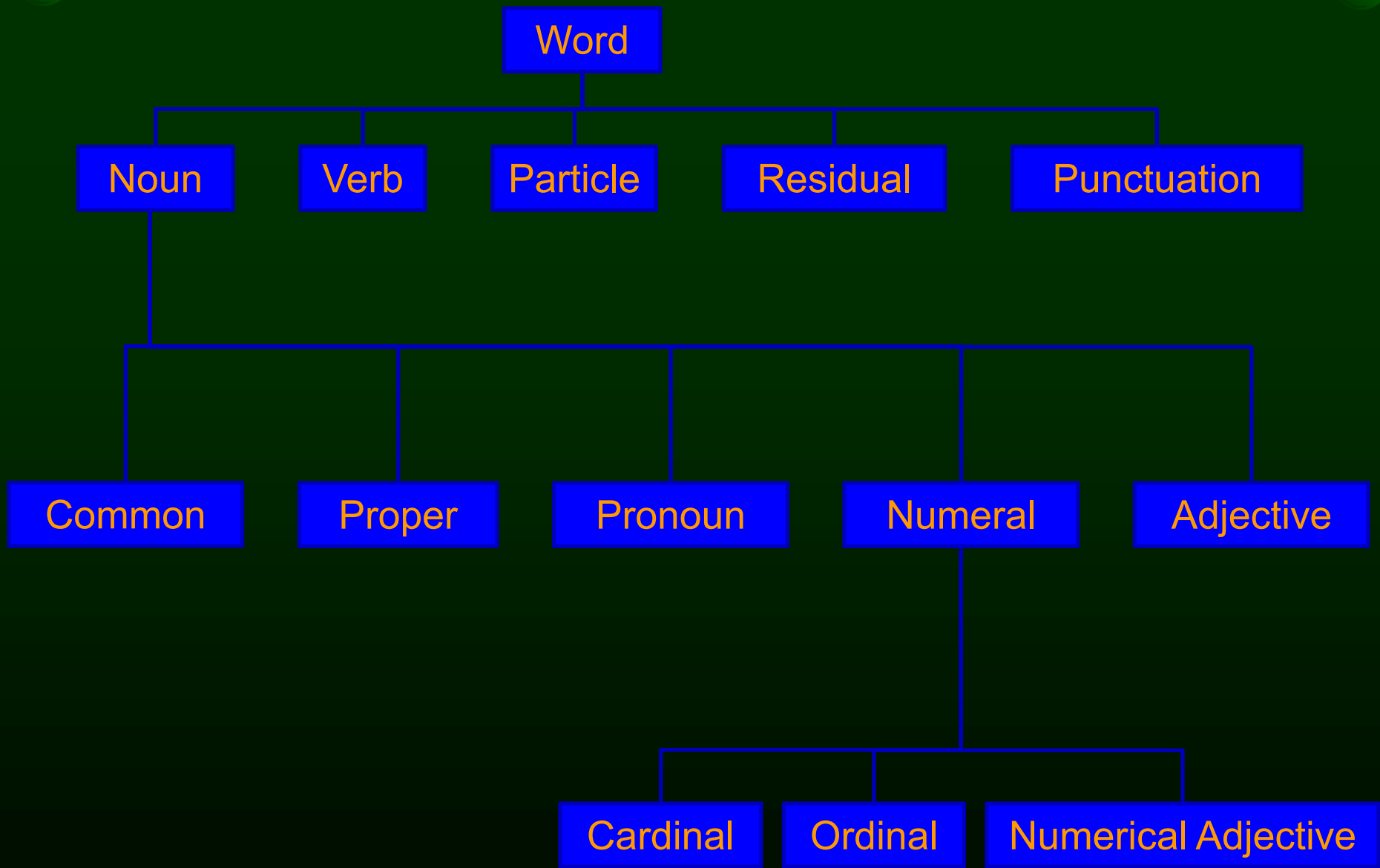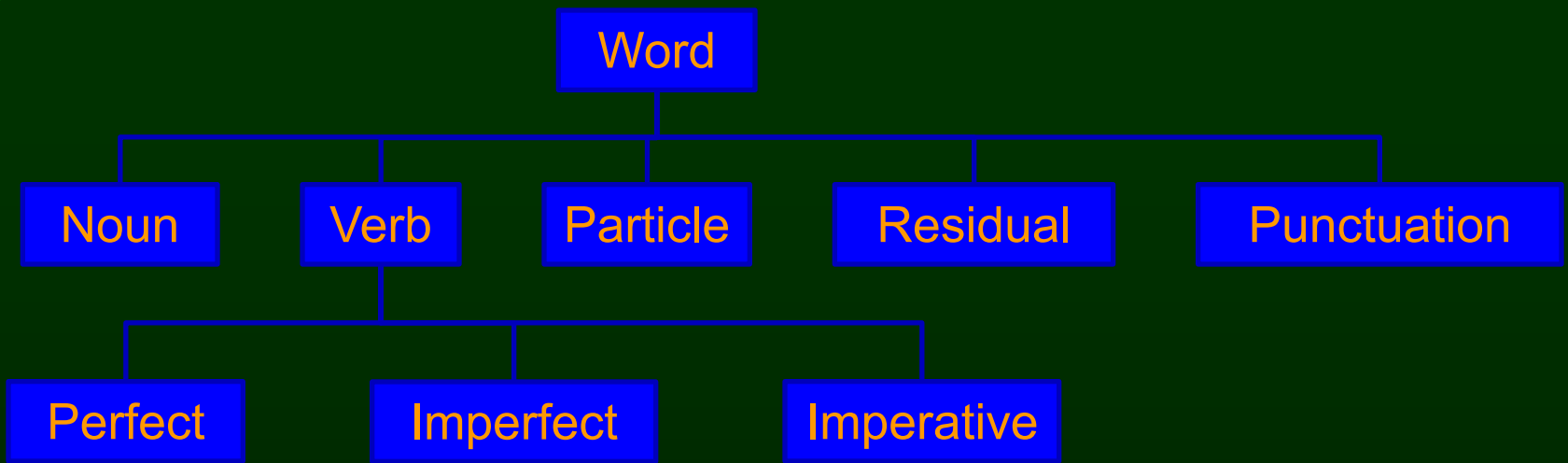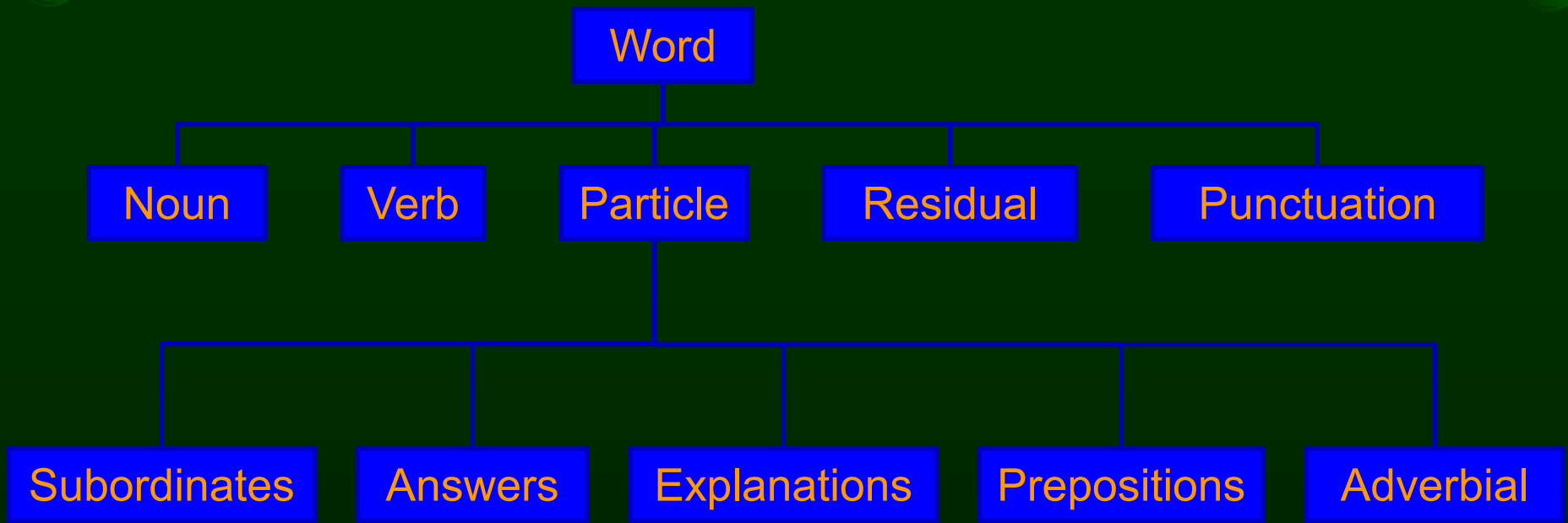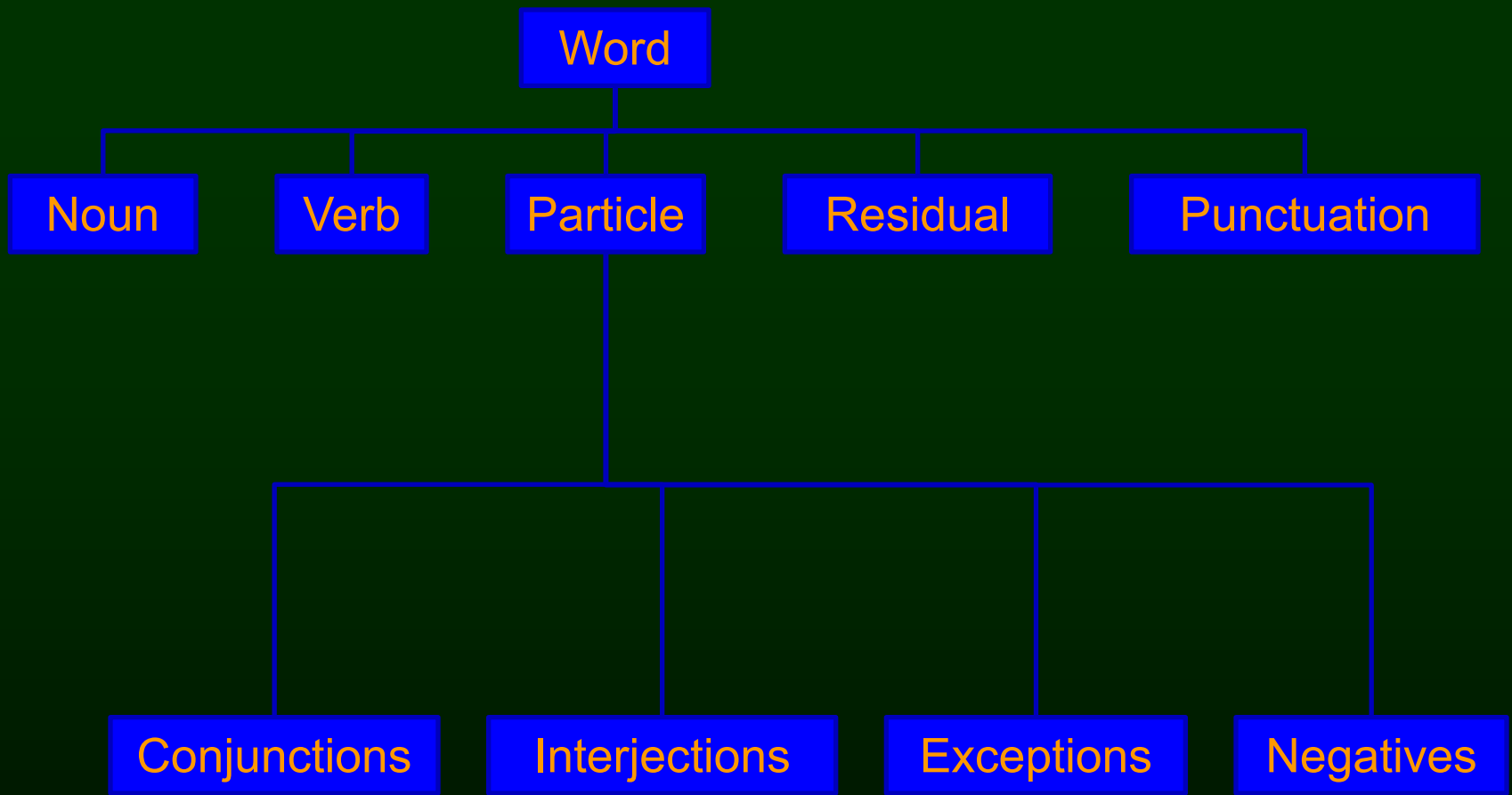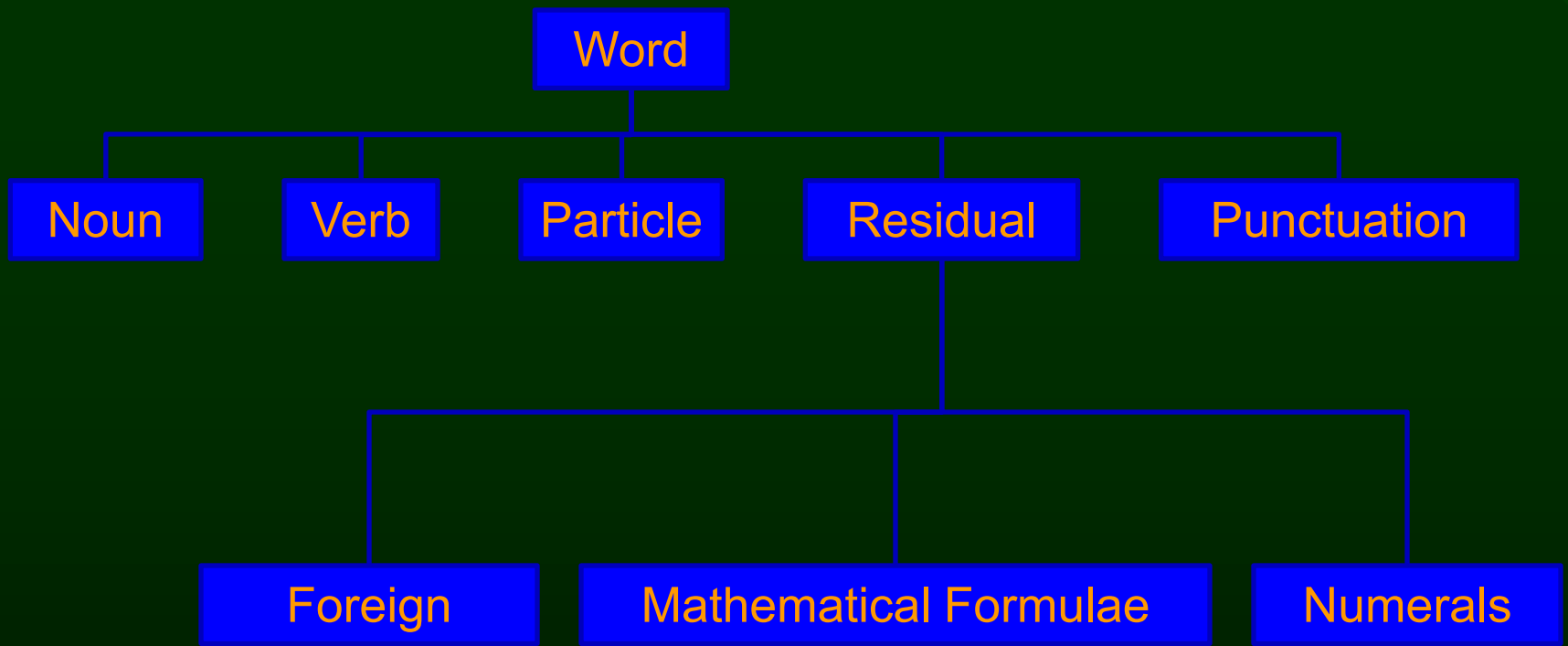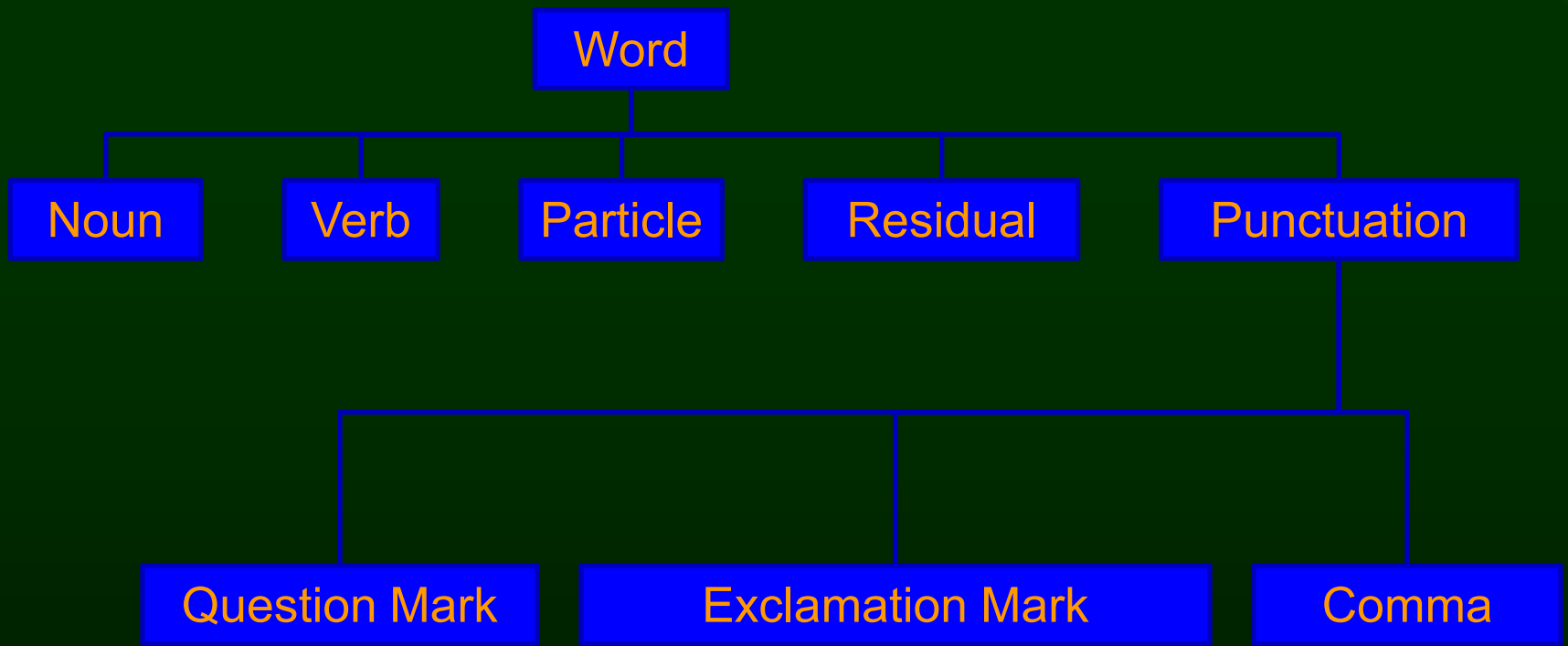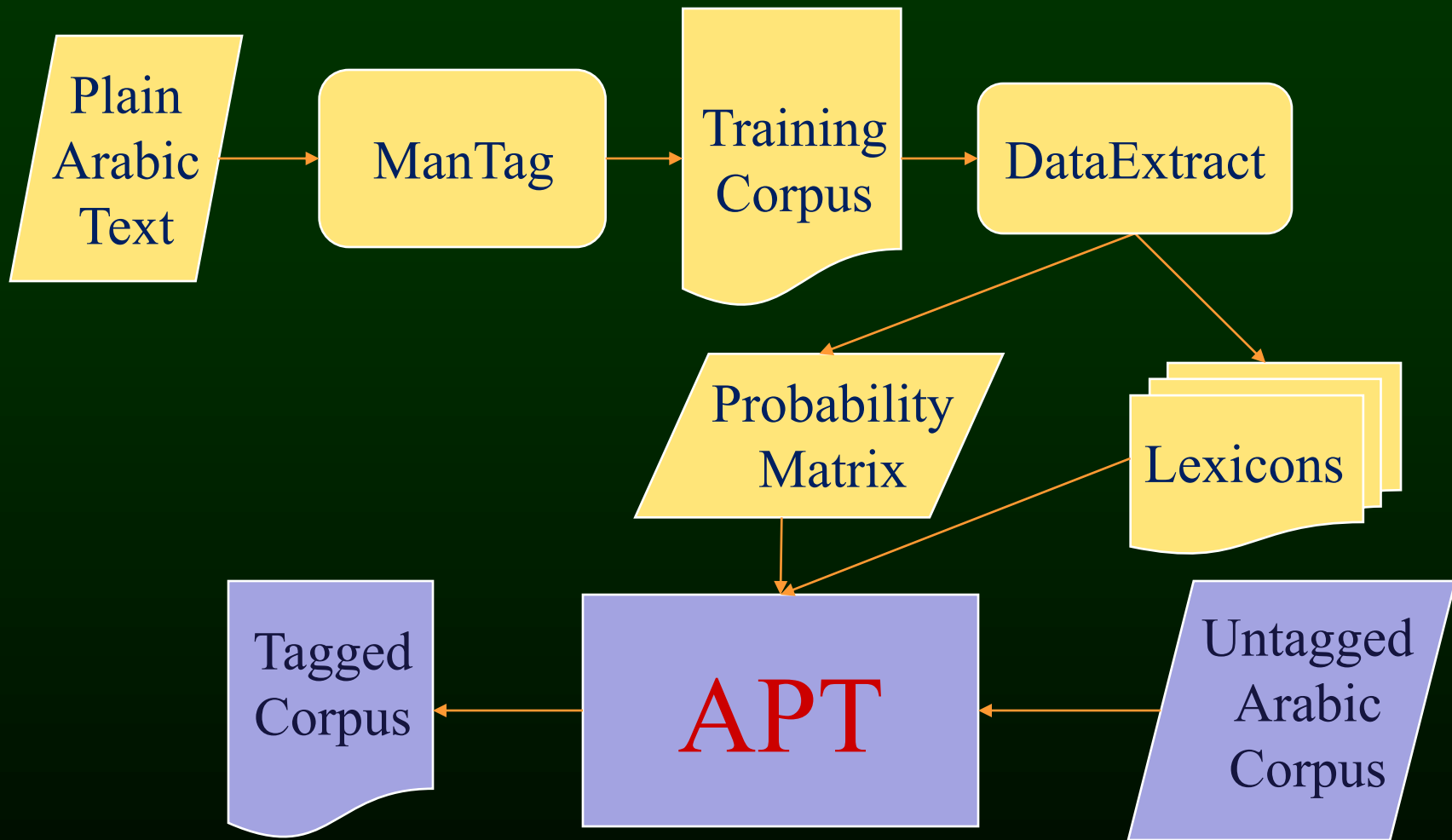
DataExtract Process

- Takes in a tagged corpus and extracts various lexicons and the probability matrix
  - Lexicon that includes all clitics.
    - (Sprout, 1992) defines a clitic as "a *syntactically* separate word that functions *phonologically* as an affix"
  - Lexicon that removes all clitics before adding the word

DataExtract Process

- Produces a probability matrix for various levels of the tagset

    - Lexical probability: probability of a word having a certain tag

    - Contextual probability: probability of a tag following another tag

DataExtract Process

	N	V	P	No.	Pu.
N	0.711	0.065	0.143	0.010	0.071
V	0.926	0.037	0.0	0.008	0.029
P	0.689	0.199	0.085	0.016	0.011
No.	0.509	0.06	0.098	0.009	0.324
Pu.	0.492	0.159	0.152	0.046	0.151

Arabic Corpora

- 59,040 words of the Saudi "al-Jazirah" newspaper, dated 03/03/1999

- 3,104 words of the Egyptian "al-Ahram" newspaper, date 25/01/2000

- 5,811 words of the Qatari "al-Bayan" newspaper, date 25/01/2000

- 17,204 words of al-Mishkat, an Egyptian published paper in social science, April 1999

APT: Arabic Part-of-speech Tagger
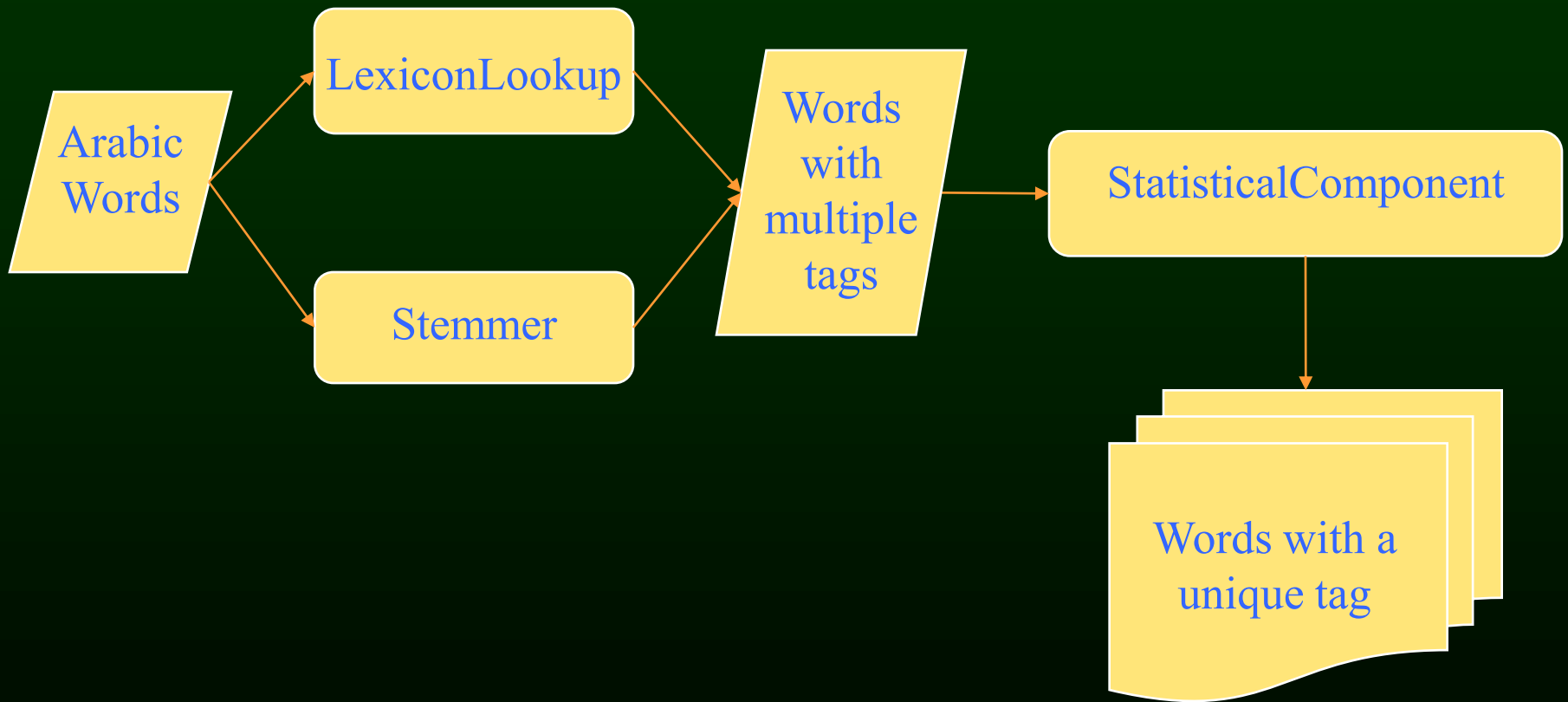
كلمة

- علامة ترقيم
- اسم
- فضلة
- أداة
- فعل

فعل:
- ماض
- مضارع
- أمر

أداة:
- تفسير
- تمني
- نفي
- استثناء
- ظرف
- إجابة
- جر
- عطف
- نداء

اسم:
- عدد
- علم
- صفة
- ضمير
- جنس

عدد:
- صفة عددية
- عددي
- ترتيبي

ضمير:
- إشارة
- شخصي
- موصول

موصول:
- عام
- محدد

٥ فئات أساسية

- اسم   1. N [noun]
- فعل   2. V [verb]
- أداة   3. P [particle]
- فضلة: الكلمات الغربية والصيغ الرياضية والأرقام
  4. R [residual]

- علامة ترقيم (كلها)
  5. PU [punctuation]: all

الاسم

1.1. C [common]	• جنس
1.2. P [proper]	• علم
1.3. Pr [pronoun]	• ضمير
1.4. Nu [numeral]	• عدد
1.5. A [adjective]	• صفة

Singular, masculine, accusative, common noun •

أخذ الولد كتابا

مفرد مذكر منصوب اسم جنس

Singular, masculine, genitive, common noun •

درست من الكتاب

مفرد مذكر مجرور اسم جنس

Singular, feminine, nominative, common noun •

هذه مدرسة

مفرد مؤنث مرفوع اسم جنس

الضمائر

- الشخصية 1.3.1. P [personal]
- منفصلة مثل **هو** detached words such as
- متصلة attached to a word
- مع الأسماء للملكية مثل **كتابها** to nouns to indicate possession
- متصلة مع الأفعال كمفعول به مثل **ضربه** to verbs as direct object
- متصلة مع حروف الجر مثل **فيه** prepositions
- الموصولة 1.3.2. R [relative]
- الإشارة 1.3.3. D [demonstrative]

الضمائر

- Third person, singular, masculine, personal pronoun

هو

ضمير الغائب مفرد مذكر شخصي

- Singular, feminine, demonstrative pronoun

هذه

اسم إشارة مؤنث مفرد

Relative Pronoun الضمائر الموصولة

- محددة   1.3.2.1. S [specific]
- عامة   1.3.2.2. C [common]
- أمثلة
  - اللتان

  ضمير موصول محدد مؤنث مثنى
  - Dual, feminine, specific, relative pronoun
  - الذين

  ضمير موصول محدد مذكر جمع
  - Plural, masculine, specific, relative pronoun
  - ضمير موصول عام
  - Common, relative pronoun

العدد

- عددي   1.4.1. Ca [cardinal]
- ترتيب   1.4.2.O [ordinal]
- صفة عددية   1.4.3. Na [numerical adjective]:
  - تصف عدد جهات شكل **ثماني**
  - وصف شخصين مثلا **ثنائي**
- أمثلة
- Singular, masculine, nominative, indefinite cardinal number

أربعة

مفرد مذكر مرفوع نكرة عدد عددي

- Singular, masculine, nominative, indefinite ordinal number

رابع

مفرد مذكر مرفوع نكرة عدد ترتيبي

- Singular, masculine, numerical adjective

رباعي

مفرد مذكر صفة عددية

الخواص اللغوية للاسم

- الخطاب Person
  - متكلم [first] 1
  - مخاطب [second] 2
  - غائب [third] 3
- الحالة Case
  - رفع [nominative] N
  - نصب [accusative] A
  - جر [genitive] G
- التعريف Definiteness
  - معرفة [definite] D
  - نكرة [indefinite] I

- الجنس Gender
- مذكر [masculine] M
- مؤنث [feminine] F
- محايد [neuter] N
- العدد Number
- مفرد [singular] Sg
- مثنى [dual] Du
- جمع [plural] Pl

Verbs الأفعال

- تام (ماض)   1. P [perfect]
- مستمر (مضارع)   2. I [imperfect]
- أمر   3. Iv [imperative]
- أمثلة
- كَسَرْتُ

فعل ماض حيادي مفرد متكلم

- First person, singular, neuter, perfect verb
- أَكْسِرُ

فعل مضارع مرفوع مفرد متكلم

- First person, singular, neuter, indicative, imperfect verb
- إِكْسِرْ

فعل أمر مذكر مفرد مخاطب

- Second person, singular, masculine, imperative verb

Verbal Attributes الخواص اللغوية للفعل Used

- الخطاب Person
  - متكلم [first] 1
  - مخاطب [second] 2
  - غائب [third] 3
- الحالة Mood
  - رفع [indicative] I
  - نصب [subjunctive] S
  - جزم [jussive] J

- الجنس Gender
  - مذكر [masculine] M
  - مؤنث [feminine] F
  - حيادي [neuter] N
- العدد Number
  - مفرد [singular] Sg
  - مثنى [plural] Pl
  - جمع [dual] Du

الأدوات

- جر [prepositions] 1.1. Pr
- ظرف [adverbial] 1.2. A
- عطف [conjunctions] 1.3. C
- نداء [interjections] 1.4. I
- استثناء [exceptions] 1.5. E
- نفي [negatives] 1.6 N
- جواب [answers] 1.7. A
- تفسير [explanations] 1.8. X
- تمني [subordinates] 1.9. S

أمثلة على الأدوات

- جر: في   Prepositions "in"
- ظرف: سوف   Adverbial particles "shall"
- عطف: و   Conjunctions "and"
- نداء: يا   Interjections "you"
- استثناء: سوى   Exceptions "Except"
- نفي: لم   Negatives "Not"
- جواب: أجل   Answers "yes"
- تفسير: أي   Explanations "that is"
- تمني: لو   Subordinates "if"

اوسمة شيرين

Tag	Description of word category	Example (Arabic)
NCSgMNI	Singular, masculine, nominative, indefinite common noun	كتابٌ
NCSgMAI	Singular, masculine, accusative, indefinite common noun	كتاباً
NCSgMGI	Singular, masculine, genitive, indefinite common noun	كتابٍ
NCSgMND	Singular, masculine, nominative, definite common noun	الكتابُ
NCSgMAD	Singular, masculine, accusative, definite common noun	الكتابَ
NCSgMGD	Singular, masculine, genitive, definite common noun	الكتابِ
NCSgFNI	Singular, feminine, nominative, indefinite common noun	مدرسةٌ
NCSgFAI	Singular, feminine, accusative, indefinite common noun	مدرسةً
NCSgFGI	Singular, feminine, genitive, indefinite common noun	مدرسةٍ
NCSgFND	Singular, feminine, nominative, definite common noun	المدرسةُ

اوسمة شيرين

	common noun	
NCDuFAD	Dual, feminine, accusative, definite common noun	المدرستين
NCDuFGD	Dual, feminine, genitive, definite common noun	المدرستين
NCPlMNI	Plural, masculine, nominative, indefinite common noun	كتبٌ - مسلمون
NCPlMAI	Plural, masculine, accusative, indefinite common noun	كتباً - مسلمين
NCPlMGI	Plural, masculine, genitive, indefinite common noun	كتبٍ - مسلمين
NCPlMND	Plural, masculine, nominative, definite common noun	الكتبُ - المسلمون
NCPlMAD	Plural, masculine, accusative, definite common noun	الكتبَ - المسلمين
	Plural, masculine, genitive, definite	

اوسمة شيرين

NCPlFGD	Plural, feminine, genitive, definite common noun	المسلمات-المدارس
NP	Proper noun	شيرين-جده
NPrPSg1	First person, singular, neuter, personal pronoun	كتابي - ضربني-أنا
NPrPSg2M	Second person, singular, masculine, personal pronoun	كتابك-أنت
NPrPSg2F	Second person, singular, feminine, personal pronoun	كتابك-أنت
NPrPSg3M	Third person, singular, masculine, personal pronoun	هو - كتابه
NPrPSg3F	Third person, singular, feminine, personal	هي - كتابها

NACPlFAD	Plural, feminine, accusative, definite adjective	السعيدات
NACPlFGD	Plural, feminine, genitive, definite adjective	السعيدات
VPSg1	First person, singular, neuter, perfect verb	كسرتُ
VPSg2M	Second person, singular, masculine, perfect verb	كسرتَ
VPSg2F	Second person, singular, feminine, perfect verb	كسرت
VPSg3M	Third person, singular, masculine, perfect verb	كسر
VPSg3F	Third person, singular, feminine, perfect verb	كسرتْ
VPDu2	Second person, dual, neuter, perfect verb	كسرتما

VPPl2F	Second person, plural, feminine, perfect verb		كسرتن
VPPl3M	Third person, plural, masculine, perfect verb		كسروا
VPPl3F	Third person, plural, feminine, perfect verb		كسرن
VISg1I	First person, singular, neuter, indicative, imperfect verb		اكسرَ
VISg1S	First person, singular, neuter, subjunctive, imperfect verb		اكسرِ

PN	Negatives	
PW	Answers	لا-لجل
PX	Explanations	أي
PS	Subordinates	لو-سما
RF	Residual, foreign	روجور
RM	Residual, mathematical	÷
RN	Residual, number	3
RD	Residual, day of the week	الاثنين
Rmy	Residual, month of the year	محرم
RA	Residual, abbreviation	واس
RO	Residual, other	ال
PU	Punctuation	؟

بعث_VPSg3M خادم_NCSgMNI الحرمين_NCDuMAD الشريفين_NCDuMGD الملك_NCSgMND فهد_NP بن_

NCSgMNI عبد_NCSgMAI العزيز_NCSgMAD آل_R سعود_NP برقية_NCSgFNI تهنئة_NCSgFGI الى_PPr

فخامة_NCSgFGI الرئيس_NCSgMGD الكسندر_RF كراسنوفيسكي_RF رئيس_NCSgMNI جمهورية_NCSgFGI

بولندا_RF بمناسبة_PPr NCSgFGI اليوم_NCSgMAD الوطني_NCSgMND لبلاده_NPrPSg3M NCPlFGI PPr

وأعرب_PC_VPSg3M الملك_NCSgMND المفدى_NCSgMAD باسمه_NPrPSg3M NCSgMGI PPr وباسم_

PC_PPr NCSgM شعب_NCSgMGI وحكومة_PC NCSgFGI المملكة_NCSgFGD العربية_NCSgFGD السعودية_

NCSgFGD عن_PPr اخلص_NCSgFNI التهاني_NCPlMND متمنيا_NCSgMAI لفخامته_

PC_NCSgFGD والسعادة_NCSgFGD الصحة_NCSgFGD دوام_NCSgMNI PPr NCSgFGI NPrPSg3M ولشعب_

PC_PPr NCSgMGI بولندا_RF الصديق_NCSgMND الازدهار_NCSgMND الدائم_NCSgMN . PU

أوسام صالح العصيمي
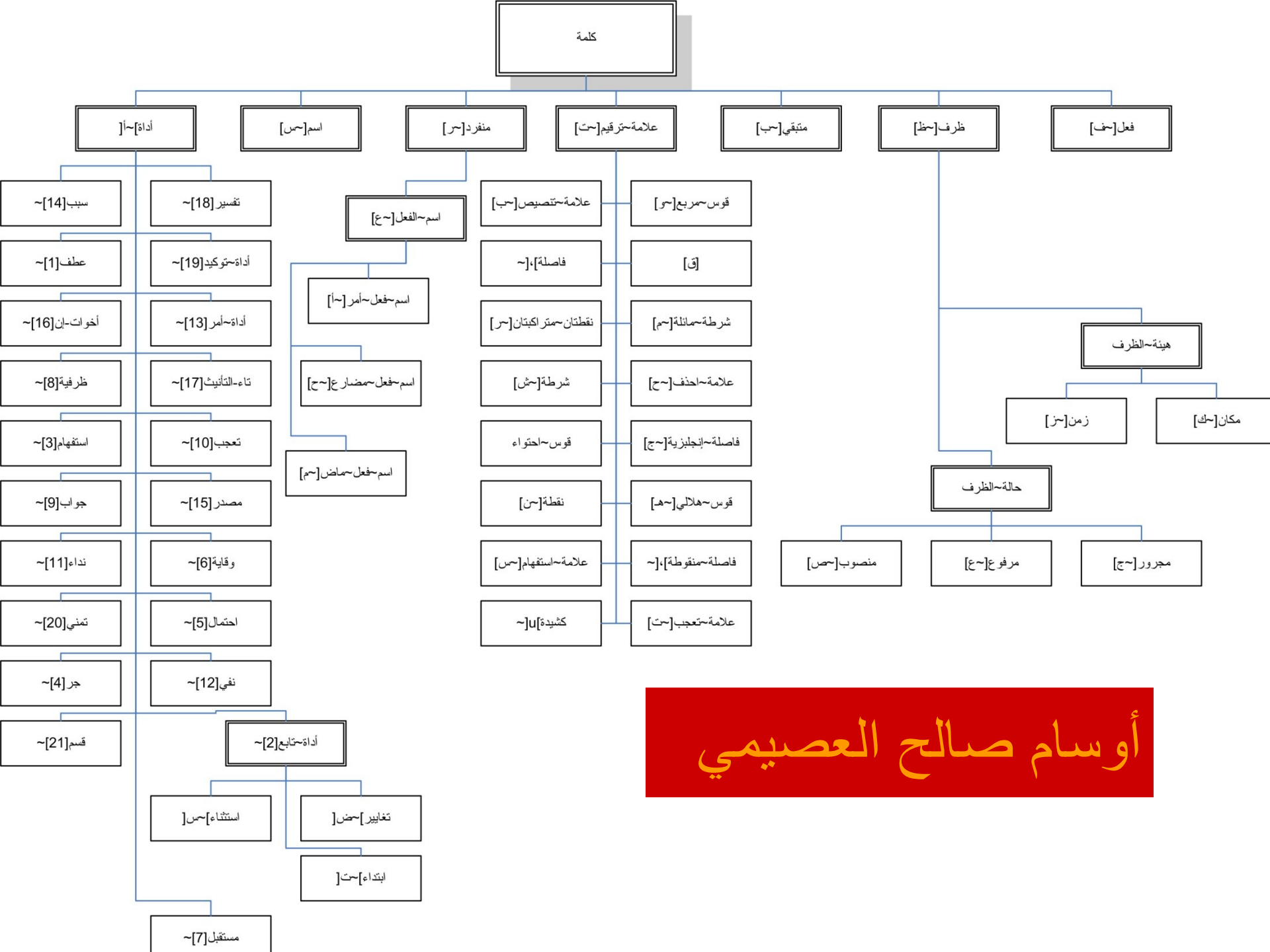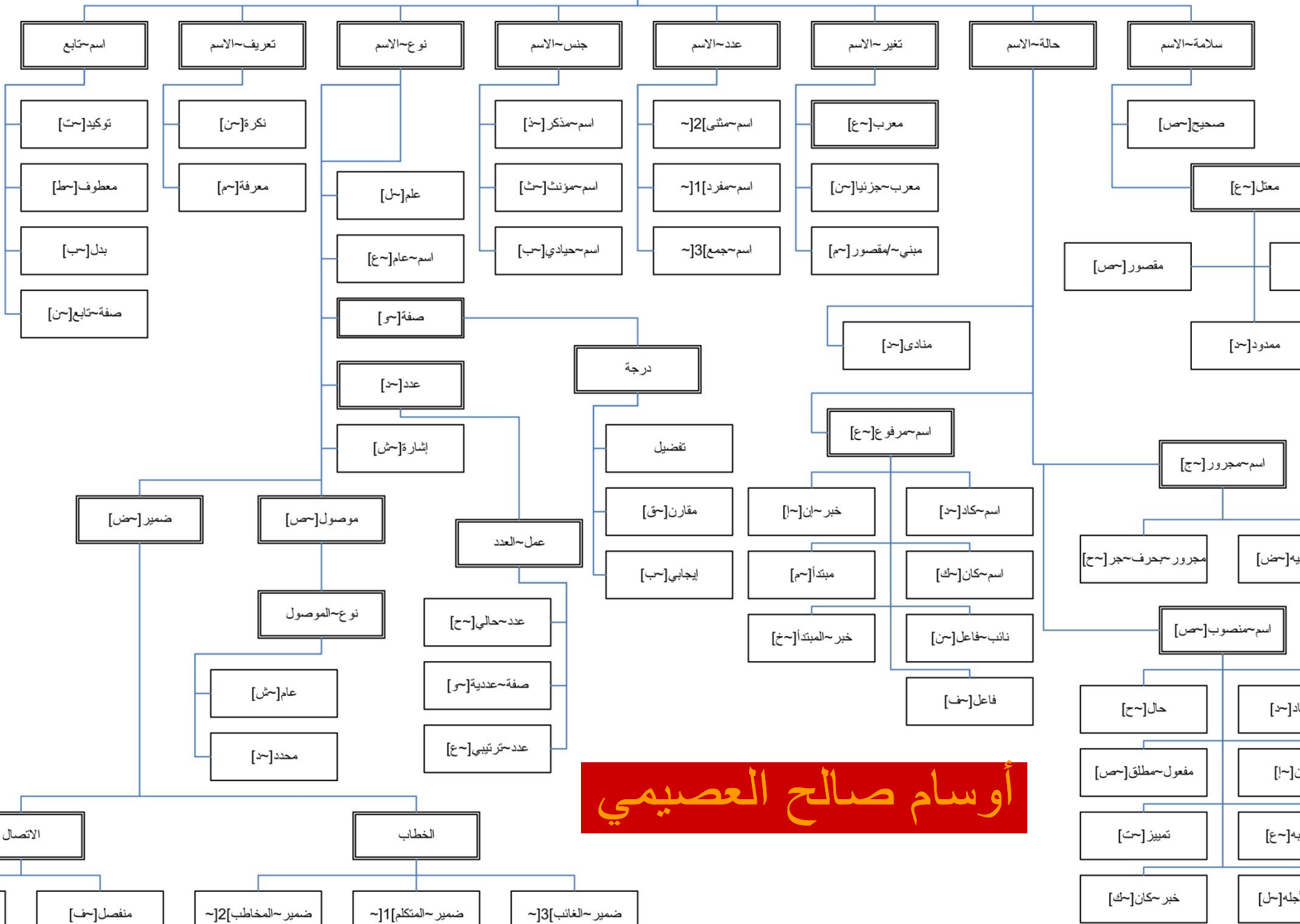
أوسام صالح العصيمي
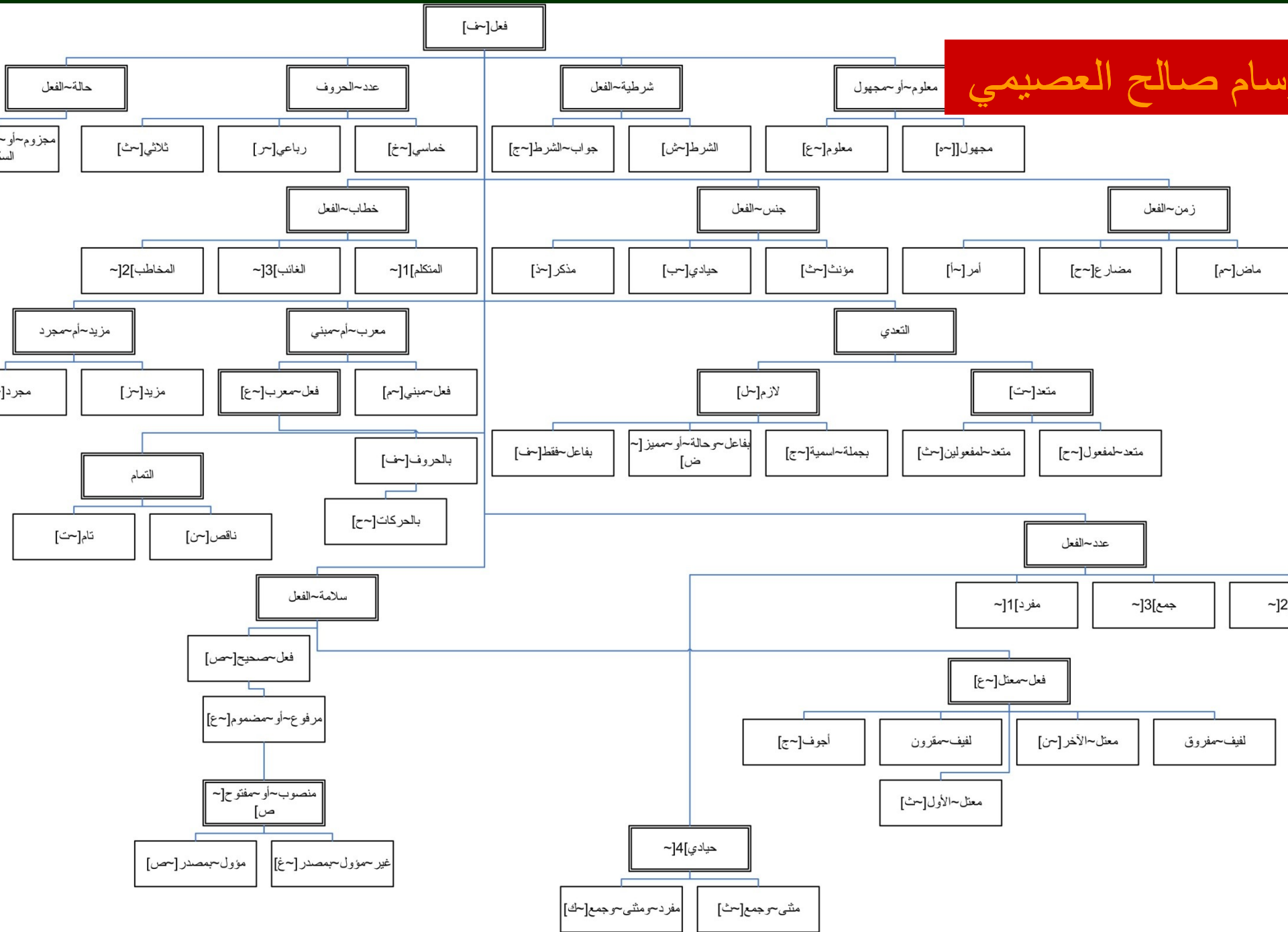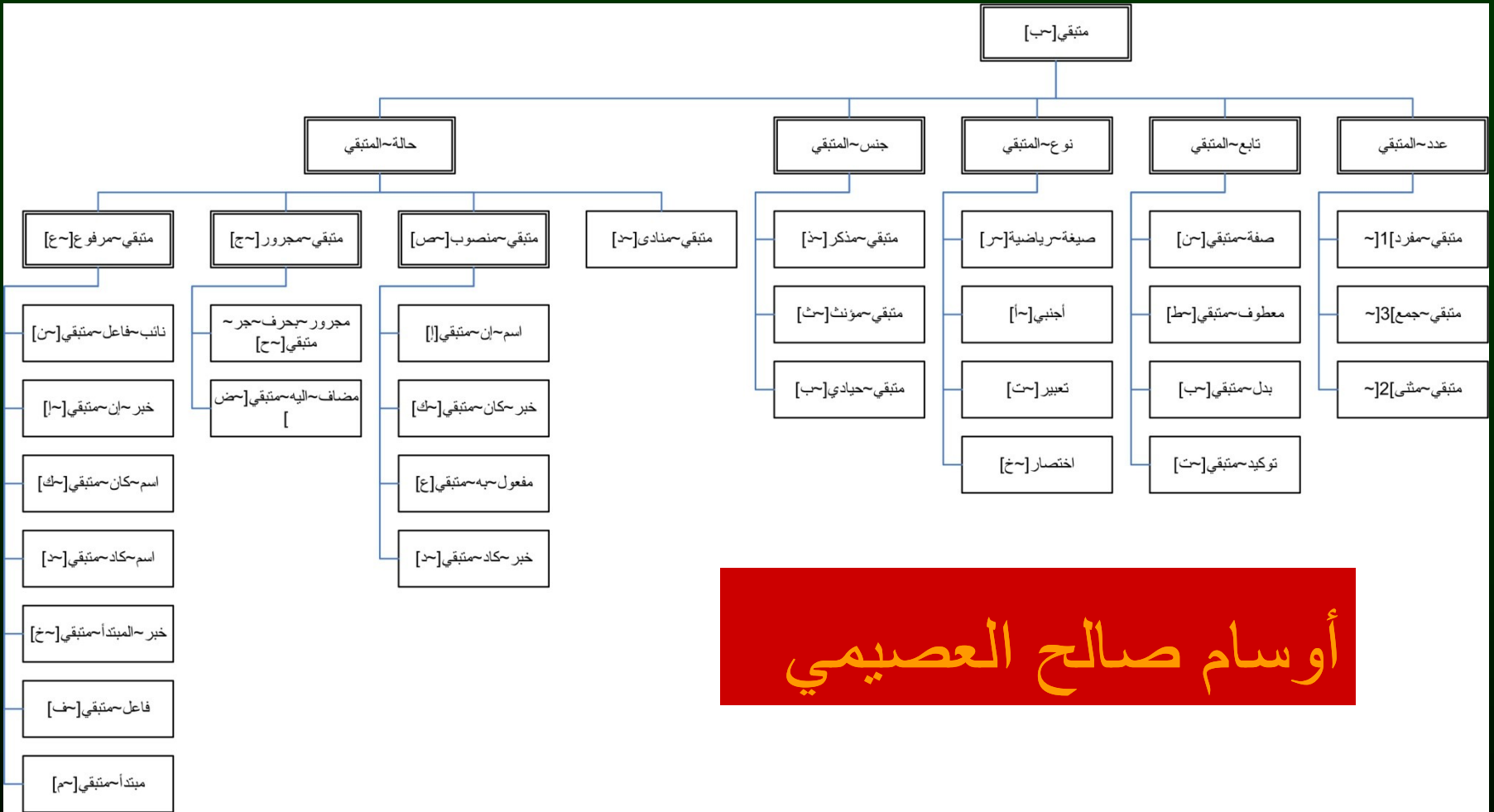
أوسام صالح العصيمي

أوسام صالح العصيمي

Parts of Speech

Part of Speech
أجزاء الكلام

Nouns
الأسماء

Adverbs
الظروف

Verbs
الأفعال

Particles
الأدوات

Unique
منفرد

Residual
المتبقي ، الفضالة

Punctuation
علامات الترقيم

1. Noun

Nouns (N – س)
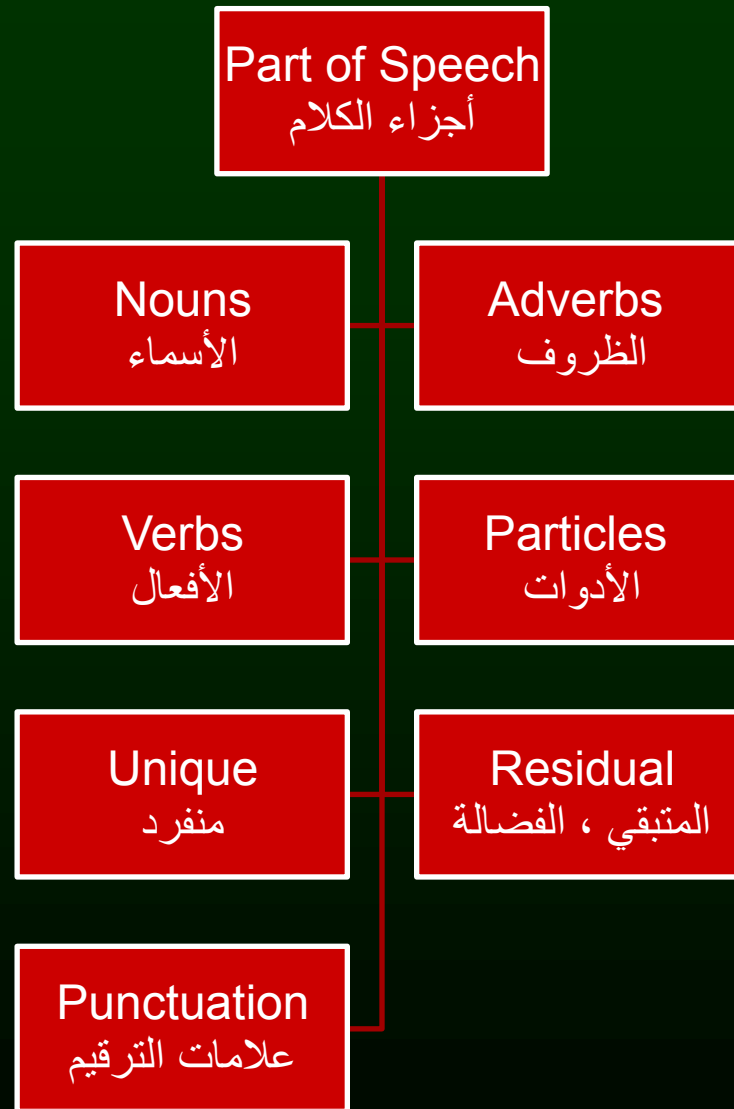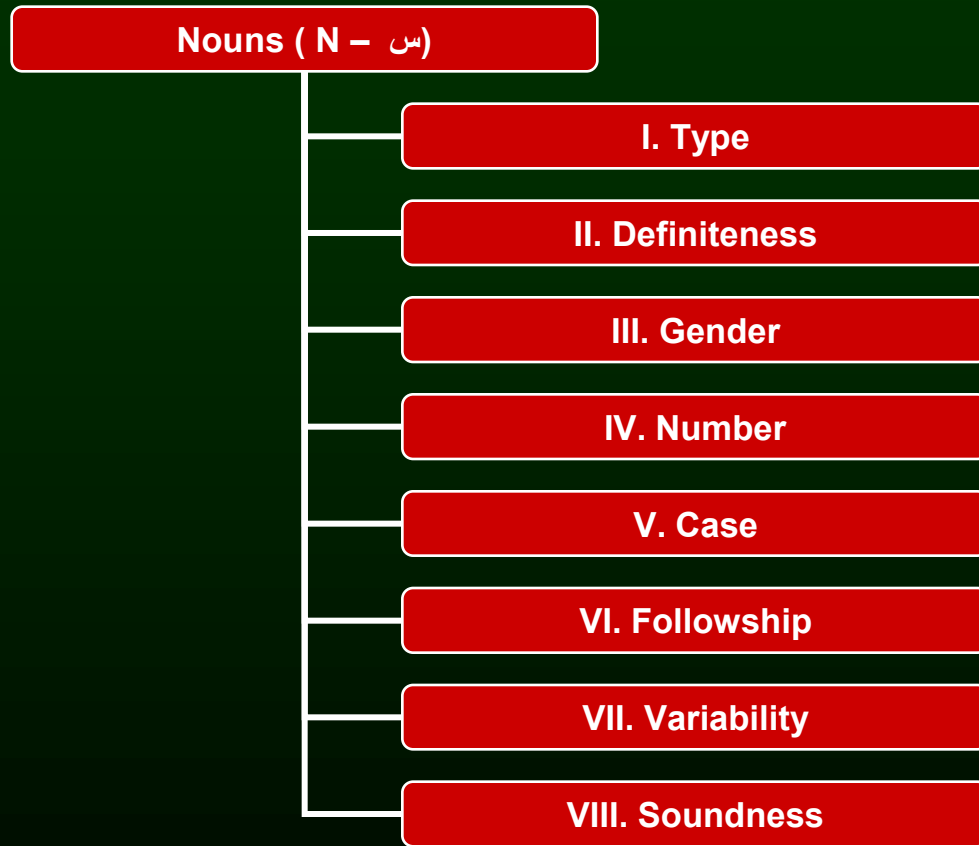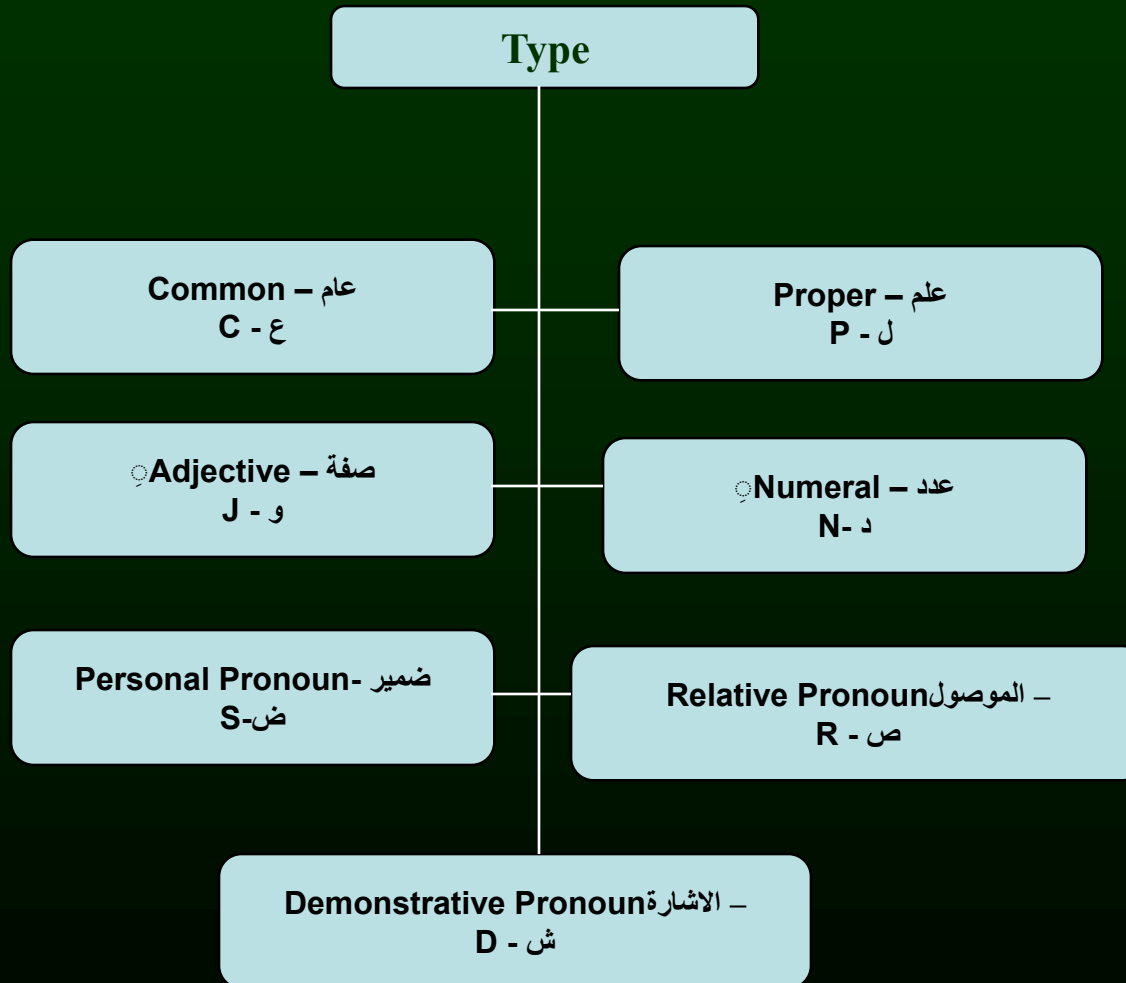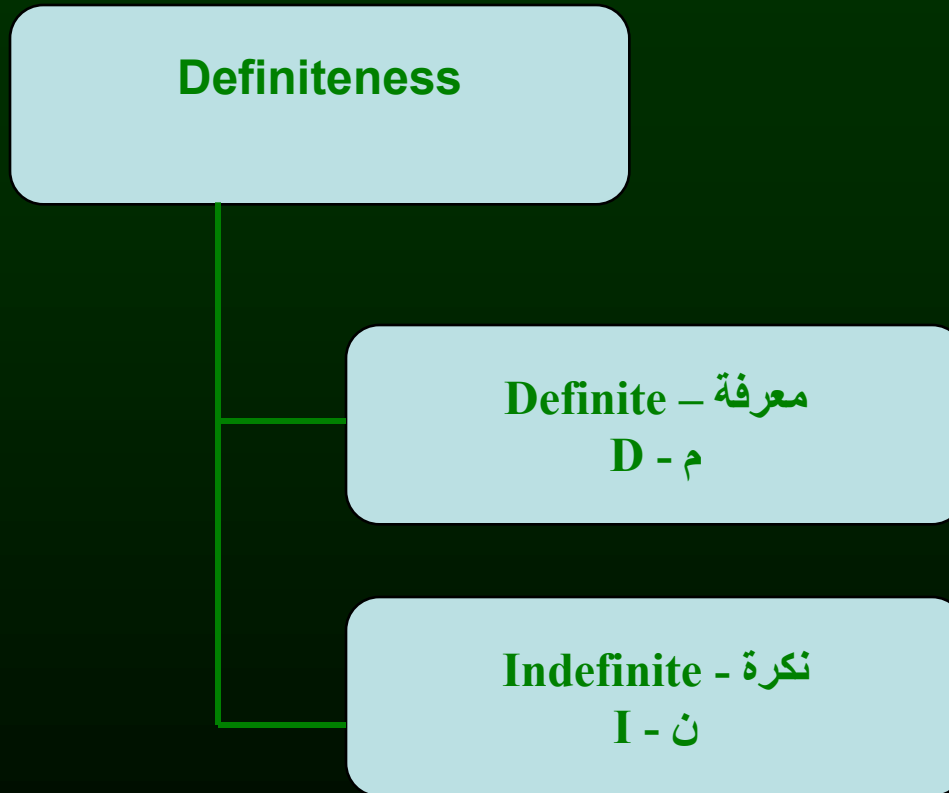
- I. Type
- II. Definiteness
- III. Gender
- IV. Number
- V. Case
- VI. Followship
- VII. Variability
- VIII. Soundness
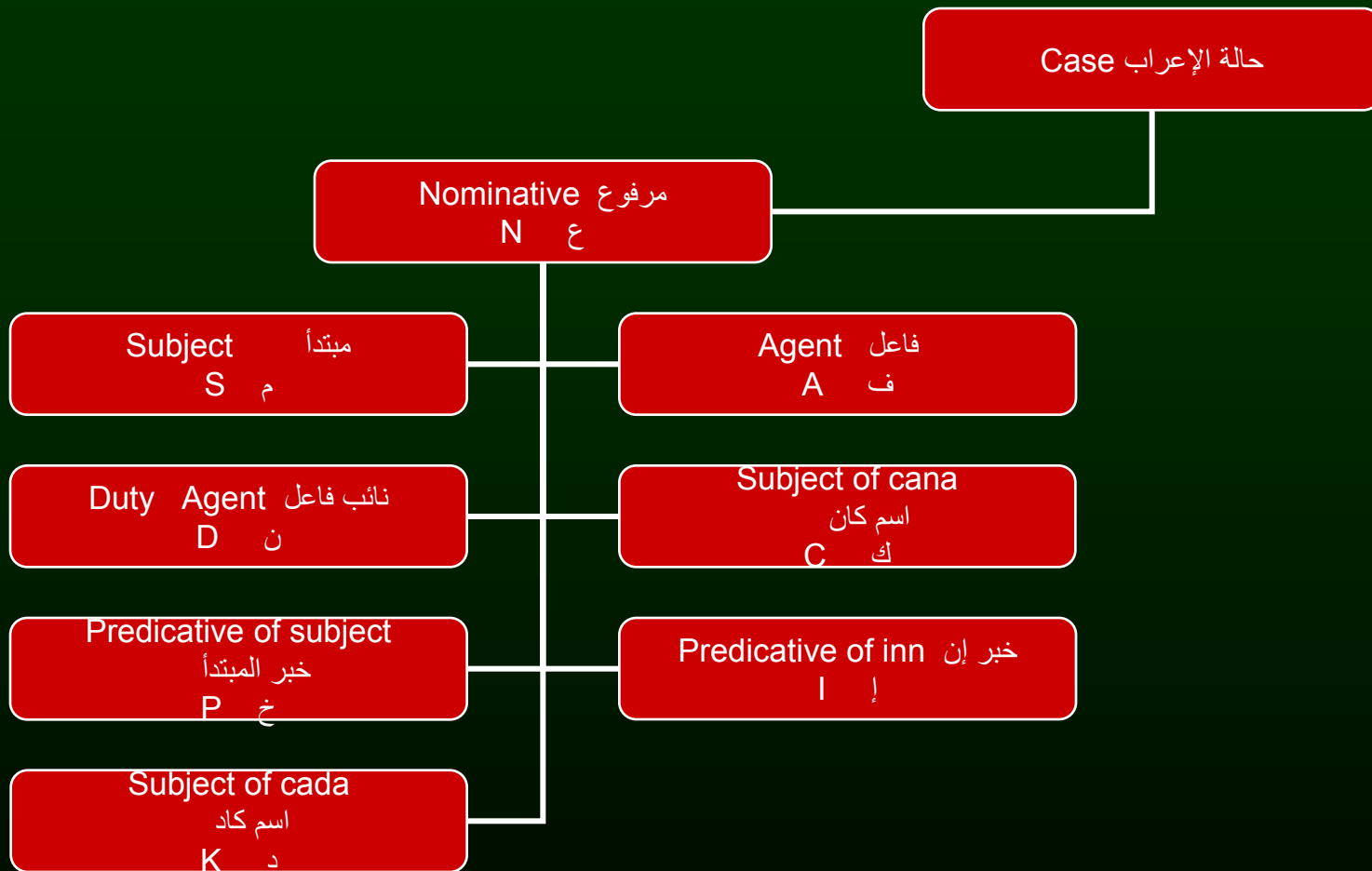
I. Type

```
┌─────────────────────┐
│        Type         │
└─────────────────────┘

┌──────────────────────┐        ┌──────────────────────┐
│   Common – عام       │        │   Proper – علم       │
│      C - ع           │        │      P - ل           │
└──────────────────────┘        └──────────────────────┘

┌──────────────────────┐        ┌──────────────────────┐
│ ○Adjective – صفة     │        │ ○Numeral – عدد       │
│      J - و           │        │      N- د            │
└──────────────────────┘        └──────────────────────┘

┌──────────────────────┐        ┌──────────────────────┐
│ Personal Pronoun- ضمير│        │ الموصول –Relative Pronoun│
│      S-ض             │        │      R - ص           │
└──────────────────────┘        └──────────────────────┘

        ┌────────────────────────────────┐
        │ Demonstrative Pronoun الاشارة – │
        │           D - ش                │
        └────────────────────────────────┘
```

II. Definiteness

Definiteness

Definite – معرفة
D - م

Indefinite - نكرة
I - ن

III. Gender

Gender

مذكر – Masculine
ذ - M
مثل / كتاب

مؤنث Feminine
ث - F
مثل / مدرسة

غير محدد Unmarked
ب U

IV. Number

```
┌─────────────────┐
│     Number      │
└─────────────────┘
         │
         │   ┌──────────────────────┐
         ├───│  Singular – مفرد     │
         │   │      1 - 1           │
         │   └──────────────────────┘
         │   ┌──────────────────────┐
         ├───│  Dual مثنى           │
         │   │      2 - 2           │
         │   └──────────────────────┘
         │   ┌──────────────────────┐
         ├───│  Plural - جمع        │
         │   │    3  -   3          │
         │   └──────────────────────┘
         │            │   ┌──────────────────────┐
         │            ├───│     Sound            │
         │            │   │    S  -  س           │
         │            │   └──────────────────────┘
         │            │   ┌──────────────────────┐
         │            ├───│     Broken           │
         │            │   │    B  -  ت           │
         │            │   └──────────────────────┘
         │            │   ┌──────────────────────┐
         │            └───│     Mass             │
         │                │    M  -  ج           │
         │                └──────────────────────┘
         │   ┌──────────────────────────────┐
         └───│  Unmarked  غير محدد          │
                  │        4  -   3         │
                  └────────────────────────┘
                       │   ┌──────────────────────────────────────┐
                       ├───│  Singular & Dual & Plural (man)       │
                       │   │        A   كـ                         │
                       │   └──────────────────────────────────────┘
                       │   ┌──────────────────────────────────────┐
                       └───│  Dual & Plural (na, nahno)            │
                           │        T  -  ث                        │
                           └──────────────────────────────────────┘
```

V. Case

حالة الإعراب Case

مرفوع Nominative
N ع

Subject مبتدأ S م	Agent فاعل A ف
Duty Agent نائب فاعل D ن	Subject of cana اسم كان C ك
Predicative of subject خبر المبتدأ P خ	Predicative of inn خبر إن I إ
Subject of cada اسم كاد K د	

حالة الإعراب Case

مجرور Genitive
ج G

Post – preposition
مجرور بحروف الجر
ح P

Adjunct (post noun)
مضاف إليه
ض A

حالة الإعراب Case

منادى Vocative
د V

VI. Followship

Followship

Assertion	Coordinated	Attributive	Substitute
توكيد	معطوف	صفة	بدل
A ت	C ط	T ن	S ب

VII. Variability

Variability

Invariable (static)
المبني / المقصور
I م

Variable
معرب
V ع

Semi-Variable
جمع المؤنث السالم
المنقوص / الممنوع من الصرف
S ن

Vowels حروف علة
W ح

Letters
سائر الحروف
L ف

١٠٣

VII. Soundness

Soundness
السلامة

Defective
D   ع

Soundٌ
S   صٌ

Ending with ya
منقوص
Y   ق

Ending with alif + hamza
ممدود
H   د

Ending with alif
مقصور
A   ص

.. Type .. Adjective

الصفة Adjective
و - J

Degree

Positive بسيط
P - ب

Comparative مقرون
C - ق

Superlative تفضيل
S - ت

Type …. Numeral العدد

Numeral عدد N - د

Function

Cardinal أصلي R ح

Ordinal ترتيبي O ع

Numerical adjective A و

Personal Pronoun ... Type   الضمائر

Personal Pronoun
ضمير
ض - S

Person

Attachment

First   متكلم
1   1

Second   مخاطب
2   2

Third   غائب
3   3

Attached
متصل
ص   T

Detached
منفصل
ف   D

Relative Pronoun
الأسماء الموصولة
ص - R

Type

Specific
د - F
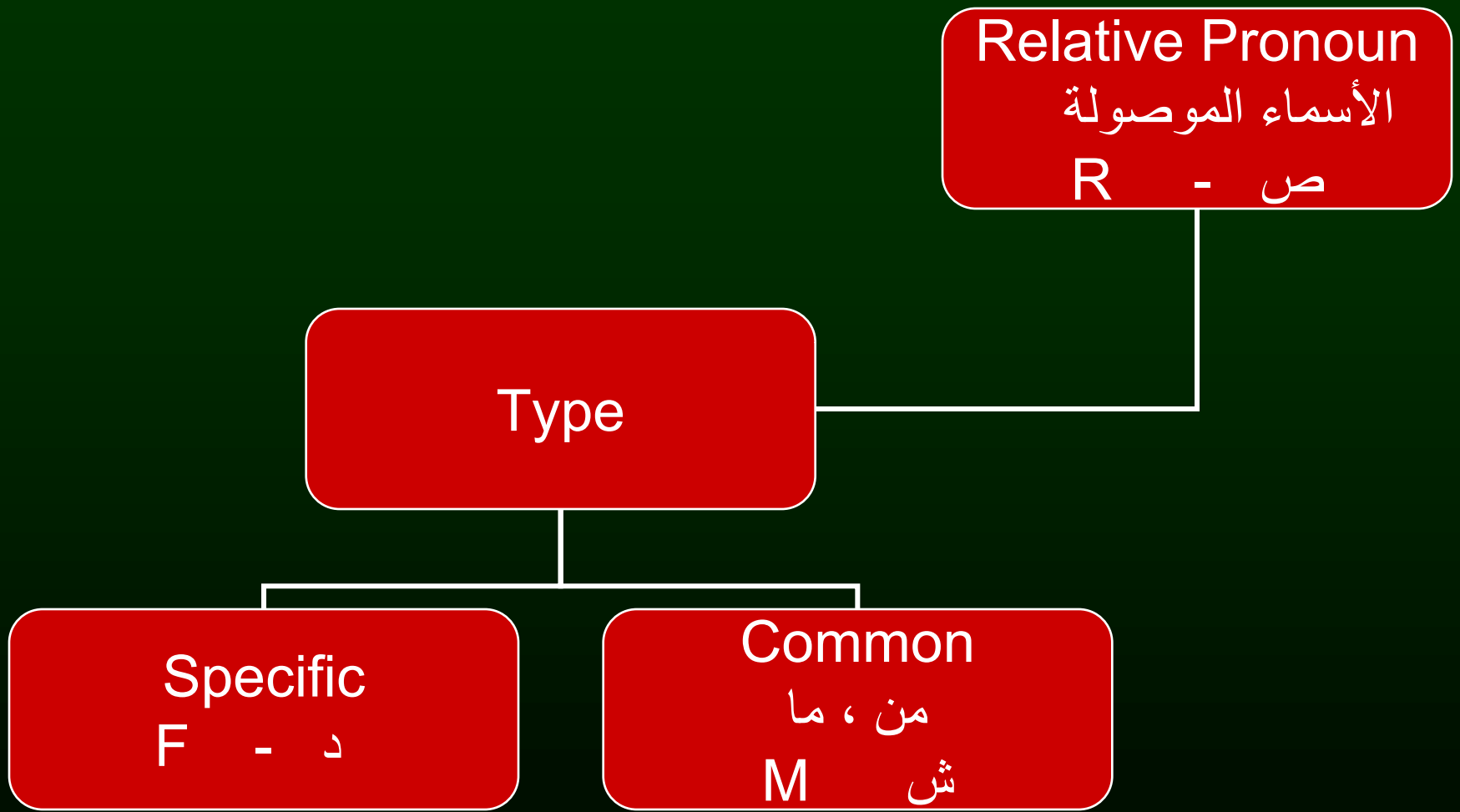
Common
من ، ما
ش M

Example

- فتحت المدرسة أبوابها

- المدرسة< Noun , Common, Definite , Feminine, Singular , Nominative (Agent) , Φ , Variable- Vowels, Sound>

<N-C-D-F-1-NA-Φ-VW-S>

- أبواب
  < N –C- I – F- 3B – AP- Φ - V W- S>

...

- ها < Noun , Personal Pronoun ,Definite , Feminine , Singular , Genitive post –noun (Adjunct), Φ , Invariable (static) , Φ, Third , attached

- < N – S – D – F – 1 – GA – Φ – I – Φ – 3 – T >

2. Adverbs

Adverbs
الظروف
D   ظ

Aspect ِ   Case

Time   Place   Nominative   Accusative   Genitive
T   ز   P   ك
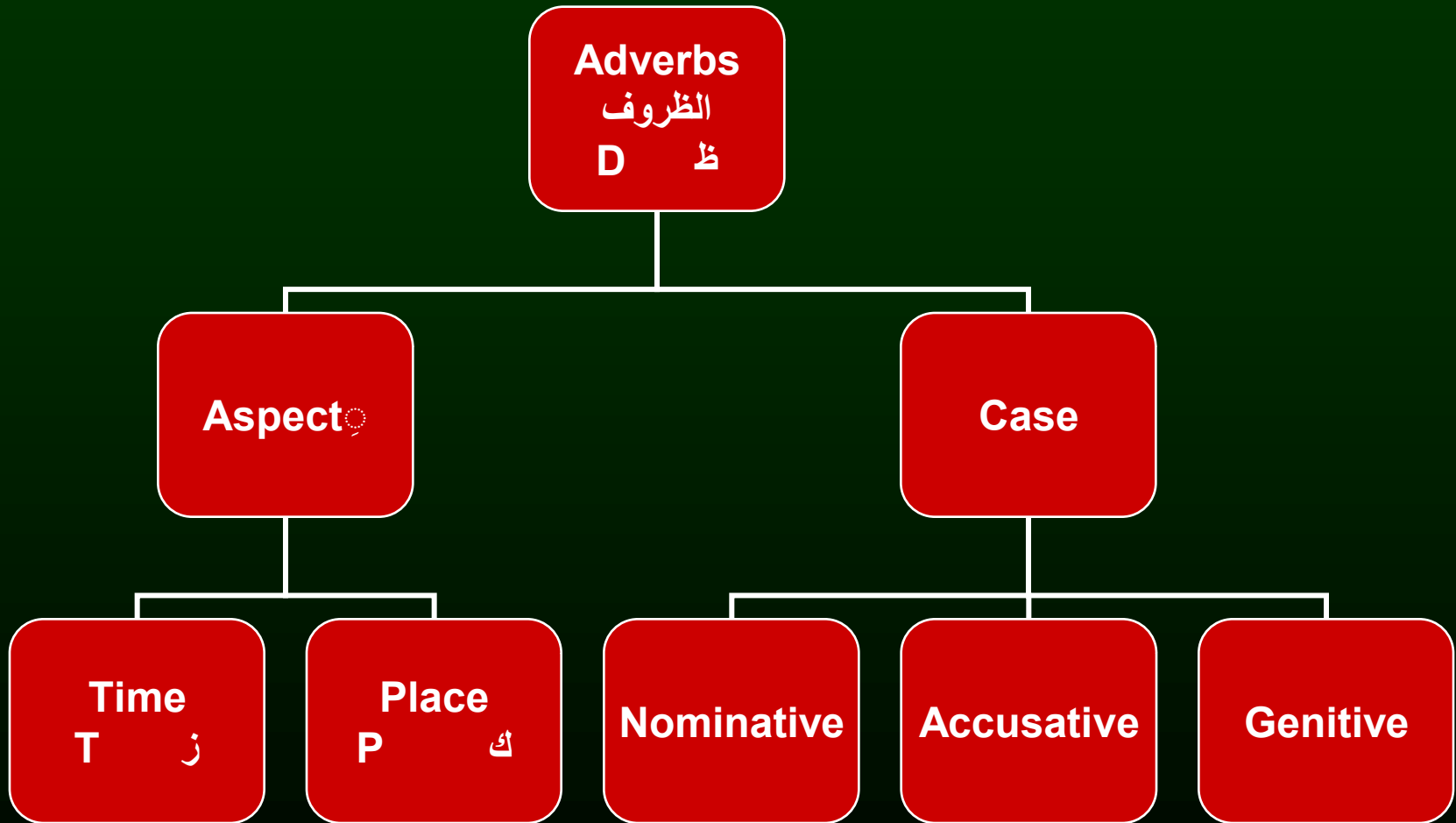
- مثال
  - اتجه إلى يمين الشارع
  - يمين< Adverb , Place ,Genitive >
    <D-P-G>

3. Verbs

	Verbs الأفعال **V ف**
I. Tens (Aspect)	**II. Gender**
III. number	**IV. Person**
V. Case	**VI. Conditional**
VII. Voice	**VIII. Variability**
IX. Perfectness	**X. Augmentation**
XI. Amount	**XII. Soundness**
XIII. Transitivity	
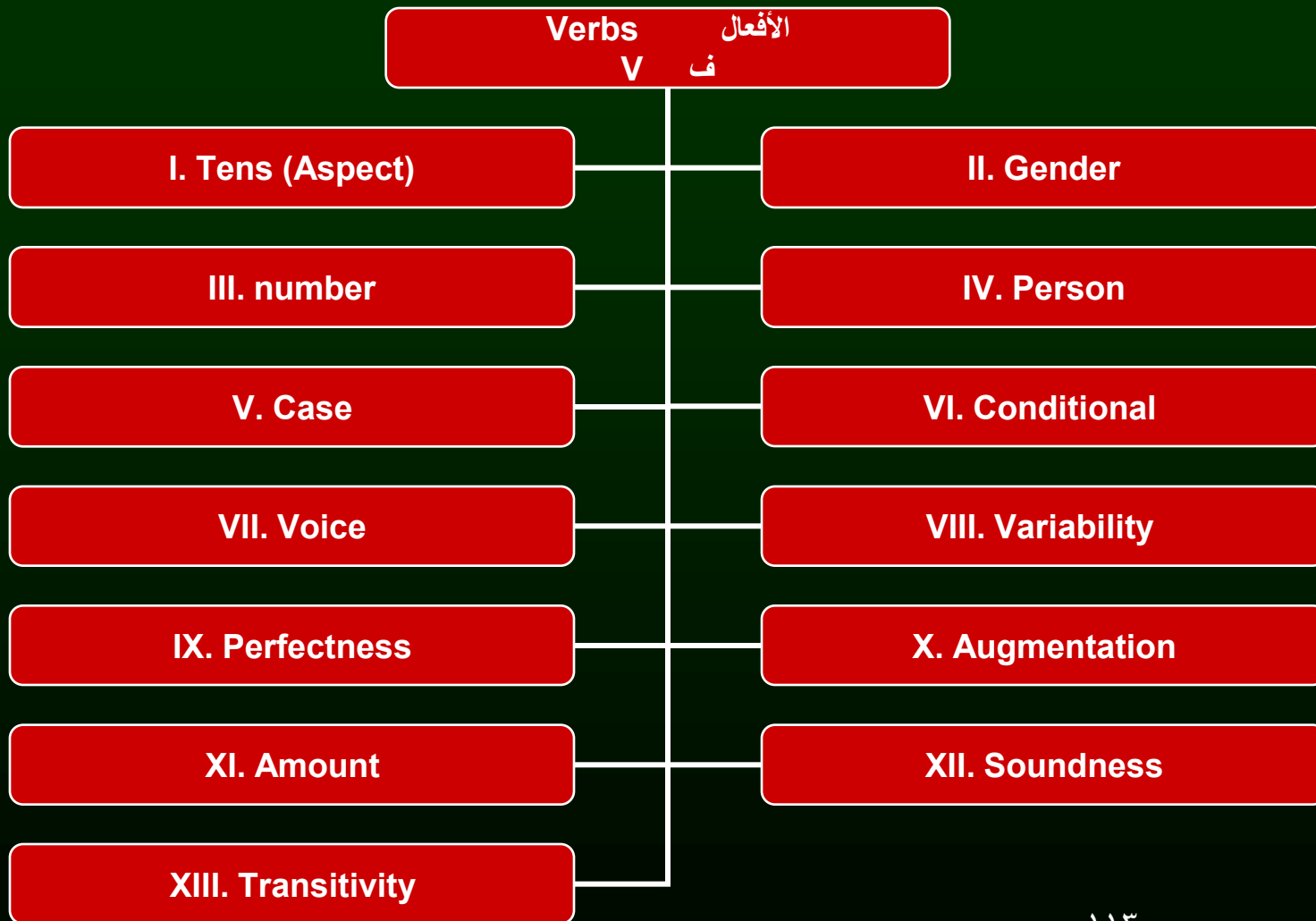
I. Tense

1. Past (P-   م   )
2. Present (Durative \ Future).(R - ح)
3. Imperative (I   -   أ)

II. Gender

1. Masculine (M   -   ذ   )
2. Feminine (F - ث)
3. Unmarked (U   -   ب)

III. Number

- Singular (1-1)
- Dual (2   -2)
- Plural (3-3)
- Unmarked (4-4)
  - Singular & Dual & Plural : verb of (man)
    (A   ك)
  - Dual & Plural : verb of (ma , nahno)
    (T   - ث)

IV. Person

1. First (1-1).
2. Second (2-2).
3. Third (3-3).

V. Case

1. Indicative (مرفوع أو مضموم) (N - ع)
2. Subjunctive (منصوب او مفتوح) (A - ص)
   - Infinitive (مؤول بمصدر) (F - ص)
   - Non – Infinitive (N - غ)
3. Jussive (مجزوم أو مبني) (G - ج)

VI. Conditional

1. The condition (C - ش)
2. The answer (A - ج)

V. Voice

1. Active مبني للمعلوم (A - ع)
2. Passive مبني للمجهول (P - ج)

VIII. Variability

1. Invariable (Static) (I - م)
2. Variable (V - ع)
   - Vowels (W - ح).
   - Letters (L - ف)

IX. Perfectness

1. Perfect تام(P - ت)
2. Imperfect ناقص (Can and cada) (I – ن)

X. Augmentation

- Augmentation زائد (A   -   ز)
- Non – Augmentation (N   -   ج)

XI. Amount

- Trilateral ثلاثي (T -   ث)
- Quadric-Literal   رباعي (Q- ر).
- **Penta – Literal خماسي(P-خ).**

XII. Soundness

- Defective   (D   -   ع):
  - Initial   في البداية (I   -   ث)
    - مثل / يبس
  - Hollow (Meddle) (H-خ)
    - مثل / ضاع
  - Last   (L - ن)
    - مثل / نسى
  - Initial + last   (T   -   ف)
    - مثل / وعى
  - Hollow + Last (O - ق)
    - مثل / حوى

- Sound   (S   - ص)

XIII. Transitivity

- Transitive   متعدي   (T - ت)
    - One Patient   مفعول واحد (O - ح)
        - مثل / أكل
    - Two Patient   مفعولين (T - ث)
        - مثل / أعطى

- Intransitive   لازم(I   -   ل)
    - Agent only   فاعل فقط (A - ف)
    - Agent + State or Distinquitor (S - ض)
    - Nominal   Sentence   (N - ج)

- مثال /

  - نامت الطفلة

نامت <Verb , Past , Feminine , Singular , Third , Subjunctive non infinitive , Φ , Active , Invariable (static) , Perfect , Augmented , Trilateral ,Sound, intransitive Agent only >

<V – P – F – 1- 3 – A N – Φ – A – I – P- A – T – S – I A >

4. Particles   أدوات
أ   P

- Coordinating   حروف العطف   (1   -   1)
  و / ف / ثم / أو / حتى لو / حتى / حتى /حتى ولو

- Subordinating   (2   -   2)
  - Contrast (بل/ لكنّ/ لكن)   (ض - c)
  - Exception (عدا / ما عدا / سوى / غير / إلا) (E - س)
  - Initial (الإبتدائية فـ / و) (I - ت) .

- Interrogative (أ / هل) (3   -   3) .

- Preposition (حروف الجر) (4 - 4) .

...

- Possibility (قد قبل المضارع) (5).
- Protection (نون الوقاية) (6) .
- Future (سـ / سوف) (7).
- Conditional (ما / مهما / إذا / أيّ)(8).
- Answer (نعم / أجل / لا / إذن)(9).
- Exclamation (ما أو إنفعال)(10).
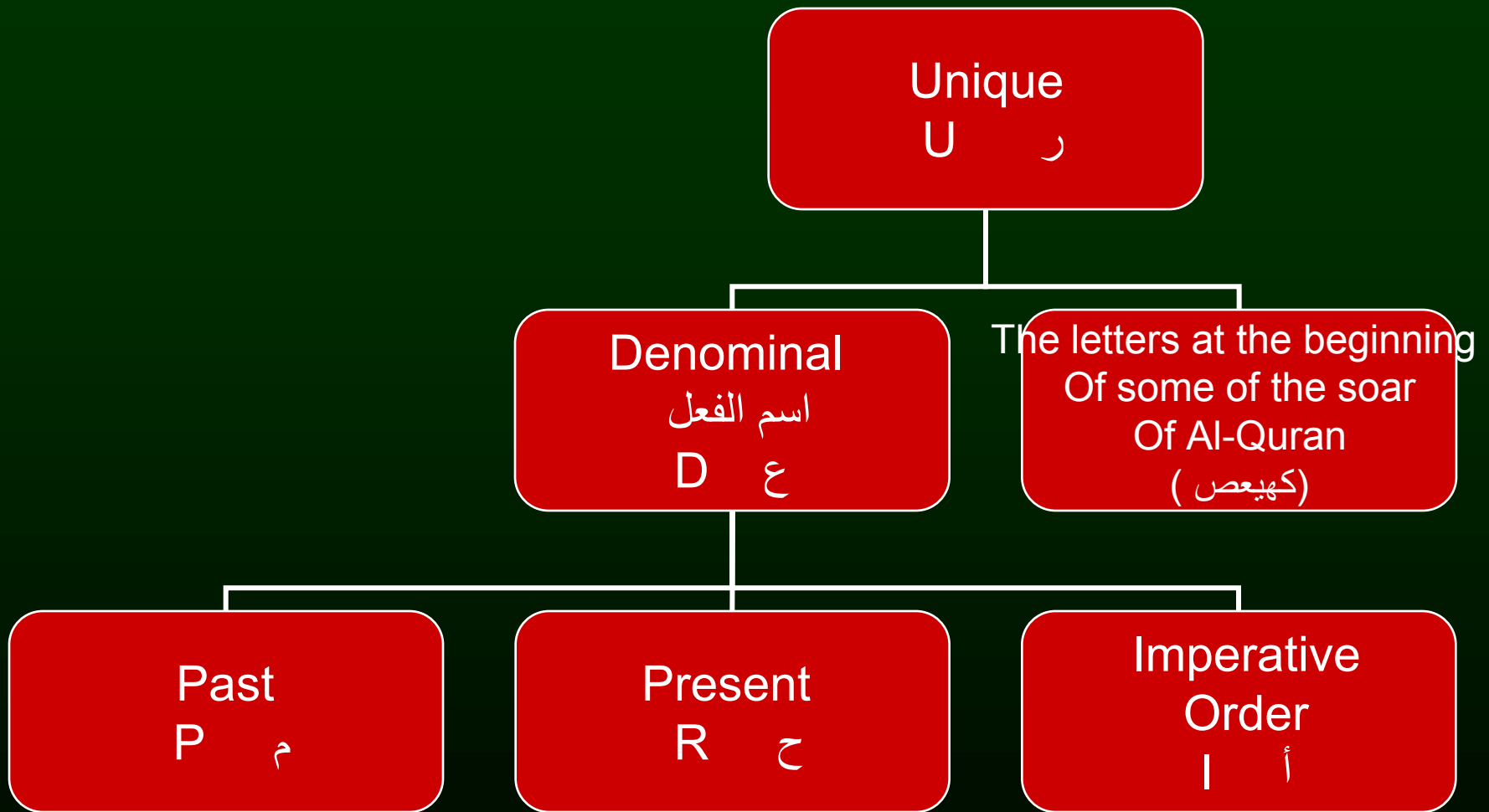- 11.Interjection/Introgative (يا / ايها / أيتها) (11).

- Negative (ليس / لا / لم / لن)(12).
- Imperative (Order)(لـ) (13).
- Cause (لـ / حتى / لأجل / من أجل / لكي)(14).
- Gerund (أن / أنّ)(15).
- Deporticle (اخوات إن)(16).
- (17) تاء التأنيث.
- Explanation (أي)(18).

- Assertion (إنّ / أنّ / نونا التوكيد / قد و لقد على الماضي)
  (19).
- Wishing (لعل / ليت) . (20)
- Swearness (و القسم)(21).

- مثال /
  - نامت الطفلة
  - تاء التأنيث < Particles , *ta* of Femininity >

  <P - 17>.
  - قال تعالى {وَالْمُرْسَلاتِ عُرْفاً }
  - و <Particles , Swearness >
     <P - 21>

5. Unique (U ر -)

Unique
U ر

Denominal
اسم الفعل
D ع

The letters at the beginning
Of some of the soar
Of Al-Quran
(كهيعص)

Past
P م

Present
R ح

Imperative
Order
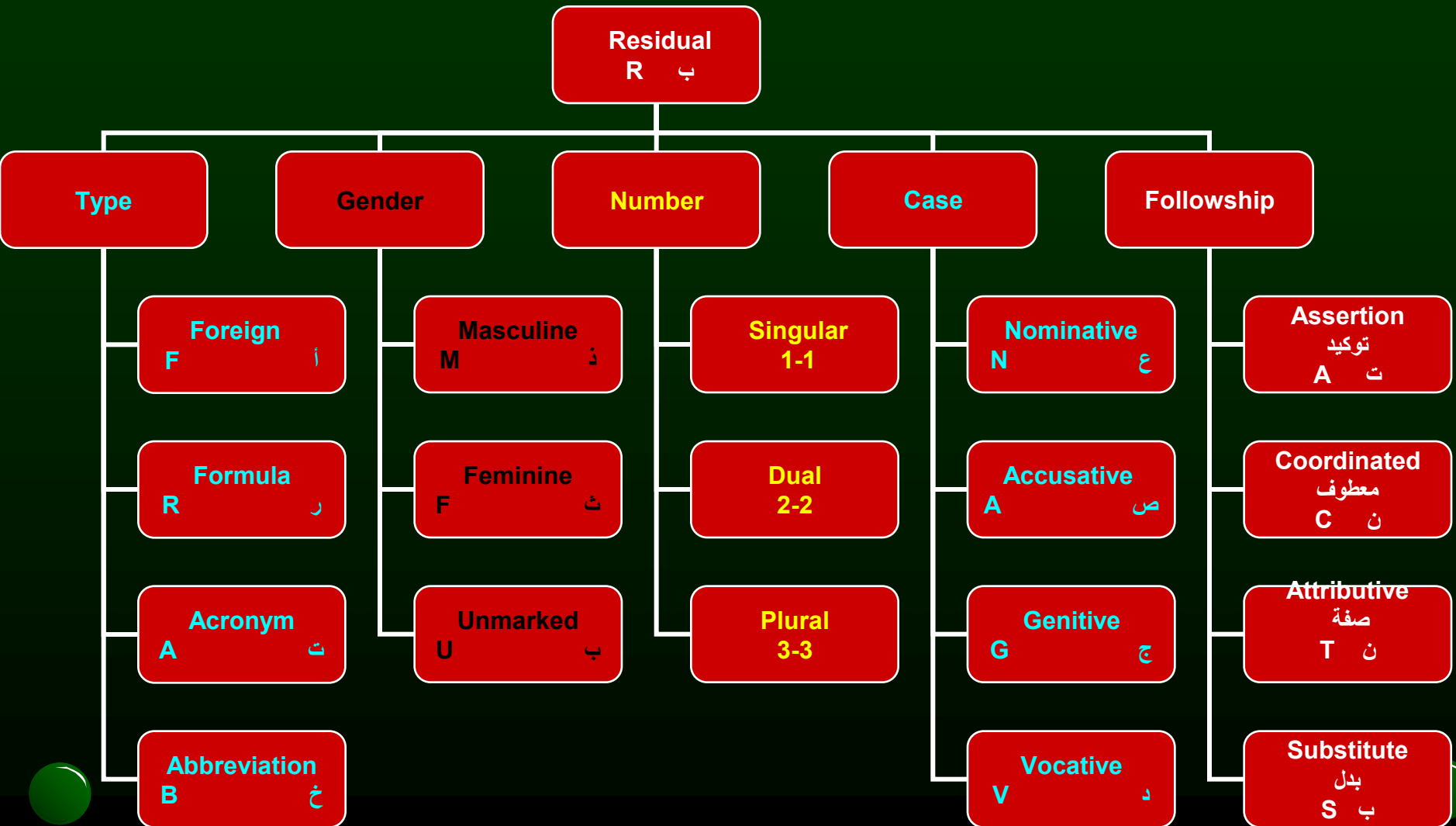I أ

١٢٨

- مثال

  قال تعالى {حم (١) تَنزِيلُ الْكِتَابِ مِنْ اللَّهِ الْعَزِيزِ الْعَلِيمِ (٢) }

- حم< Unique , The litters at the beginning Of some of the soar Of Al-Quran >

  - <U - L>

6. Residual

Residual
R ب

Type

Gender

Number

Case

Followship

Foreign
F أ

Formula
R ر

Acronym
A ت

Abbreviation
B خ

Masculine
M ذ

Feminine
F ث

Unmarked
U ب

Singular
1-1

Dual
2-2

Plural
3-3

Nominative
N ع

Accusative
A ص

Genitive
G ج

Vocative
V د

Assertion
توكيد
A ت

Coordinated
معطوف
C ن

Attributive
صفة
T ن

Substitute
بدل
S ب

مثال /

تلعب شركة أرامكو دورا كبيرا في اقتصاد الدولة

أرامكو <Residual , Foreign , Feminine , Singular , Genitive Adjunct (post noun) > <R – F – F – 1 – G A>

7. Punctuation (c – ت)

- ?   Question Mark **(Q- س)**.
- !   Exclamation Mark **(X - ت)**.
- …   Ellipsis (**E – ح**).
- .   Full Stop **(F- ن)**.
- ،   Comma **(C- ف)**.
- ;   Dotted Comma **(D-ق)**.
- -   Hyphen **(H-ش)**.

- -   Interspersion Marks **(I- ع).**
- ,   The English Comma **(G- ج).**
- , , Interspersion Marks **(R- ق).**
- ( )   Brackets **(B- هـ).**
- "   Quotation Marks **(U- ب).**
- :   Colon **(O- ر).**
- [ ] Square Brackets **(S- و).**
- {}
- / slash **(L- م).**

- مثال /

  - فتحت المدرسةُ أبوابَها .

    - \<Punctuation , Full Stop\>

    \<C - F\>

To Part 2: Arabic POS

- السلام عليكم ورحمة الله