

# Computational Linguistics

CSC 485/2501  
Fall 2021

# 5A

## 5a. Extending grammars with features

Gerald Penn

Department of Computer Science, University of Toronto

Reading: Jurafsky & Martin: 12.3.4–6, 15.0–3;  
[Allen: 4.1–5]; Bird et al: 9.

Copyright © 2017 Suzanne  
Stevenson, Graeme Hirst  
and Gerald Penn. All rights  
reserved.

# Agreement and inflection

- Problem: **Agreement** phenomena.

*Nadia {washes/\***wash**} the dog.*

*The boys {\***washes**/wash} the dog.*

*You {\***washes**/wash} the dog.*

- **Morphological inflection** of verb must match subject noun in person and number.

# Subject–verb agreement 1

## Present tense

	Singular	Plural
1	<i>I wash</i>	<i>we wash</i>
2	<i>you wash</i>	<i>you wash</i>
3	<i>he/she/it washes</i>	<i>they wash</i>
1	<i>I am</i>	<i>we are</i>
2	<i>you are</i>	<i>you are</i>
3	<i>he, she, it is</i>	<i>they are</i>

# Subject–verb agreement 2

## Past tense

	<b>Singular</b>	<b>Plural</b>
<b>1</b>	<i>I washed</i>	<i>we washed</i>
<b>2</b>	<i>you washed</i>	<i>you washed</i>
<b>3</b>	<i>he, she, it washed</i>	<i>they washed</i>
<b>1</b>	<i>I was</i>	<i>we were</i>
<b>2</b>	<i>you were</i>	<i>you were</i>
<b>3</b>	<i>he, she, it was</i>	<i>they were</i>

# Agreement features 1

- English agreement rules are fairly simple.
  - Subject : verb w.r.t. person and number.
  - No agreement required between verb and object.
- Many languages have other agreements.
  - *E.g.*, German: Article and adjective ending depends on noun gender and case:

# Agreement features 2

## Nominative Case (Subject Case)

Masculine	Feminine	Neuter	Plural
<b>der</b>	<b>die</b>	<b>das</b>	<b>die</b>

der neu**e** Wagen

the new car

die schön**e** Stadt

the beautiful city

das alte**e** Auto

the old car

die neu**en** Bücher

the new books

Masculine	Feminine	Neuter	Plural
<b>ein</b>	<b>eine</b>	<b>ein</b>	<b>keine</b>

ein neu**er** Wagen

a new car

eine schön**e** Stadt

a beautiful city

ein alte**s** Auto

an old car

keine neu**en** Bücher

no new books

# Agreement features 2

## Accusative Case (Direct Object)

Masculine	Feminine	Neuter	Plural
<b>den</b>	<b>die</b>	<b>das</b>	<b>die</b>

den neu**en** Wagen  
the new car

die schön**e** Stadt  
the beautiful city

das alte**e** Auto  
the old car

die neu**en** Bücher  
the new books

Masculine	Feminine	Neuter	Plural
<b>einen</b>	<b>eine</b>	<b>ein</b>	<b>keine</b>

einen neu**en** Wagen  
a new car

eine schön**e** Stadt  
a beautiful city

ein alte**s** Auto  
an old car

keine neu**en** Bücher  
no new books

# Agreement features 3

*E.g.*, Chinese: Numeral classifiers, often based on shape, aggregation, ....:

两条鱼	<i>liang tiao yu</i> ‘two CLASSIF-LONG-ROPELIKE fish’
两条河	<i>liang tiao he</i> ‘two CLASSIF-LONG-ROPELIKE rivers’
两条腿	<i>liang tiao tui</i> ‘two CLASSIF-LONG-ROPELIKE legs’
两条裤子	<i>liang tiao kuzi</i> ‘two CLASSIF-LONG-ROPELIKE pants’
两只胳膊	<i>liang zhi gebo</i> ‘two CLASSIF-GENERAL arms’
两件上衣	<i>liang jian shangyi</i> ‘two CLASSIF-CLOTHES-ABOVE-WAIST tops’
两套西装	<i>liang tao xizhuang</i> ‘two CLASSIF-SET suits’

Zhang, Hong (2007). Numeral classifiers in Mandarin Chinese. *Journal of East Asian Linguistics*, 16(1), 43–59. Thanks also to Tong Wang, Vanessa Wei Feng, and Helena Hong Gao.



# Agreement features 1

- English agreement rules are fairly simple.
- Many languages have other agreements.
- Some languages have multiple grammatical genders.
  - E.g. Chichewa has genders for men, women, bridges, houses, diminutives, men inside houses, etc. Between 12-18 in total.
- Some languages overtly realize many of these distinctions.
  - E.g. some Hungarian verbs have as many as 4096 inflected forms.

# Inflectional morphology

- Word may be inflected ...
  - ... to indicate paradigmatic properties, e.g. singular / plural, past / present, ...
  - ... to indicate some (other) semantic properties
  - ... to agree with inflection of other words.
- Each (open-class) word-type has a **base form / stem / lemma**.
- Each occurrence of a word includes inflection by a (possibly null) morphological change.

# Rule proliferation 1

- **Problem:** How to account for this in grammar.
- **Possible solution:** Replace all NPs, Vs, and VPs throughout the grammar.

$S \rightarrow NP VP$

$NP \rightarrow you, dog, dogs, bear, bears,$

...

$VP \rightarrow V NP$

$V \rightarrow washes, wash, washed, is,$

$was, \dots$



$S \rightarrow NP_{3s} VP_{3s}$

$S \rightarrow NP_{3p} VP_{3p}$

$S \rightarrow NP_2 VP_2$

$S \rightarrow NP_{1s} VP_{1s}$

$S \rightarrow NP_{1p} VP_{1p}$

$NP_2 \rightarrow you$

$\vdots$

$VP_{3s} \rightarrow V_{3s} NP$

$\vdots$

$V_{3s} \rightarrow is, was,$   
 $washes, washed, \dots$

$V_{1s} \rightarrow am, was, wash,$

$washed, \dots$

$\vdots$

$NP_{3s} \rightarrow dog, bear, \dots$

$NP_{3p} \rightarrow dogs, bears$

$V_{3p} \rightarrow are, were,$

$wash, washed, \dots$

# Rule proliferation 2

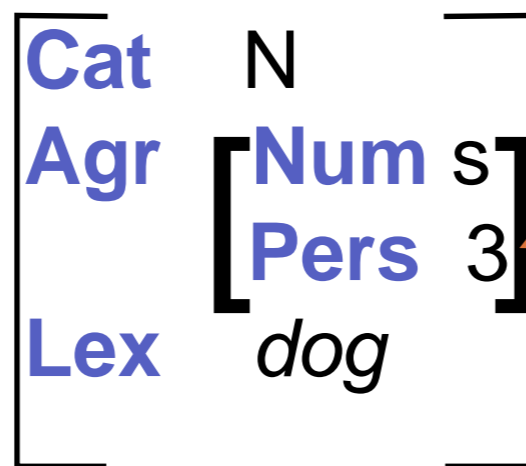
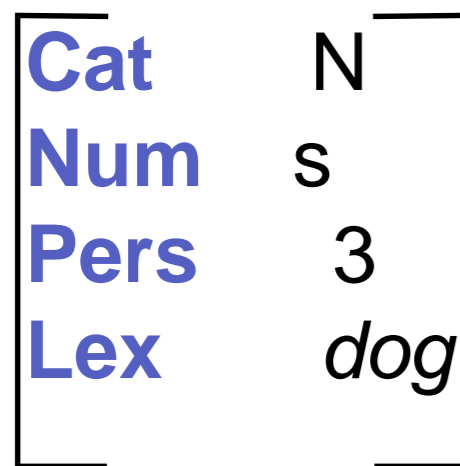
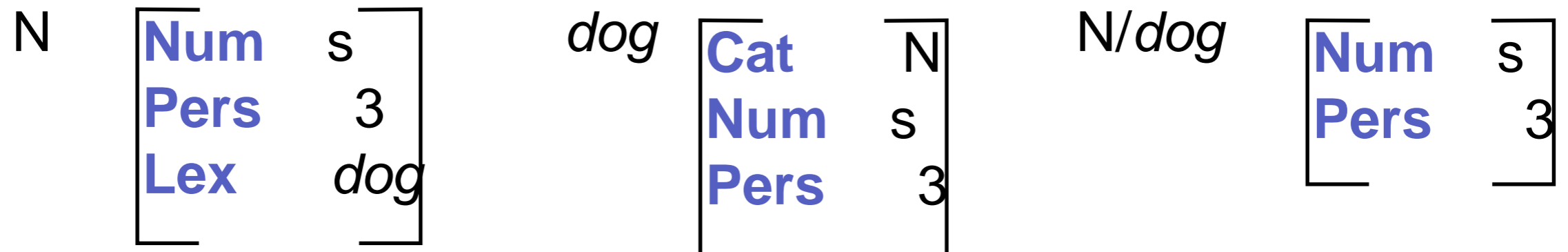
- *Drawback 1:* the result is big ... really big.
- *Drawback 2:* Losing the generalization:
  - All these Ss, NPs, VPs have the same structure.
  - Doesn't depend on particular verb, noun, and number.
- CF rules collapse together structural and featural information.
- All information must be completely and directly specified.
  - *E.g.*, can't just say that values must be equal for some feature without saying exactly what values.

# Feature structures 1

- *Solution:* Separate feature information from syntactic, structural, and lexical information.
- A **feature structure** is a list of pairs:  
    [*feature-name feature-value*]
- Feature-values may be atoms or feature structures.
- Can consider syntactic category or word to be bundle of features too.
- Can represent syntactic structure.

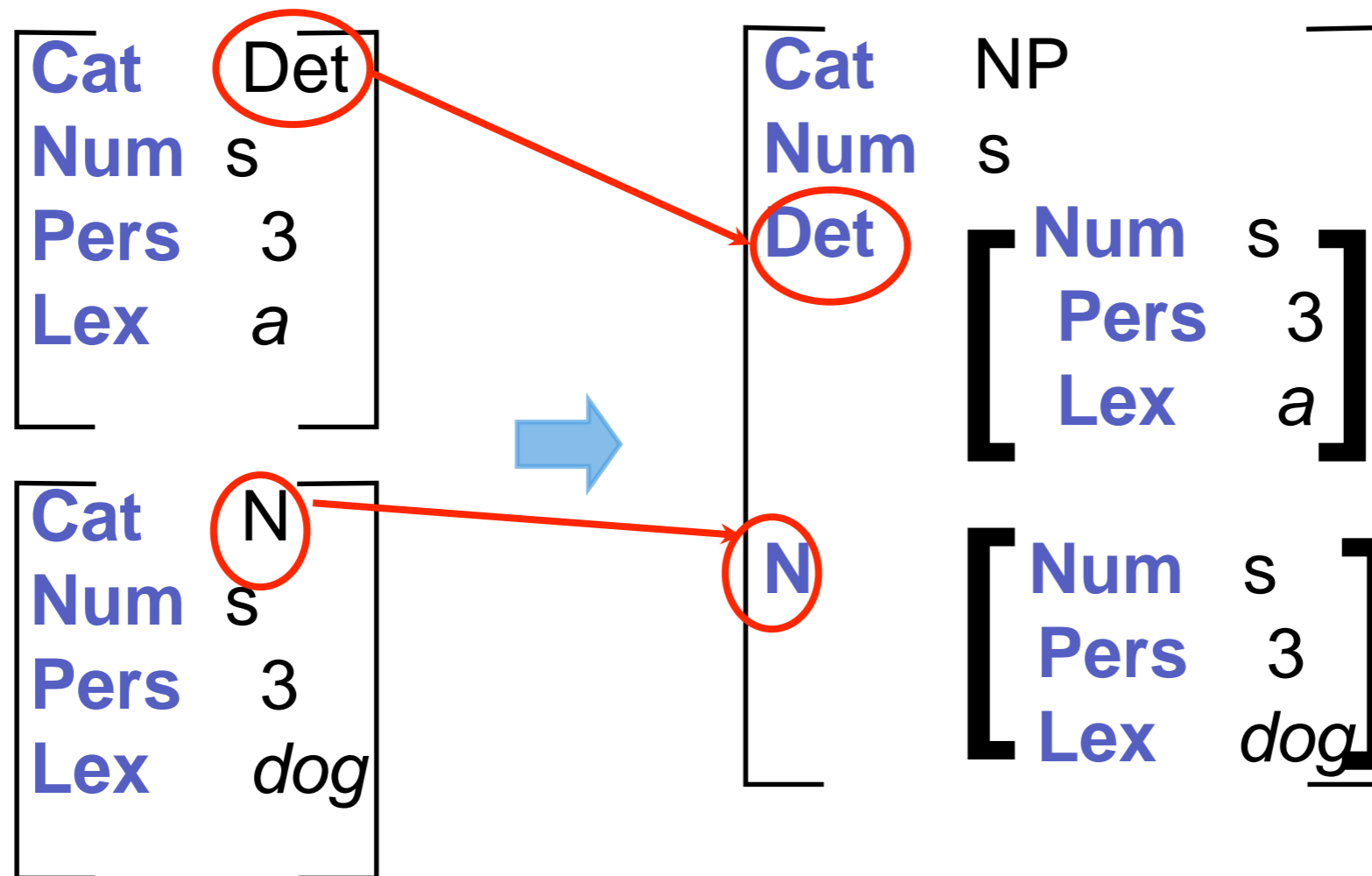
# Feature structures 2

- *Drawback:* many equivalent notations.



*Feature paths:*  
features of  
features; e.g.,  
(Agr Pers 3)

# Feature structures 3



NP formed from Det and N.  
Feature values in components become  
feature names in new constituent.

# Components of feature use

- **1. Lexical specification:**

Description of *properties* of a word:  
morphological, syntactic, semantic, ...

*dog*: [Cat N  
Agr 3s]

*sleeps*: [Cat V  
Agr 3s]

*dogs*: [Cat N  
Agr 3p]

*sleep*: [Cat V  
Agr {1s,2s,1p,2p,3p}]

Or: ^3s

Or: N → *dog*  
(N Agr) = 3s

V → *sleeps*  
(V Agr) = 3s

N → *dogs*  
(N Agr) = 3p

V → *sleep*  
(V Agr) = {1s,2s,1p,2p,3p}



# Components of feature use

- **2. Agreement:**
  - **Constraints** on co-occurrence in a rule — within or across phrases.
  - Typically are equational constraints.

NP  $\rightarrow$  Det N  
(Det **Num**) = (N **Num**)

S  $\rightarrow$  NP VP  
(NP **Agr**) = (VP **Agr**)

# Components of feature use

- **3. Projection:**

- ***Sharing of features*** between the head of a phrase and the phrase itself.

$$\begin{aligned} VP &\rightarrow V \dots \\ (VP \text{ Agr}) &= (V \text{ Agr}) \end{aligned}$$

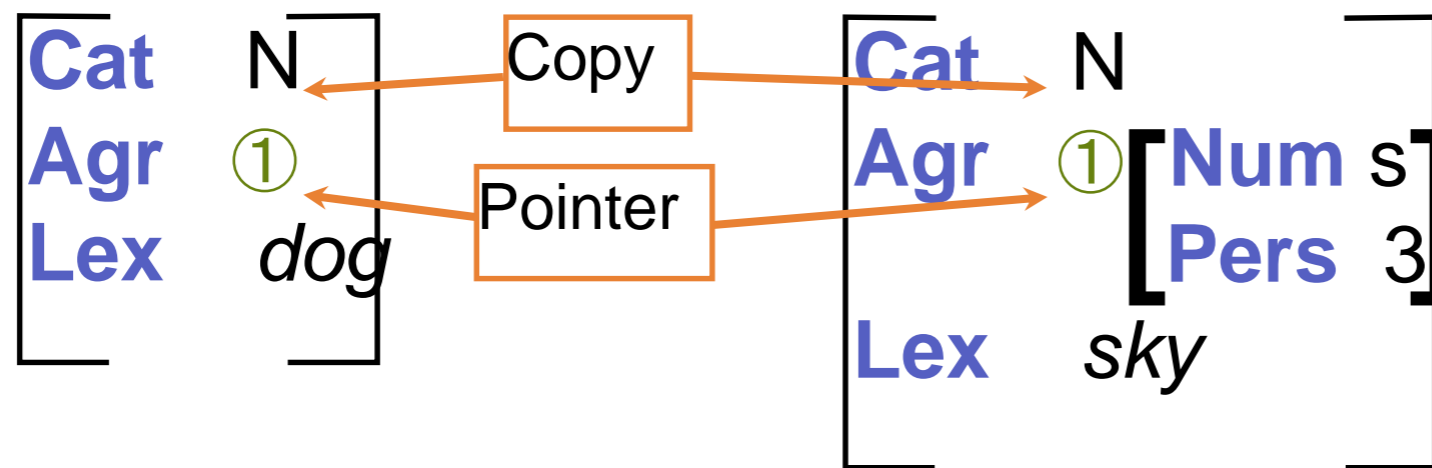
- Head features:

- **Agr** is typical, but so is the head-word itself as a feature.

(Common enough that there's usually a mechanism for "declaring" head features and omitting them from rules.)

# Constraints on feature values 1

- What does it mean for two features to be “equal”?
- A *copy* of the value or feature structure, or a *pointer* to the same value or feature structure (re-entrancy, shared feature paths).



# Constraints on feature values 2

- But: It may be sufficient that two features are not equal, just *compatible* — that they can be *unified*.

- *E.g.*, 

Cat	N
Pers	3
Num	s

 and 

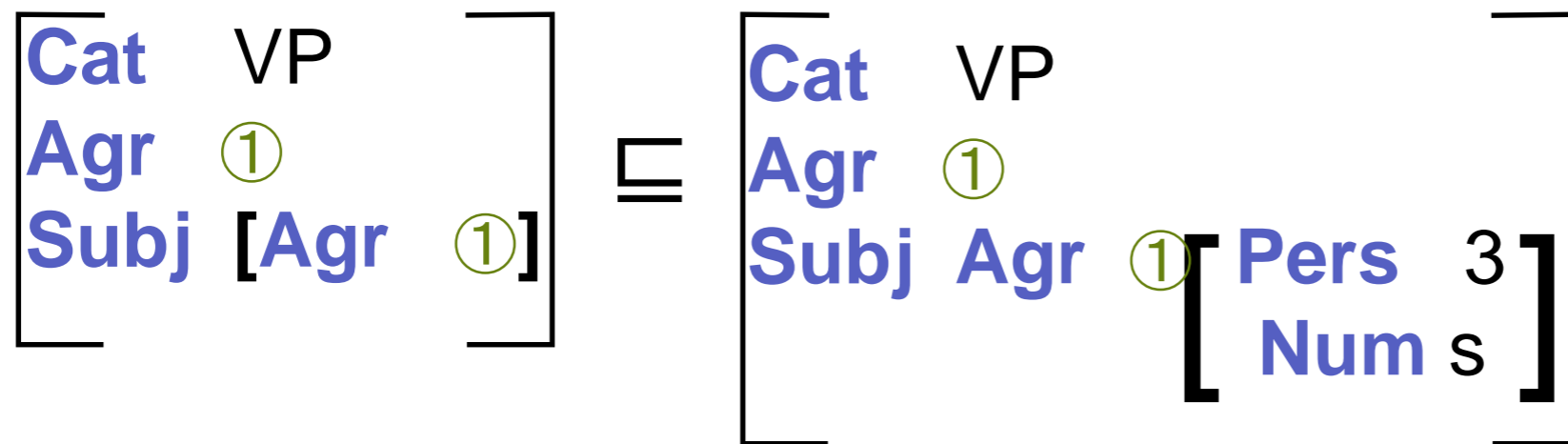
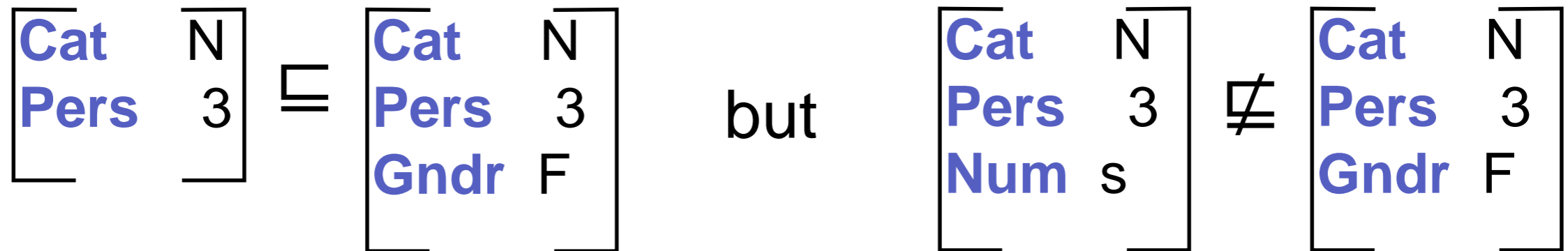
Cat	N
Pers	3
Gndr	F

# Subsumption of feature structures 1

- Feature structure  $X$  *subsumes* feature structure  $Y$  if  $Y$  is consistent with, and at least as specific as  $X$ .
  - Also say that  $Y$  *extends*  $X$ .  
 $Y$  can add (non-contradictory) features to those in  $X$ .
- **Definition:**  $X$  *subsumes*  $Y$  ( $X \sqsubseteq Y$ ) iff there is a *simulation* of  $X$  inside  $Y$ , i.e., a function s.t.:
  - $\text{sim}(X) = Y$
  - If  $X$  is atomic, so is  $Y$  and  $X = Y$
  - Otherwise, for all feature values  $X.f$ :  $Y.f$  is defined, and  $\text{sim}$  simulates  $X.f$  inside  $Y.f$ .

# Subsumption of feature structures 2

- Examples:



# Unification 1

- The *unification* of  $X$  and  $Y$  ( $X \sqcup Y$ ) is the most general feature structure  $Z$  that is subsumed by both  $X$  and  $Y$ .
  - $Z$  is the smallest feature structure that extends both  $X$  and  $Y$ .
- Unification is a constructive operation.
  - If any feature values in  $X$  and  $Y$  are incompatible, it fails.
  - Else it produces a feature structure that includes all the features in  $X$  and all the features in  $Y$ .

# Unification 2

$$\begin{bmatrix} \text{Cat} & \text{N} \\ \text{Pers} & 3 \\ \text{Num} & \text{s} \\ \text{Gndr} & \text{F} \end{bmatrix} \sqcup \begin{bmatrix} \text{Cat} & \text{N} \\ \text{Pers} & 3 \\ \text{Gndr} & \text{F} \end{bmatrix} = \begin{bmatrix} \text{Cat} & \text{N} \\ \text{Pers} & 3 \\ \text{Num} & \text{s} \\ \text{Gndr} & \text{F} \end{bmatrix}$$



# Features in chart parsing

- Each constituent has an associated feature structure.
  - Constituents with children have a feature structure for each child.
- Arc addition:
  - The feature structure of the new arc is initialized with all known constraints.
- Arc extension:
  - The feature structure of the predicted constituent must unify with that of the completed constituent extending the arc.

# Sample grammar fragment

$S \rightarrow NP VP$

$(NP \text{ Agr}) = (VP \text{ Agr})$

$NP \rightarrow Det N$

$(NP \text{ Agr}) = (N \text{ Agr})$

$(Det \text{ Agr}) = (N \text{ Agr})$

$VP \rightarrow V$

$(VP \text{ Agr}) = (V \text{ Agr})$

$Det \rightarrow a$

$[Agr \ 3s]$

$Det \rightarrow all$

$[Agr \ 3p]$

$Det \rightarrow the$

$[Agr \ \{3s,3p\}]$

$N \rightarrow dog$

$[Agr \ 3s]$

$N \rightarrow dogs$

$[Agr \ 3p]$

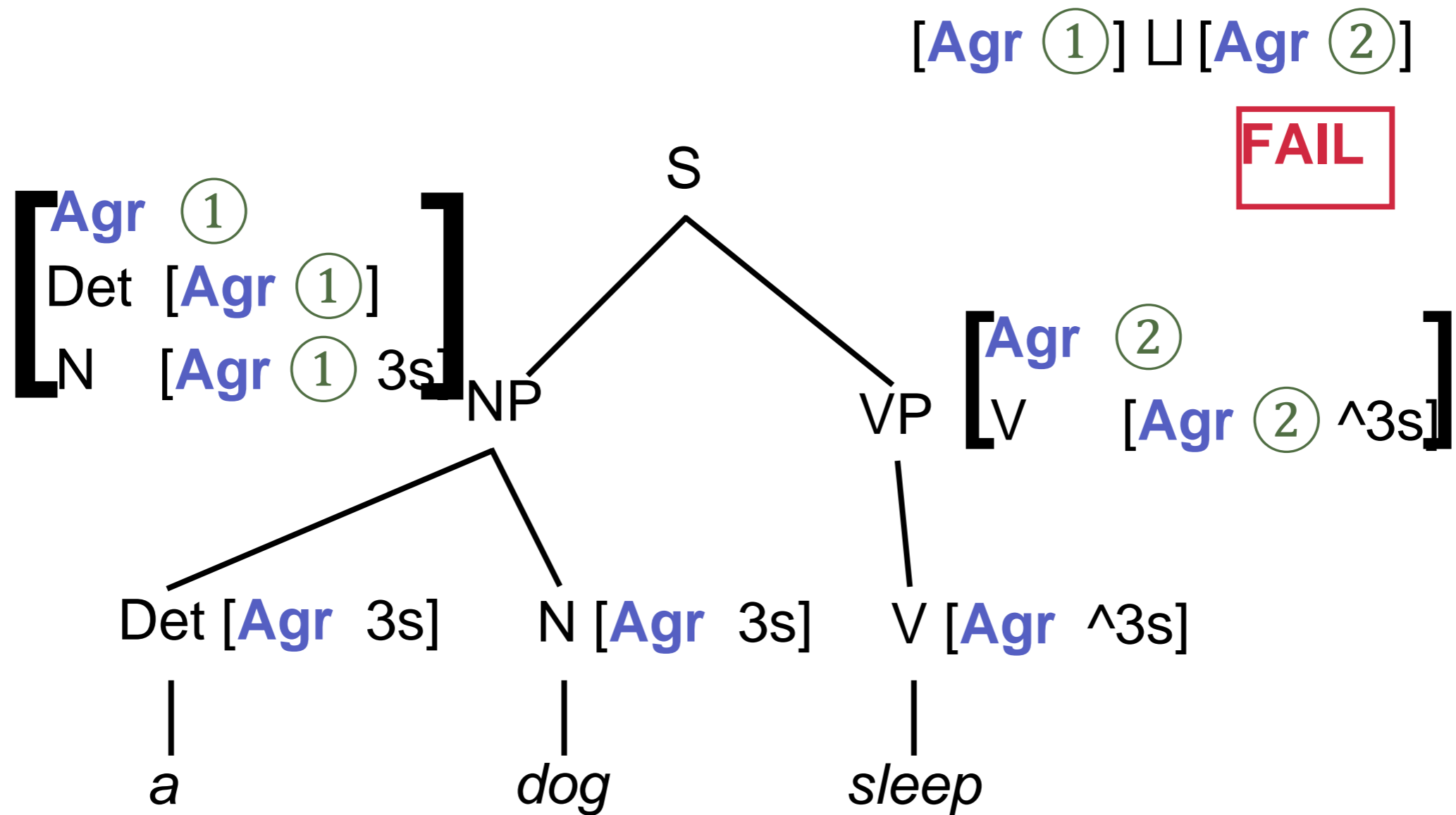
$V \rightarrow sleep$

$[Agr \ ^3s]$

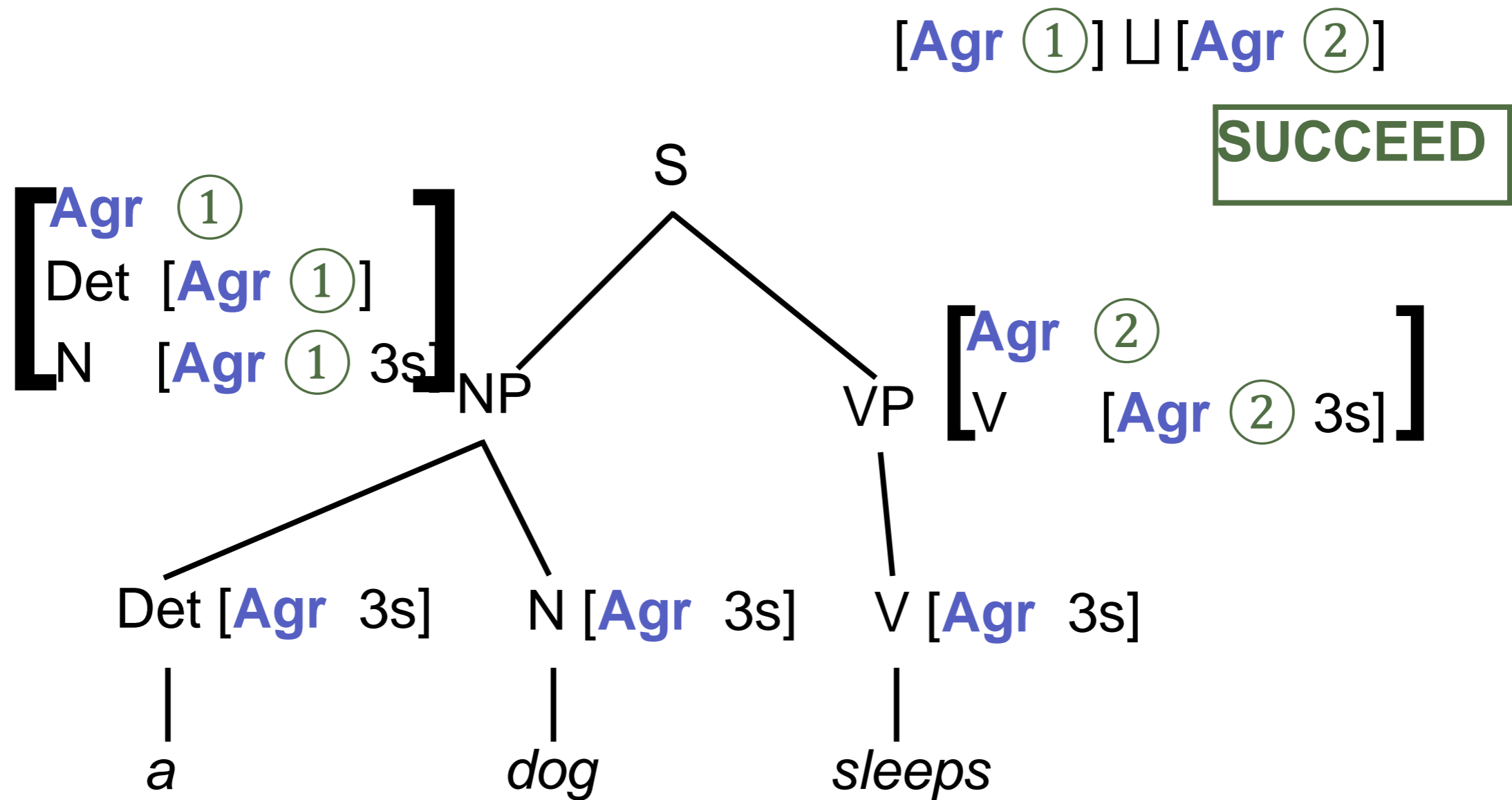
$V \rightarrow sleeps$

$[Agr \ 3s]$

# Mismatched features fail




# Unifiable features succeed



# Advantages of this approach

- Distinguishes structure from "functional" info.
- Allows for economy of specification:
  - Equations in rules:  
$$S \rightarrow NP \ VP$$
$$(NP \text{ Agr}) = (VP \text{ Agr})$$

Must unify with


  - Sets of values in lexicon:  
$$N \rightarrow \textit{fish}$$
$$(N \text{ Agr} \{3s, 3p\})$$
- Allows for indirect specification and transfer of information, *e.g.*, head features.

# Features and the lexicon

- Lexicon may contain each inflected form.
  - Feature values and base form listed.
- Lexicon may contain only base forms.
  - Process of *morphological analysis* maps inflected form to base form plus feature values.
  - Time–space trade-off, varies by language.
- Lexicon may contain *semantics* for each form.

# Morphological analysis

- Morphological analysis is simple in English.
- Reverse the rules for inflections, including spelling changes.

*dogs* → *dog* [**Agr** 3p]

*eats* → *eat* [**Agr** 3s, **Tns** pres]

*dog* → *dog* [**Agr** 3s]

*ripped* → *rip* [**Tns** past]

*berries* → *berry* [**Agr** 3p]

*tarried* → *tarry* [**Tns** past]

*buses* → *bus* [**Agr** 3p]

*running* → *run* [**Tns** pp]

- Irregular forms will always have to be explicitly listed in lexicon.

*children* → *child* [**Agr** 3p]

*sang* → *sing* [**Tns** past]

# Morphology in other languages

- Rules may be more complex in other (even European) languages.
- Languages with compounding (e.g., German) or agglutination (e.g., Finnish) require more-sophisticated methods.
  - E.g., *Verdauungsspaziergang*, a stroll that one takes after a meal to assist in digestion.



# Semantics as a lexical feature

- Add a **Sem** feature:

<b>Cat</b>	N
<b>Num</b>	s
<b>Pers</b>	3
<b>Lex</b>	<i>dog</i>
<b>Sem</b>	dog

Typewriter font  
for semantic objects

- The meaning of *dog* is dog.  
The meaning of *chien* and *Hund* are both dog.  
The meaning of *dog* is G52790.  
Not uncommon, but pretty unimpressive...

# Verb subcategorization 1

- **Problem:** Constraints on verbs and their complements.

*Nadia told / instructed / \*said / \*informed Ross to sit down.*

*Nadia \*told / \*instructed / said / \*informed to sit down.*

*Nadia told / \*instructed / \*said / informed Ross of the requirement to sit down.*

*Nadia gave / donated her painting to the museum.*

*Nadia gave / \*donated the museum her painting.*

*Nadia put / ate the cake in the kitchen.*

*Nadia \*put / ate the cake.*

# Verb subcategorization 2

- VPs are much more complex than just V with optional NP and/or PP.
  - Can include more than one NP.
  - Can include clauses of various types:  
*that Ross fed the marmoset*  
*to pay him the money*
- **Subcat**: A feature on a verb indicating the kinds of verb phrase it allows:  
\_np, \_np\_np, \_inf, \_np\_inf, ...

Write this way to distinguish from constituents.

# Verb tense and aspect 1

- **Tense and aspect** markings on verb:
  - Locate the event in time (relative to another time).
  - Mark the event as complete/finished or in progress.

*Nadia rides the horse.* — In progress now.

*Nadia rode the horse.* — Completed before now.

*Nadia had ridden the horse.* — Completed before before now.

*Nadia was riding the horse.* — In progress before now.

⋮

# Verb tense and aspect 2

- Tense: past or present
- Aspect: simple, progressive, or perfect

*Nadia ...*

	Auxiliary verb	In progress	Complete
	Simple	Progressive	Perfect
Present	<i>rides</i>	<i>is riding</i>	<i>has ridden</i>
Past	<i>rode</i>	<i>was riding</i>	<i>had ridden</i>

*... the horse*

Present participle

Past participle

# Verb tense and aspect 3

- Tense: past or present
- Aspect: simple, progressive, or perfect

<i>Nadia ...</i>	Auxiliary verbs	
	Simple	Perfect progressive (continuous)
Present	<i>rides</i>	<i>has been riding</i>
Past	<i>rode</i>	<i>had been riding</i>

*... the horse*

# Modal verbs

- **Modal verbs:** Auxiliary verbs that express degrees of certainty, obligation, possibility, prediction, etc.

*Nadia*

*{could, should, must, ought to, might, will, ...}*

*{ride, be riding, have ridden, have been riding}*

*the horse.*

# English auxiliary system

- Structure (so far):  
[MODAL] [HAVE] [BE] MAIN-VERB
- General pattern:  
VP → AUX VP  
AUX → MODAL | HAVE | BE
  - Use features to capture necessary agreements.



# Voice 1

*The goalie kicked the ball.*

**ACTIVE**

Event: *kicked*

Role: Agent  
(doer)

Role: Theme  
(thing affected)

Thing: *the goalie*

Thing: *the ball*

kick (agent=goalie, theme=ball)

# Voice 2

*The ball was kicked.*

**PASSIVE**

Event: *kicked*

Role: Theme  
(thing affected)

Thing: *the ball*

kick (agent=?, theme=ball)

# Voice 3

*The ball was kicked by the goalie.* **PASSIVE**

Event: *kicked*

Role: Theme  
(thing affected)

Role: Agent  
(doer)

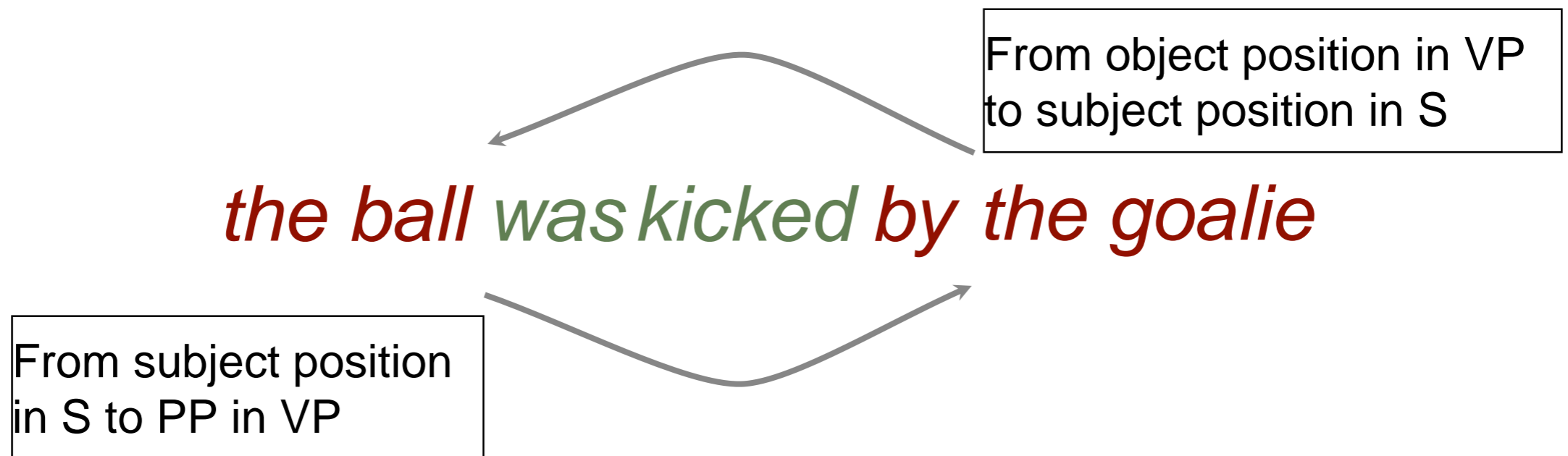
Thing: *the ball*    Thing: *the goalie*

kick (agent=goalie, theme=ball)

# Passive as *Diathetic alternation*

*the goalie*      *kicked*      *the ball*

# Passive as *Diathetic alternation*



But the semantic representation doesn't change

# Voice 4

- **Voice:** System of assigning thematic roles to syntactic positions.
  - English has **active** and **passive** voices.
- Passive expressed with *be*+past participle.  
Other auxiliaries may also apply, including progressive *be*.
- *Nadia was kissed.*                      *Nadia was being kissed.*  
*Nadia had been kissed.*                *Nadia had been being kissed.*  
*Nadia could be kissed.*                 *Nadia could have been being*  
*kissed.*
- Structure:  
[MODAL] [HAVE] [BE<sub>1</sub>] [BE<sub>2</sub>] MAIN-VERB

# Some useful features

- **VForm**: The tense/aspect form of a verb: passive, pastprt, ...
- **CompForm**: The tense/aspect form of the complement of an auxiliary.

# Augmenting rules for passive voice

- For all rules of the form:

$$\begin{array}{ccc} VP \rightarrow V \ NP \ X & \xrightarrow{\text{ADD}} & VP \rightarrow V \ X \\ (V \ \mathbf{Subcat}) = \_y & \rightarrow & (V \ \mathbf{Subcat}) = \_y \\ & & (V \ \mathbf{VForm}) = \text{passive} \\ & & (VP \ \mathbf{VForm}) = \text{passive} \end{array}$$

Metarule to ease grammar coding

- Augment Aux+VP rules:

$$\begin{array}{l} VP \rightarrow AUX \ VP \\ (AUX \ \mathbf{Root}) = \text{Be2} \\ (AUX \ \mathbf{CompForm}) = (VP_2 \ \mathbf{VForm}) \\ (VP_2 \ \mathbf{VForm}) = \text{passive} \end{array}$$



# The GAP feature for passive voice

S → NP VP

- <sup>1</sup> (NP **Agr**) = (VP **Agr**)
- <sup>2</sup> (VP **VForm**) = passive
- <sup>3</sup> (VP **Gap Cat**) = NP
- <sup>4</sup> (VP **Gap Agr**) = (NP **Agr**)
- <sup>5</sup> (VP **Gap Sem**) = (NP **Sem**)

VP → AUX VP

- <sup>1</sup> (VP<sub>1</sub> **Agr**) = (AUX **Agr**)
- <sup>2</sup> (VP<sub>1</sub> **VForm**) = (VP<sub>2</sub> **VForm**)
- <sup>3</sup> (VP<sub>1</sub> **Gap**) = (VP<sub>2</sub> **Gap**)
- <sup>4</sup> (AUX **Lex**) = be2
- <sup>5</sup> (VP<sub>2</sub> **VForm**) = passive

V → *kicked*

- <sup>1</sup> (V **VForm**) = {pastprt, passive}
- <sup>2</sup> (V **Subcat**) = \_np
- <sup>3</sup> (V **Lex**) = *kick*
- <sup>4</sup> (V **Sem**) = kick

VP → V NP

- <sup>1</sup> (VP **VForm**) = (V **VForm**)
- <sup>2</sup> (VP **Gap**) = (NP **Gap**)
- <sup>3</sup> (V **Subcat**) = \_np

Empty string

NP → ε

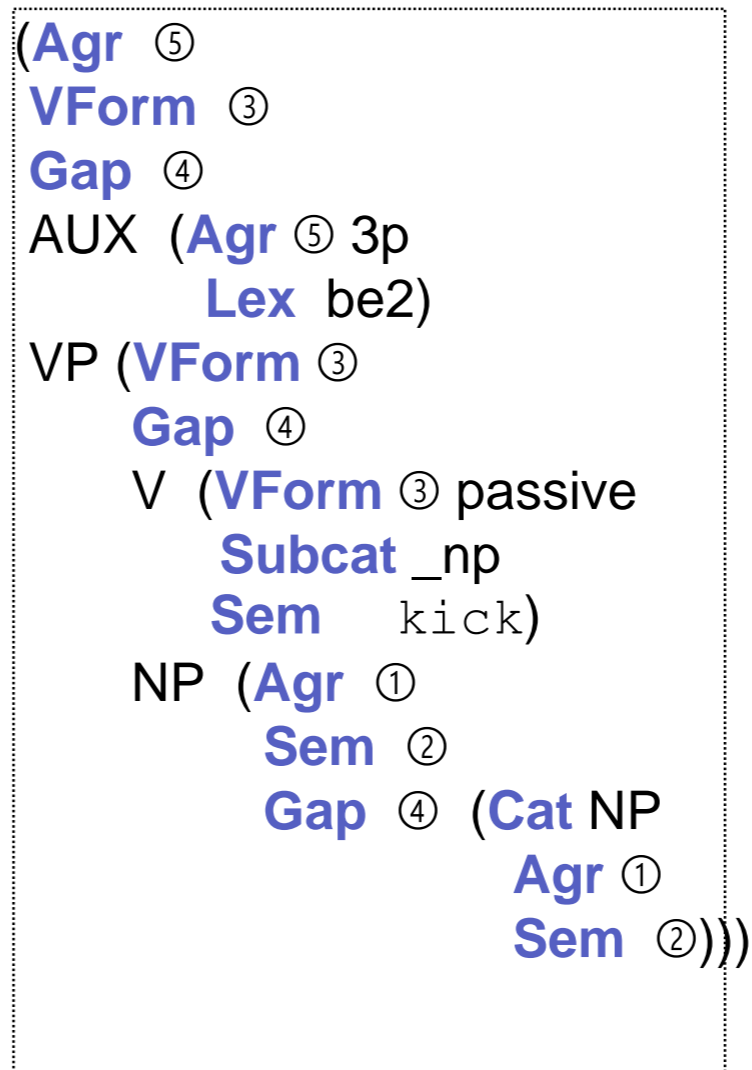
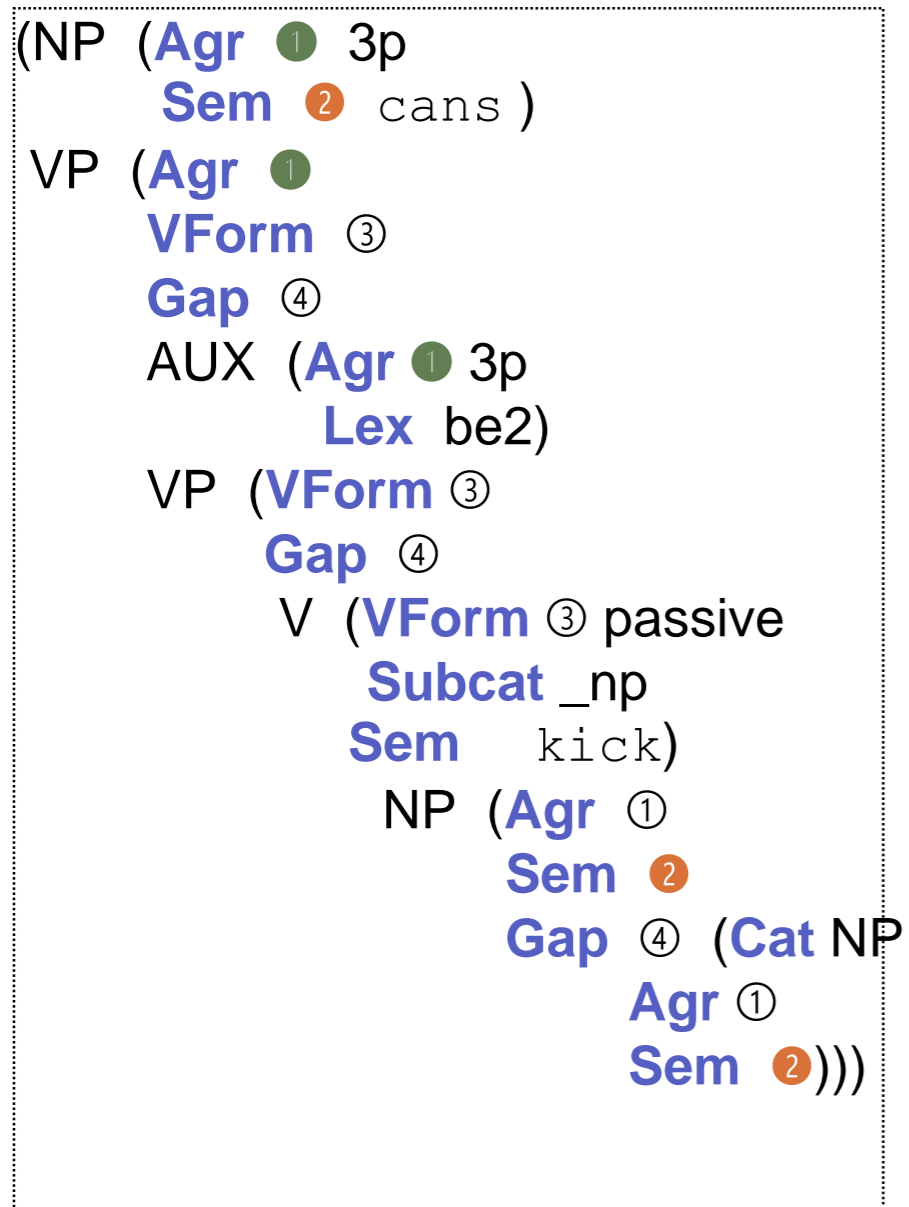
- <sup>1</sup> (NP **Gap Cat**) = NP
- <sup>2</sup> (NP **Gap Agr**) = (NP **Agr**)
- <sup>3</sup> (NP **Gap Sem**) = (NP **Sem**)

NP → *cans*

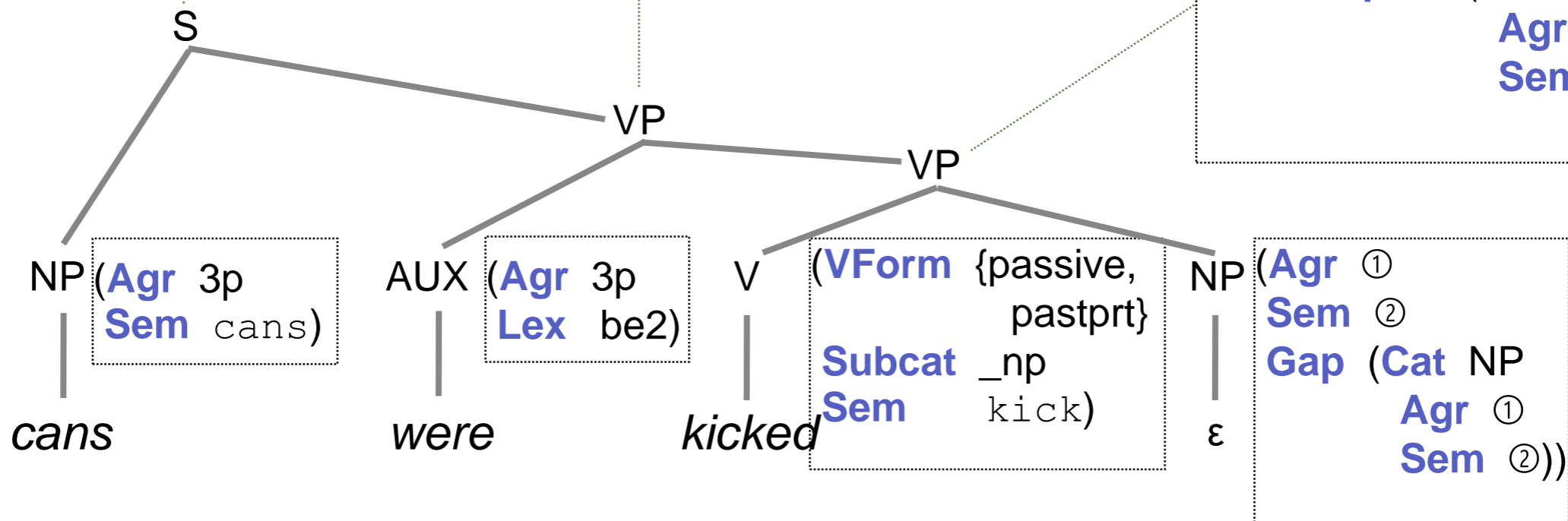
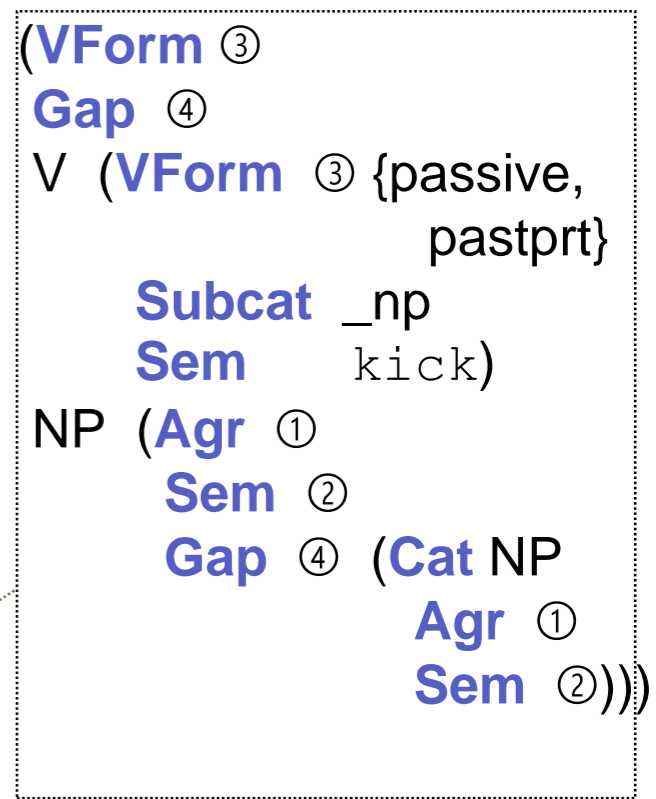
- <sup>1</sup> (NP **Agr**) = 3p
- <sup>2</sup> (NP **Lex**) = *can*
- <sup>3</sup> (NP **Sem**) = cans

AUX → *were*

- <sup>1</sup> (AUX **Agr**) = 3p
- <sup>2</sup> (AUX **Lex**) = be2



Note: The green ①'s of the S were ⑤'s until the 4th constraint of the rule  $S \rightarrow NP VP$ . The 5th constraint fills in the **Sem** of the **Gap** ②.



# Other cases of *gap percolation*

- Other constructions involve NPs in syntactic configurations where they would not get the right thematic roles using linear order alone.

*Nadia seems to like Ross.*

*Nadia seems to be liked.*

*Nadia is easy to like.*

*Who did Nadia like?*

*I fed the dog that Nadia likes to walk.*

- Can use grammar rules with gap features to ensure correct structure/interpretation of these as well.

# Summary

- Features help capture syntactic constructions in a general and elegant grammar.
- Features can encode the compositional semantics of a sentence as you parse it.
- Features can accomplish mapping functions between syntax and semantics that simplify the interpretation process.