Computational Lexical Semantics

Lecture 7: Semantic Role Labeling

Linguistic Institute 2005
University of Chicago

Semantic Role Labeling

Introduction

Semantic Role Labeling

Who did *what* to *whom* at *where?*

The police officer detained the suspect at the scene of the crime

Agent Predicate Theme Location

Can we figure out that these have the same meaning?

XYZ corporation **bought** the stock.

They **sold** the stock to XYZ corporation.

The stock was **bought** by XYZ corporation.

The **purchase** of the stock by XYZ corporation...

The stock **purchase** by XYZ corporation...

A Shallow Semantic Representation: Semantic Roles

Predicates (bought, sold, purchase) represent an **event**

semantic roles express the abstract role that arguments of a predicate can take in the event

More specific ⟷ More general

buyer **agent** **proto-agent**

Semantic Role Labeling

Semantic Roles

Getting to semantic roles

Neo-Davidsonian event representation:

Sasha broke the window

Pat opened the door

$$\exists e,x,y\ Breaking(e) \wedge Breaker(e,Sasha) \wedge BrokenThing(e,y) \wedge Window(y)$$

$$\exists e,x,y\ Opening(e) \wedge Opener(e,Pat) \wedge OpenedThing(e,y) \wedge Door(y)$$

Subjects of break and open: **Breaker** and **Opener**

Deep roles specific to each event (breaking, opening)

Hard to reason about them for NLU applications like QA

Thematic roles

- **Breaker** and **Opener** have something in common!
  - Volitional actors
  - Often animate
  - Direct causal responsibility for their events
- Thematic roles are a way to capture this semantic commonality between *Breakers* and *Eaters*.
- They are both AGENTS.
- The *BrokenThing* and *OpenedThing*, are THEMES.
  - prototypically inanimate objects affected in some way by the action

Thematic roles

- One of the oldest linguistic models
  - Indian grammarian Panini between the 7th and 4th centuries BCE
- Modern formulation from Fillmore (1966,1968), Gruber (1965)
  - Fillmore influenced by Lucien Tesnière's (1959) *Éléments de Syntaxe Structurale,* the book that introduced dependency grammar
  - Fillmore first referred to roles as *actants* (Fillmore, 1966) but switched to the term *case*

Thematic roles

- A typical set:

Thematic Role	**Definition**	**Example**
AGENT	The volitional causer of an event	*The waiter* spilled the soup.
EXPERIENCER	The experiencer of an event	*John* has a headache.
FORCE	The non-volitional causer of the event	*The wind* blows debris from the mall into our yards.
THEME	The participant most directly affected by an event	Only after Benjamin Franklin broke *the ice*...
RESULT	The end product of an event	The city built a *regulation-size baseball diamond*...
CONTENT	The proposition or content of a propositional event	Mona asked *"You met Mary Ann at a supermarket?"*
INSTRUMENT	An instrument used in an event	He poached catfish, stunning them *with a shocking device*...
BENEFICIARY	The beneficiary of an event	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	The origin of the object of a transfer event	I flew in *from Boston*.
GOAL	The destination of an object of a transfer event	I drove *to Portland*.

Thematic grid, case frame, θ-grid

Example usages of "break"

John	*broke the window.*	
AGENT	THEME	
John	*broke the window with a rock.*	
AGENT	THEME	INSTRUMENT
The rock	*broke the window.*	
INSTRUMENT	THEME	
The window broke.		
THEME		
The window was broken by John.		
THEME		AGENT

thematic grid, case frame, θ-grid

Break:

AGENT, THEME, INSTRUMENT.

Some realizations:

AGENT/Subject, THEME/Object
AGENT/Subject, THEME/Object, INSTRUMENT/PP_{with}
INSTRUMENT/Subject, THEME/Object
THEME/Subject

Diathesis alternations (or verb alternation)

Doris	*gave*	*the book*	*to Cary.*
AGENT		THEME	GOAL
Doris	*gave*	*Cary*	*the book.*
AGENT		GOAL	THEME

Break: AGENT, INSTRUMENT, or THEME as subject

Give: THEME and GOAL in either order

Dative alternation: particular semantic classes of verbs, “verbs of future having” (*advance, allocate, offer, owe*), “send verbs” (*forward, hand, mail*), “verbs of throwing” (*kick, pass, throw*), etc.

Levin (1993): 47 semantic classes (“**Levin classes**”) for 3100 English verbs and alternations. In online resource VerbNet.

Problems with Thematic Roles

Hard to create standard set of roles or formally define them

Often roles need to be fragmented to be defined.

Levin and Rappaport Hovav (2015): two kinds of INSTRUMENTS

intermediary instruments that can appear as subjects

The cook opened the jar with the new gadget.

The new gadget opened the jar.

enabling instruments that cannot

Shelly ate the sliced banana with a fork.

*The fork ate the sliced banana.

Alternatives to thematic roles

1. **Fewer roles**: generalized semantic roles, defined as prototypes (Dowty 1991)

   PROTO-AGENT

   PROTO-PATIENT

   PropBank

2. **More roles**: Define roles specific to a group of predicates

   FrameNet

Semantic Role Labeling

The Proposition Bank (PropBank)

PropBank

- Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106

PropBank Roles

Following Dowty 1991

Proto-Agent

- Volitional involvement in event or state
- Sentience (and/or perception)
- Causes an event or change of state in another participant
- Movement (relative to position of another participant)

Proto-Patient

- Undergoes change of state
- Causally affected by another participant
- Stationary relative to movement of another participant

PropBank Roles

- Following Dowty 1991
  - Role definitions determined verb by verb, with respect to the other roles
  - Semantic roles in PropBank are thus verb-sense specific.
- Each verb sense has numbered argument: Arg0, Arg1, Arg2,…

Arg0: PROTO-AGENT

Arg1: PROTO-PATIENT

Arg2: usually: benefactive, instrument, attribute, or end state

Arg3: usually: start point, benefactive, instrument, or attribute

Arg4 the end point

(Arg2-Arg5 are not really that consistent, causes a problem for labeling)

PropBank Frame Files

agree.01

Arg0: Agreer
Arg1: Proposition
Arg2: Other entity agreeing

Ex1: [$_{\text{Arg0}}$ The group] *agreed* [$_{\text{Arg1}}$ it wouldn't make an offer].
Ex2: [$_{\text{ArgM-TMP}}$ Usually] [$_{\text{Arg0}}$ John] *agrees* [$_{\text{Arg2}}$ with Mary] [$_{\text{Arg1}}$ on everything].

fall.01

Arg1: Logical subject, patient, thing falling
Arg2: Extent, amount fallen
Arg3: start point
Arg4: end point, end state of arg1
Ex1: [$_{\text{Arg1}}$ Sales] *fell* [$_{\text{Arg4}}$ to $25 million] [$_{\text{Arg3}}$ from $27 million].
Ex2: [$_{\text{Arg1}}$ The average junk bond] *fell* [$_{\text{Arg2}}$ by 4.2%].

Advantage of a ProbBank Labeling

increase.01 "go up incrementally"

Arg0: causer of increase

Arg1: thing increasing

Arg2: amount increased by, EXT, or MNR

Arg3: start point

Arg4: end point

This would allow us to see the commonalities in these 3 sentences:

[$_{\text{Arg0}}$ Big Fruit Co.] increased [$_{\text{Arg1}}$ the price of bananas].

[$_{\text{Arg1}}$ The price of bananas] was increased again [$_{\text{Arg0}}$ by Big Fruit Co.]

[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].

Modifiers or adjuncts of the predicate: Arg-M

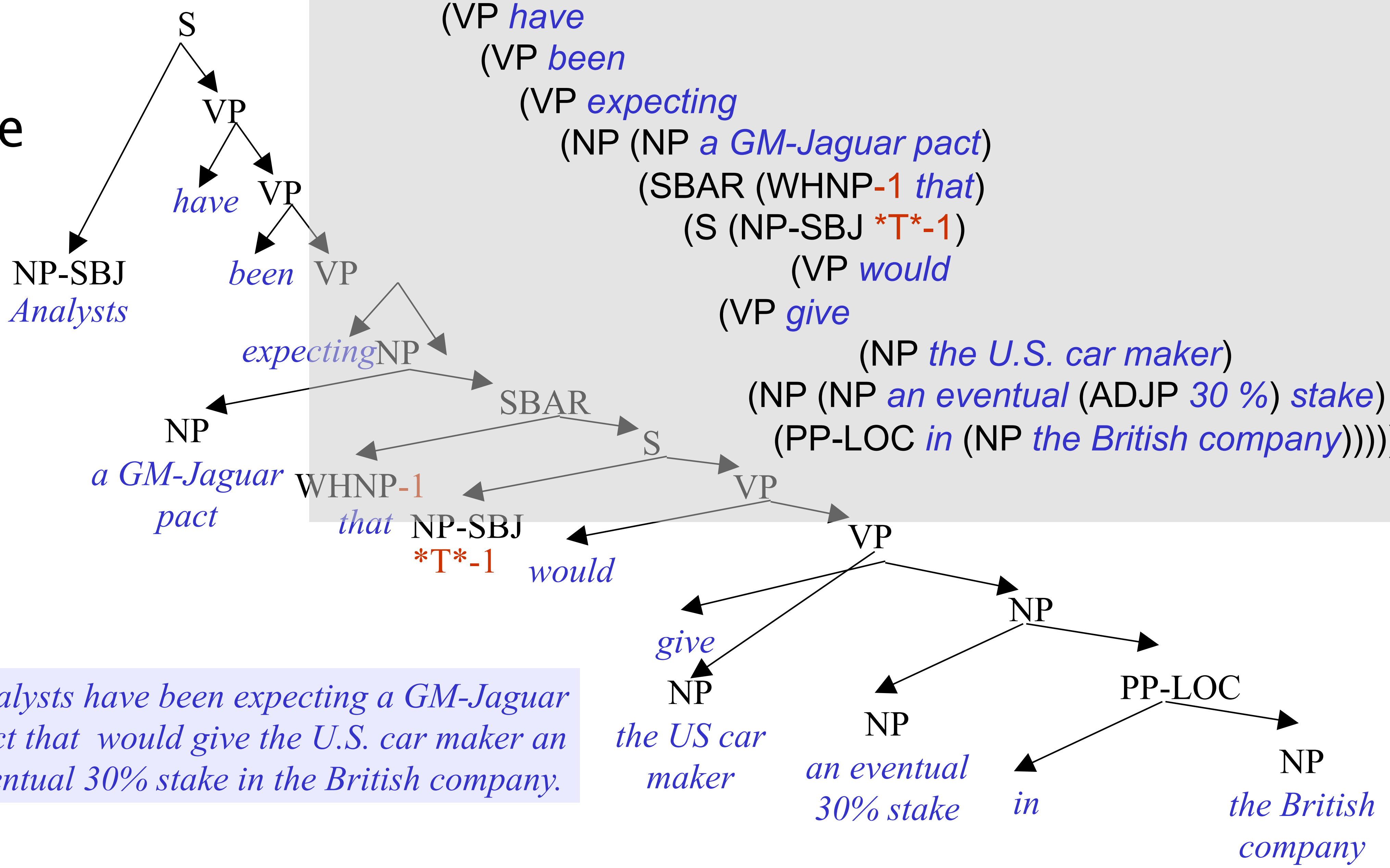
ArgM-TMP	when?	yesterday evening, now
LOC	where?	at the museum, in San Francisco
DIR	where to/from?	down, to Bangkok
MNR	how?	clearly, with much enthusiasm
PRP/CAU	why?	because ... , in response to the ruling
REC		themselves, each other
ADV	miscellaneous	
PRD	secondary predication	...ate the meat raw

PropBanking a Sentence

Martha Palmer 2013

A sample parse tree

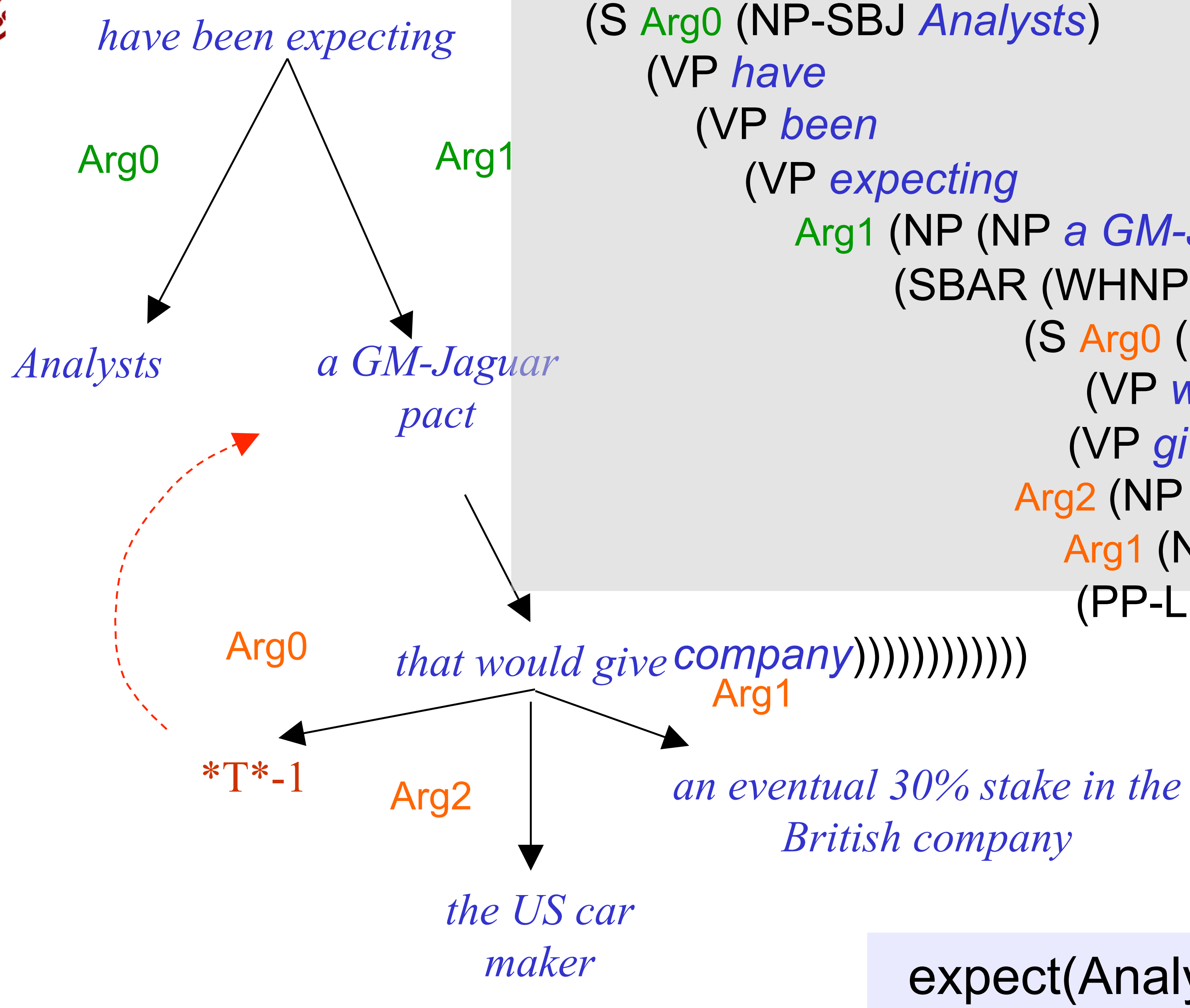
```
(S (NP-SBJ Analysts)
    (VP have
        (VP been
            (VP expecting
                (NP (NP a GM-Jaguar pact)
                    (SBAR (WHNP-1 that)
                        (S (NP-SBJ *T*-1)
                            (VP would
                        (VP give
                            (NP the U.S. car maker)
                        (NP (NP an eventual (ADJP 30 %) stake)
                            (PP-LOC in (NP the British company))))))))))))
```



Analysts have been expecting a GM-Jaguar pact that would give the U.S. car maker an eventual 30% stake in the British company.

The same parse tree PropBanked

Martha Palmer 2013



(S Arg0 (NP-SBJ *Analysts*)
(VP *have*
(VP *been*
(VP *expecting*
Arg1 (NP (NP *a GM-Jaguar pact*)
(SBAR (WHNP-1 *that*)
(S Arg0 (NP-SBJ *T*-1)
(VP *would*
(VP *give*
Arg2 (NP *the U.S. car maker*)
Arg1 (NP (NP *an eventual* (ADJP *30 %*) *stake*)
(PP-LOC *in* (NP *the British*
company))))))))))))

expect(Analysts, GM-J pact)
give(GM-J pact, US car maker, 30% stake)

Annotated PropBank Data

- Penn English TreeBank, OntoNotes 5.0.
  - Total ~2 million words
- Penn Chinese TreeBank
- Hindi/Urdu PropBank
- Arabic PropBank

2013 Verb Frames Coverage
Count of word sense (lexical units)

Language	*Final Count*
English	**10,615***
Chinese	**24, 642**
Arabic	**7,015**

From Martha Palmer 2013 Tutorial

Plus nouns and light verbs

Example Noun: *Decision*

- Roleset: Arg0: decider, Arg1: decision…

- "…[your$_{ARG0}$] [decision$_{REL}$]
  [to say look I don't want to go through this anymore$_{ARG1}$]"

Example within an LVC: *Make a decision*

- "…[the President$_{ARG0}$] [made$_{REL-LVB}$]
  the [fundamentally correct$_{ARGM-ADJ}$]
  [decision$_{REL}$] [to get on offense$_{ARG1}$]"

Slight from Palmer 2013

Semantic Role Labeling

FrameNet

Capturing descriptions of the same event by different nouns/verbs

[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].

[$_{\text{Arg1}}$ The price of bananas] rose [$_{\text{Arg2}}$ 5%].

There has been a [$_{\text{Arg2}}$ 5%] rise [$_{\text{Arg1}}$ in the price of bananas].

FrameNet

- Baker et al. 1998, Fillmore et al. 2003, Fillmore and Baker 2009, Ruppenhofer et al. 2006
- Roles in PropBank are specific to a verb
- Role in FrameNet are specific to a **frame: a** background knowledge structure that defines a set of frame-specific semantic roles, called **frame elements**,
  - includes a set of pred cates that use these roles
  - each word evokes a frame and profiles some aspect of the frame

The "Change position on a scale" Frame

This frame consists of words that indicate the change of an ITEM's position on a scale (the ATTRIBUTE) from a starting point (INITIAL VALUE) to an end point (FINAL VALUE)

[$_{\text{ITEM}}$ Oil] *rose* [$_{\text{ATTRIBUTE}}$ in price] [$_{\text{DIFFERENCE}}$ by 2%].

[$_{\text{ITEM}}$ It] has *increased* [$_{\text{FINAL\_STATE}}$ to having them 1 day a month].

[$_{\text{ITEM}}$ Microsoft shares] *fell* [$_{\text{FINAL\_VALUE}}$ to 7 5/8].

[$_{\text{ITEM}}$ Colon cancer incidence] *fell* [$_{\text{DIFFERENCE}}$ by 50%] [$_{\text{GROUP}}$ among men].

a steady *increase* [$_{\text{INITIAL\_VALUE}}$ from 9.5] [$_{\text{FINAL\_VALUE}}$ to 14.3] [$_{\text{ITEM}}$ in dividends]

a [$_{\text{DIFFERENCE}}$ 5%] [$_{\text{ITEM}}$ dividend] *increase*...

The "Change position on a scale" Frame

VERBS:
advance
climb
decline
decrease
diminish
dip
double
drop
dwindle
edge
explode
fall
fluctuate
gain
grow
increase
jump
move
mushroom
plummet
reach
rise
rocket
shift
skyrocket
slide
soar
swell
swing
triple
tumble

NOUNS:
decline
decrease
escalation
explosion
fall
fluctuation
gain
growth
hike
increase
rise
shift
tumble

ADVERBS:
increasingly

The “Change position on a scale” Frame

Core Roles	
ATTRIBUTE	The ATTRIBUTE is a scalar property that the ITEM possesses.
DIFFERENCE	The distance by which an ITEM changes its position on the scale.
FINAL_STATE	A description that presents the ITEM’s state after the change in the ATTRIBUTE value as an independent predication.
FINAL_VALUE	The position on the scale where the ITEM ends up.
INITIAL_STATE	A description that presents the ITEM’s state before the change in the A TRIBUTE’s value as an independent predication.
INITIAL_VALUE	The initial position on the scale from which the ITEM moves away.
ITEM	The entity that has a position on the scale.
VALUE_RANGE	A portion of the scale, typically identified by its end points, along which t values of the ATTRIBUTE fluctuate.
Some Non-Core Roles	
DURATION	The length of time over which the change takes place.
SPEED	The rate of change of the VALUE.
GROUP	The GROUP in which an ITEM changes the value of an ATTRIBUTE in a specified way.

Relation between frames

Inherits from:
Is Inherited by:
Perspective on:
Is Perspectivized in:
Uses:
Is Used by:
Subframe of:
Has Subframe(s):
Precedes:
Is Preceded by:
Is Inchoative of:
Is Causative of:

Relation between frames

"cause change position on a scale"

Is Causative of: Change_position_on_a_scale

Adds an agent Role

[AGENT They] *raised* [ITEM the price of their soda] [DIFFERENCE by 2%].

- *add.v, crank.v, curtail.v, cut.n, cut.v, decrease.v, development.n, diminish.v, double.v, drop.v, enhance.v, growth.n, increase.v, knock down.v, lower.v, move.v, promote.v, push.n, push.v, raise.v, reduce.v, reduction.n, slash.v, step up.v, swell.v*

Relations between frames

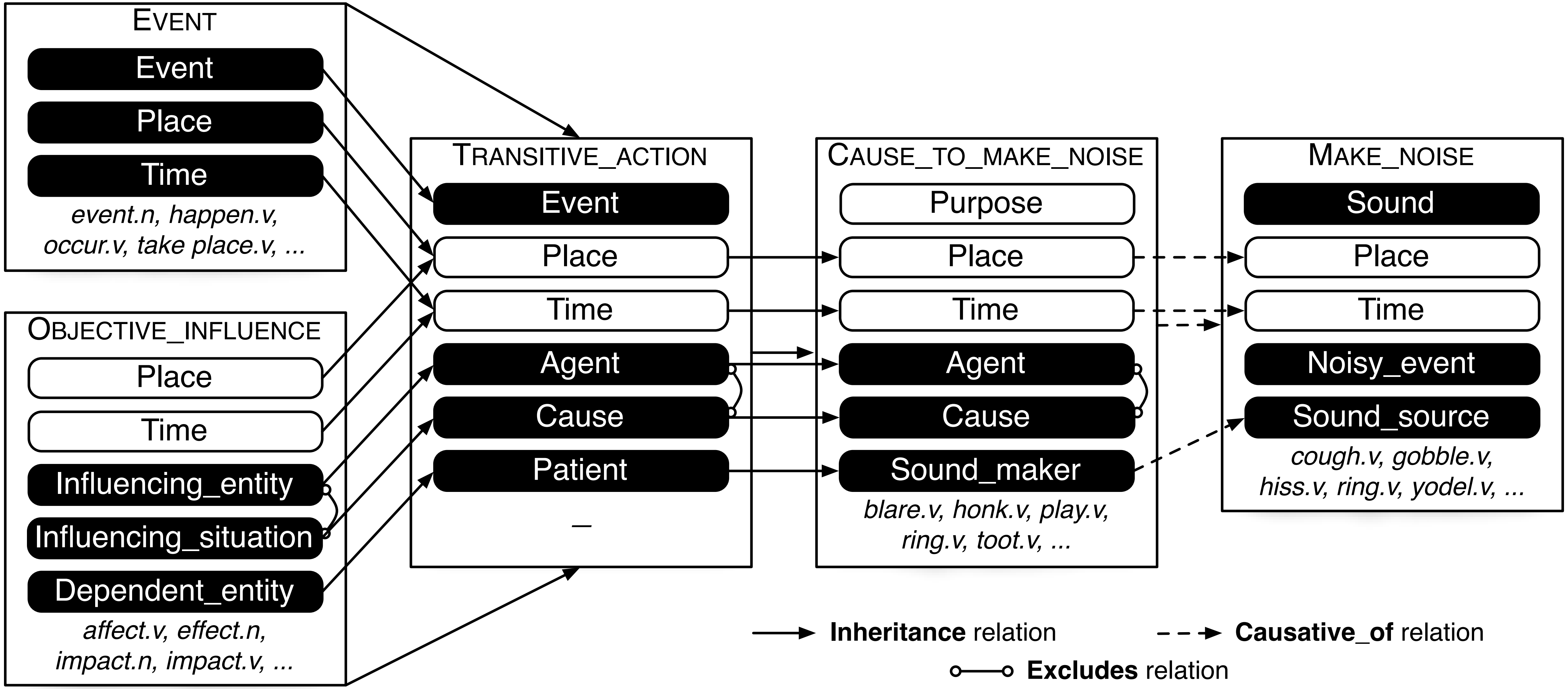


Figure from Das et al 2010

Schematic of Frame Semantics

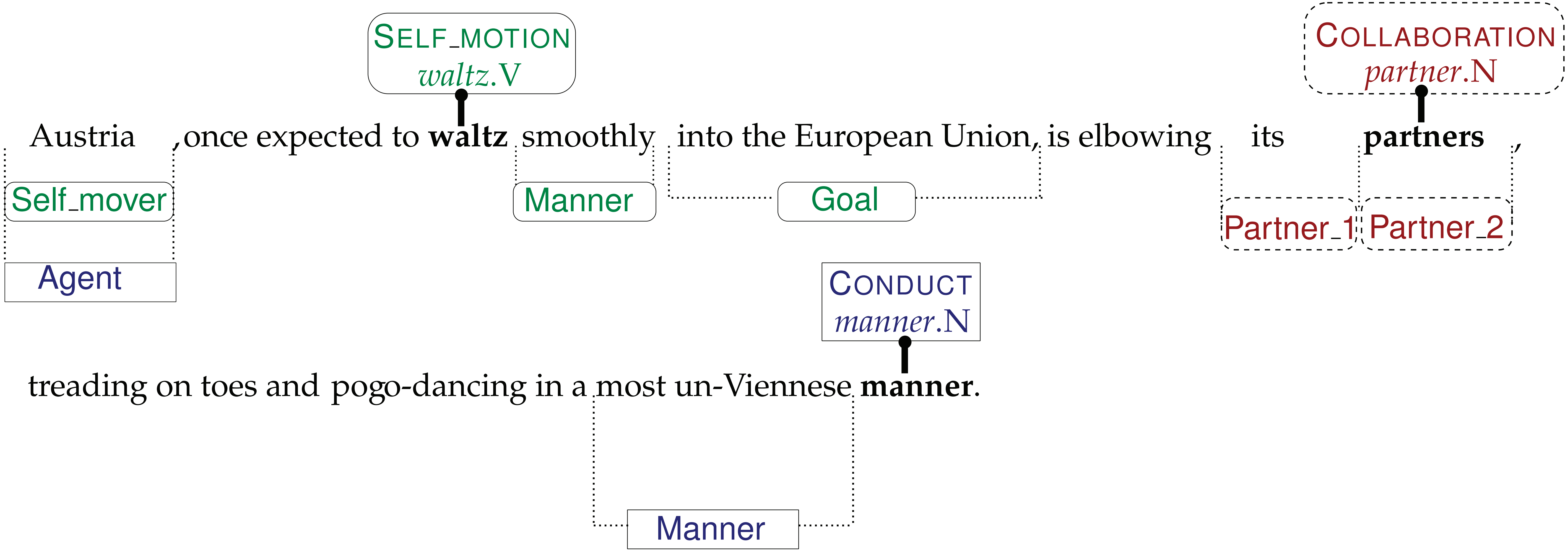
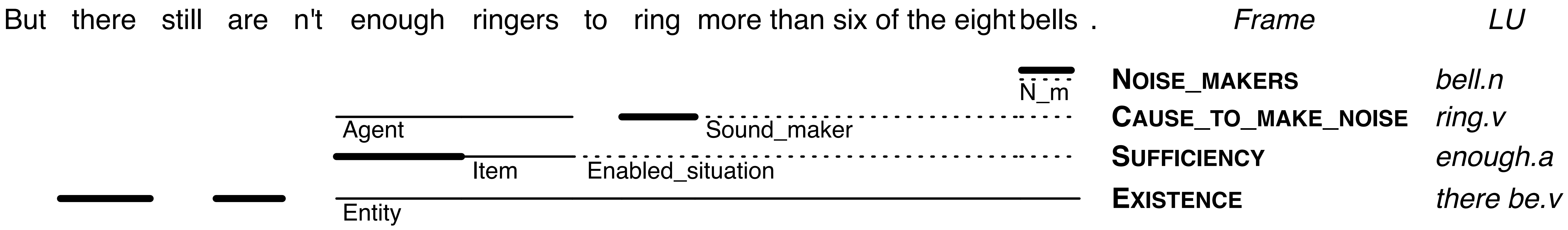


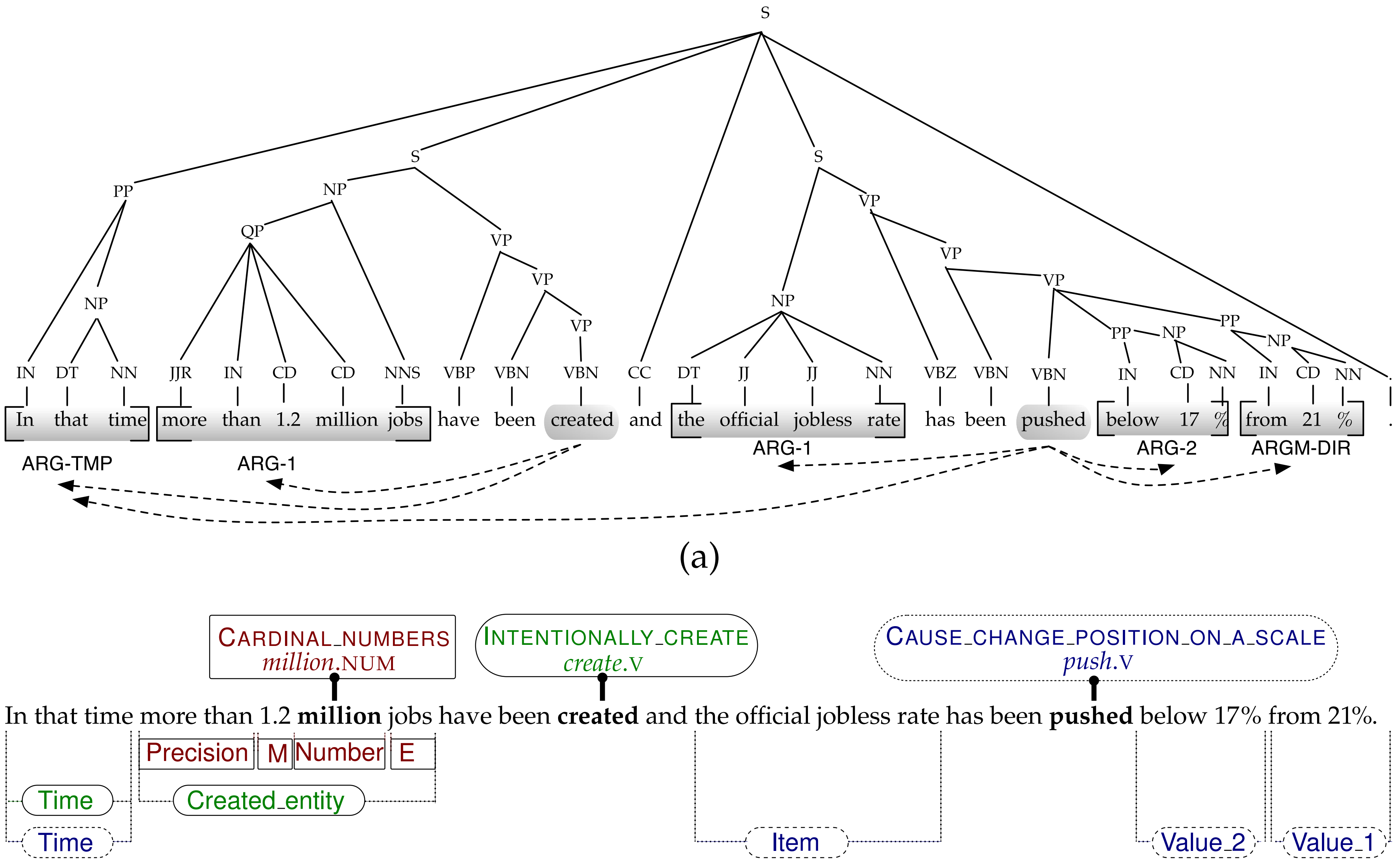
Figure from Das et al (2014)

FrameNet Complexity



From Das et al. 2010

FrameNet and PropBank representations



Semantic Role Labeling

Semantic Role Labeling Algorithm

Semantic role labeling (SRL)

- The task of finding the semantic roles of each argument of each predicate in a sentence.
- FrameNet versus PropBank:

[You]	can't	[blame]	[the program]	[for being unable to identify it]
COGNIZER		TARGET	EVALUEE	REASON

[The San Francisco Examiner]	issued	[a special edition]	[yesterday]
ARG0	TARGET	ARG1	ARGM-TMP

History

- Semantic roles as a intermediate semantics, used early in
  - machine translation (Wilks, 1973)
  - question-answering (Hendrix et al., 1973)
  - spoken-language understanding (Nash-Webber, 1975)
  - dialogue systems (Bobrow et al., 1977)
- Early SRL systems

  Simmons 1973, Marcus 1980:

  - parser followed by hand-written rules for each verb
  - dictionaries with verb-specific case frames (Levin 1977)

Why Semantic Role Labeling

- A useful shallow semantic representation
- Improves NLP tasks like:
  - question answering

    Shen and Lapata 2007, Surdeanu et al. 2011
  - machine translation

    Liu and Gildea 2010, Lo et al. 2013

A simple modern algorithm

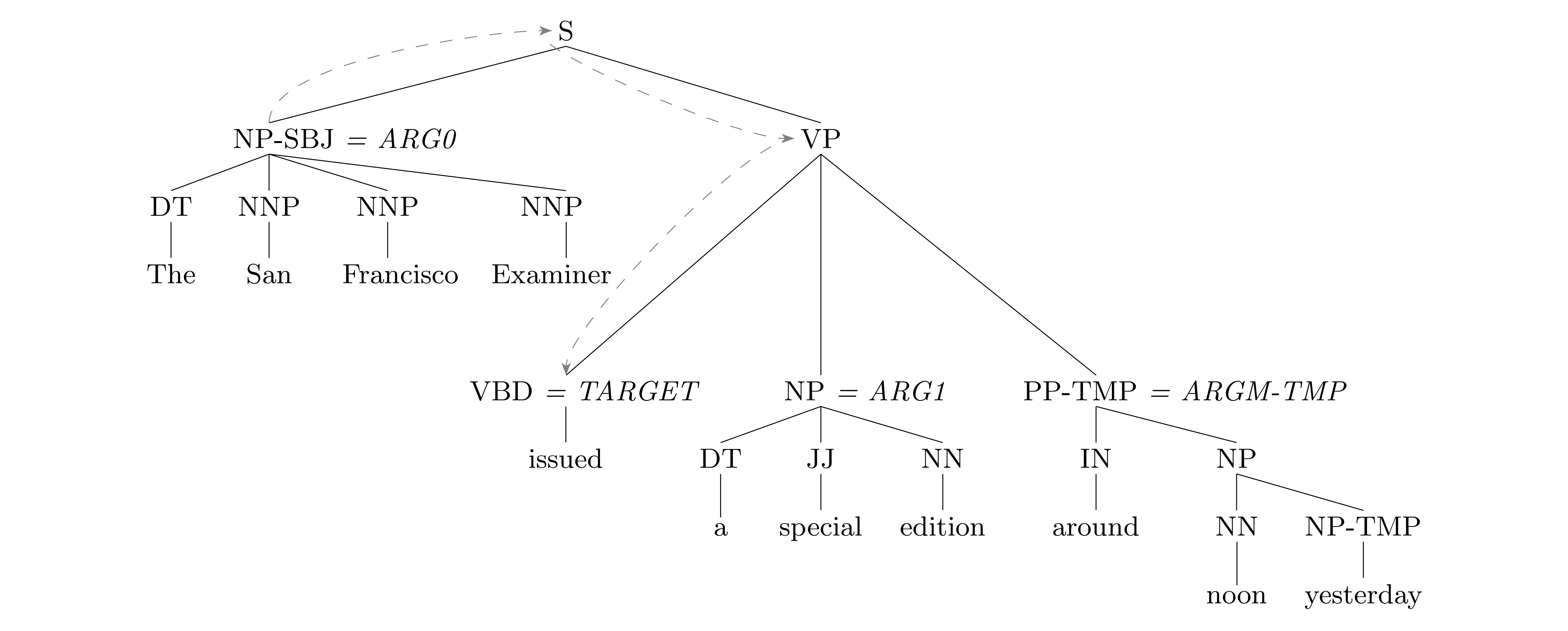
```
function SEMANTICROLELABEL(words) returns labeled tree

    parse ← PARSE(words)
    for each predicate in parse do
        for each node in parse do
            featurevector ← EXTRACTFEATURES(node, predicate, parse)
            CLASSIFYNODE(node, featurevector, parse)
```

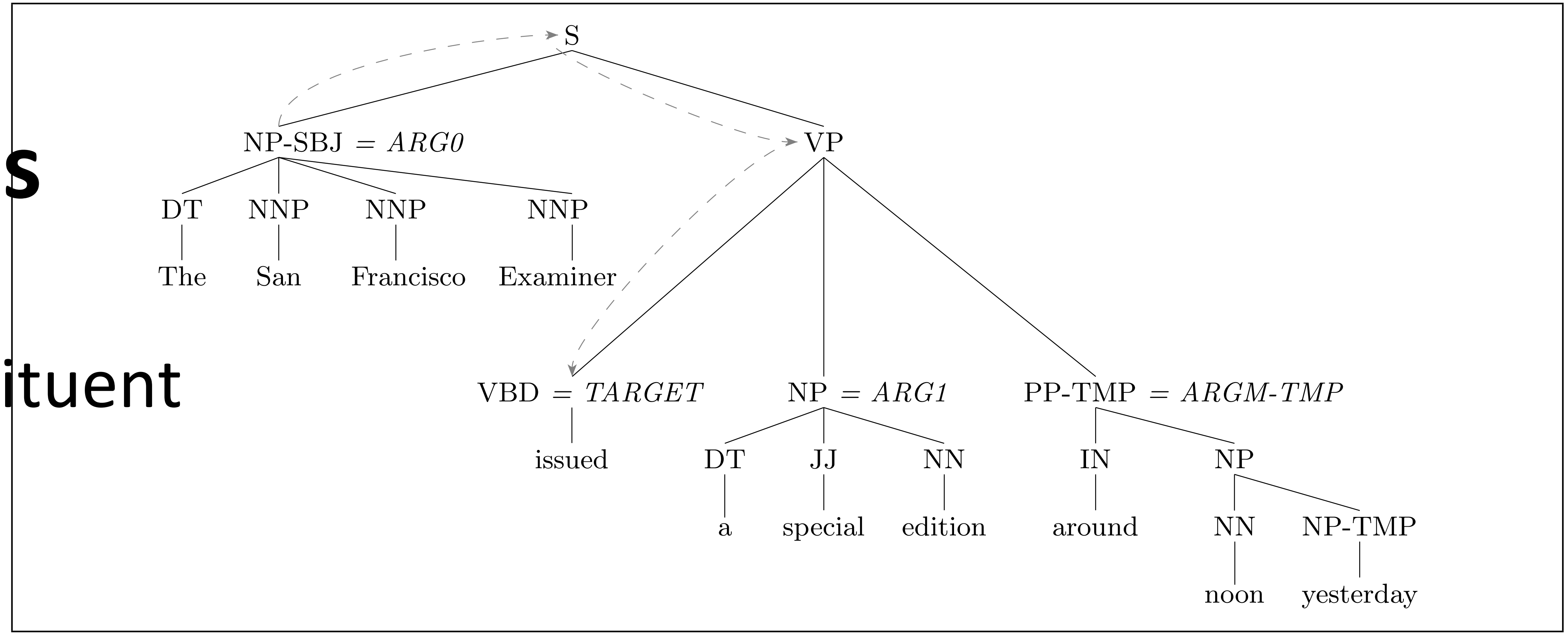
How do we decide what is a predicate

- If we're just doing PropBank verbs
  - Choose all verbs
  - Possibly removing light verbs (from a list)
- If we're doing FrameNet (verbs, nouns, adjectives)
  - Choose every word that was labeled as a target in training data

Semantic Role Labeling



Features



Headword of constituent

Examiner

Headword POS

NNP

Voice of the clause

Active

Subcategorization of pred

VP -> VBD NP PP

Named Entity type of constit

ORGANIZATION

First and last words of constit

The, Examiner

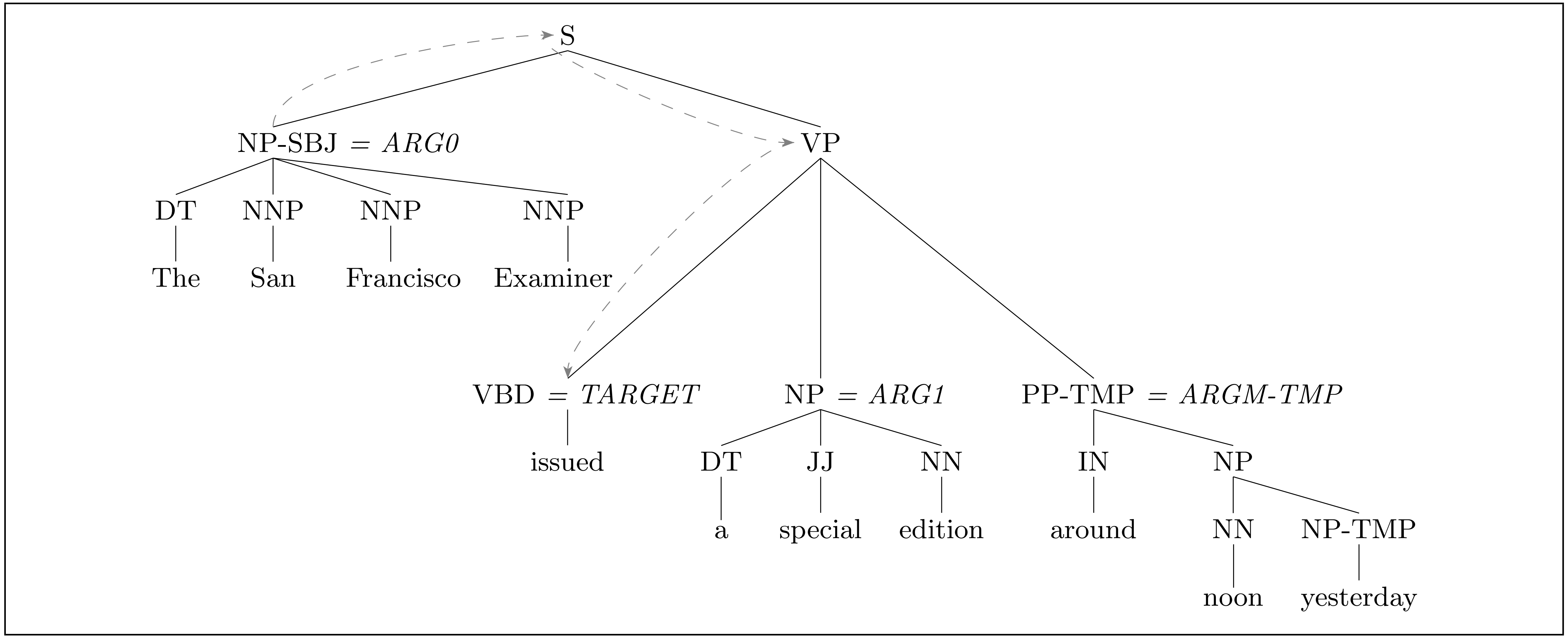
Linear position,clause re: predicate

before

Path Features

Path in the parse tree from the constituent to the predicate

$$NP\uparrow S\downarrow VP\downarrow VBD$$



Frequent path features

Frequency	Path	Description
14.2%	VB↑VP↓PP	PP argument/adjunct
11.8	VB↑VP↑S↓NP	subject
10.1	VB↑VP↓NP	object
7.9	VB↑VP↑VP↑S↓NP	subject (embedded VP)
4.1	VB↑VP↓ADVP	adverbial adjunct
3.0	NN↑NP↑NP↓PP	prepositional complement of noun
1.7	VB↑VP↓PRT	adverbial particle
1.6	VB↑VP↑VP↑VP↑S↓NP	subject (embedded VP)
14.2		no matching parse constituent
31.4	Other	

From Palmer, Gildea, Xue 2010

Final feature vector

- For "The San Francisco Examiner",
- Arg0, [issued, NP, Examiner, NNP, active, before, VP→NP PP, ORG, The, Examiner, NP↑S↓VP↓VBD]

- Other features could be used as well
  - sets of n-grams inside the constituent
  - other path features
    - the upward or downward halves
    - whether particular nodes occur in the path

3-step version of SRL algorithm

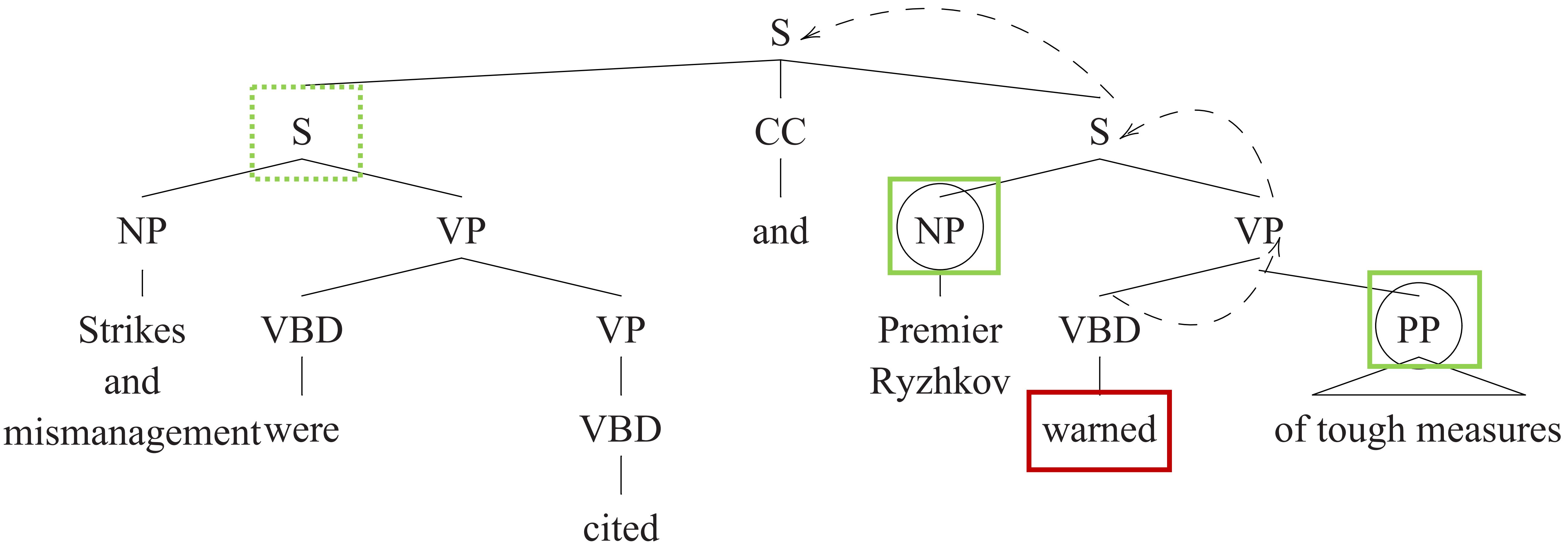
1. **Pruning**: use simple heuristics to prune unlikely constituents.
2. **Identification**: a binary classification of each node as an argument to be labeled or a NONE.
3. **Classification**: a 1-of-*N* classification of all the constituents that were labeled as arguments by the previous stage

Why add Pruning and Identification steps?

- Algorithm is looking at one predicate at a time
- Very few of the nodes in the tree could possible be arguments of that one predicate
- Imbalance between
  - positive samples (constituents that are arguments of predicate)
  - negative samples (constituents that are not arguments of predicate)
- Imbalanced data can be hard for many classifiers
- So we prune the **very** unlikely constituents first, and then use a classifier to get rid of the rest.

Pruning heuristics – Xue and Palmer (2004)

- Add sisters of the predicate, then aunts, then great-aunts, etc
  - But ignoring anything in a coordination structure



A common final stage: joint inference

- The algorithm so far classifies everything **locally** **–** each decision about a constituent is made independently of all others
- But this can’t be right: Lots of **global** or **joint** interactions between arguments
  - Constituents in FrameNet and PropBank must be non-overlapping.
    - A local system may incorrectly label two overlapping constituents as arguments
    - PropBank does not allow multiple identical arguments
      - labeling one constituent ARG0
      - Thus should increase the probability of another being ARG1

How to do joint inference

- Reranking
  - The first stage SRL system produces multiple possible labels for each constituent
  - The second stage classifier the best **global** label for all constituents
  - Often a classifier that takes all the inputs along with other features (sequences of labels)

More complications: FrameNet

We need an extra step to find the frame

function SEMANTICROLELABEL(*words*) **returns** labeled tree
    parse ← PARSE(*words*)
    for each *predicate* **in** *parse* **do**
        for each *node* **in** *parse* **do**
            featurevector ← EXTRACTFEATURES(*node*, *predicate*, *parse*)
            CLASSIFYNODE(*node*, *featurevector*, *parse*)

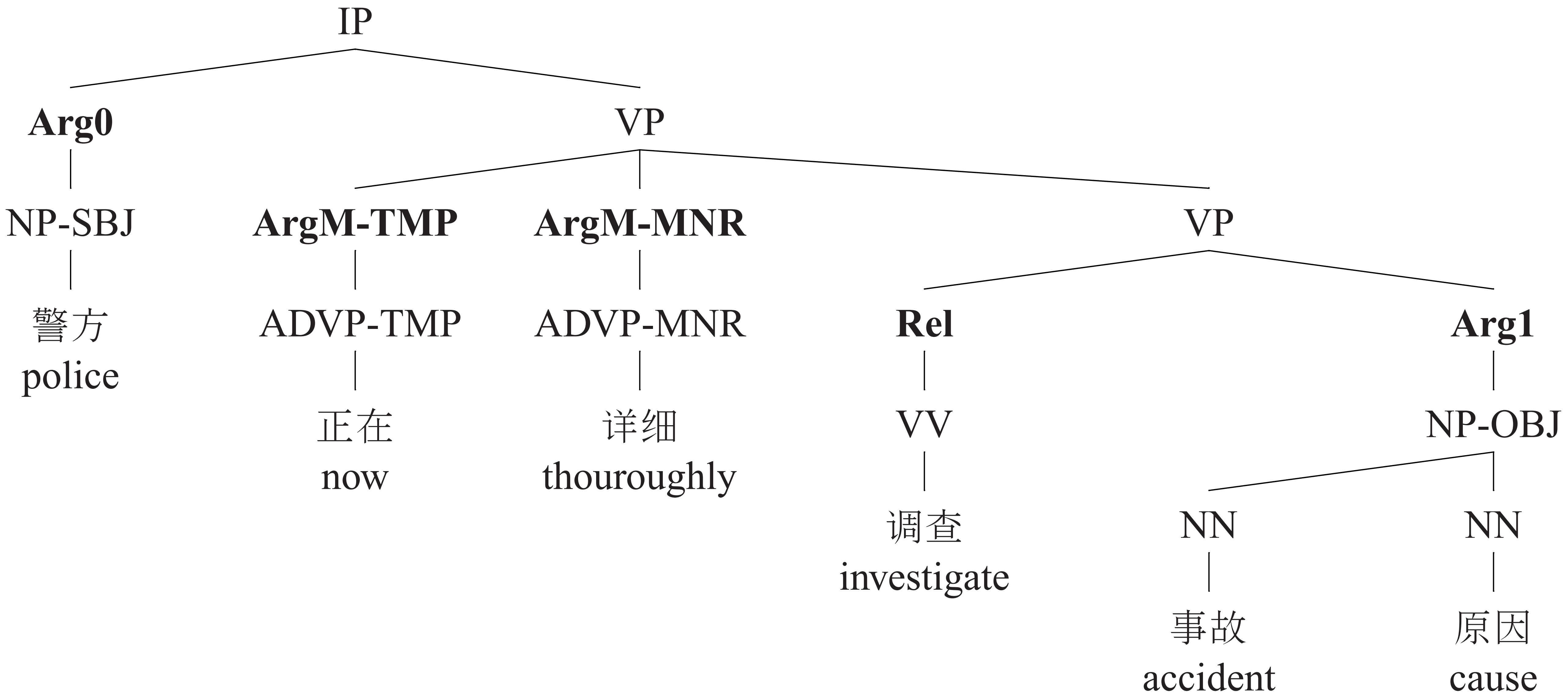
function SEMANTICROLELABEL(*words*) **returns** labeled tree
    parse ← PARSE(*words*)
    for each *predicate* **in** *parse* **do**
        Predicatevector ← ExtractFrameFeatures(predicate,parse)
        Frame ← ClassifyFrame(predicate,predicatevector)
        for each *node* **in** *parse* **do**
            featurevector ← EXTRACTFEATURES(*node*, *predicate*, *parse*)
            CLASSIFYNODE(*node*, *featurevector*, *parse* , Frame)

Features for Frame Identification

Das et al (2014)

the POS of the parent of the head word of t_i
the set of syntactic dependencies of the head word[21] of t_i
if the head word of t_i is a verb, then the set of dependency labels of its children
the dependency label on the edge connecting the head of t_i and its parent
the sequence of words in the prototype, w_ℓ
the lemmatized sequence of words in the prototype
the lemmatized sequence of words in the prototype and their part-of-speech tags π_ℓ
WordNet relation[22] ρ holds between ℓ and t_i
WordNet relation[22] ρ holds between ℓ and t_i, and the prototype is ℓ
WordNet relation[22] ρ holds between ℓ and t_i, the POS tag sequence of ℓ is π_ℓ, and the POS tag sequence of t_i is π_t

Not just English



“The police are thoroughly investigating the cause of the accident.”

Not just verbs: NomBank

Meyers et al. 2004

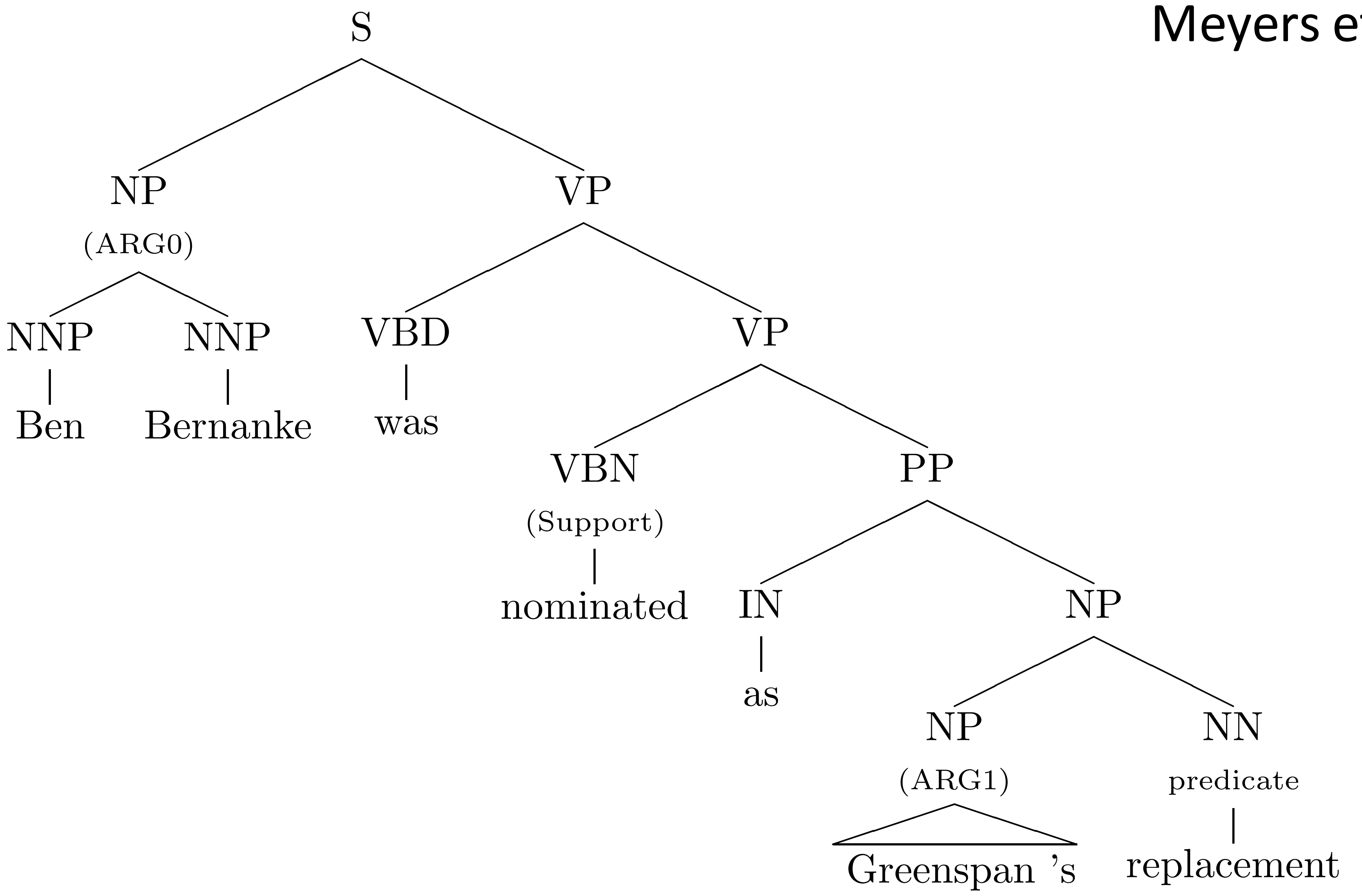


Figure from Jiang and Ng 2006

Additional Issues for nouns

- Features:
  - Nominalization lexicon (employment→ employ)
  - Morphological stem
    - Healthcare, Medicate → care
- Different positions
  - Most arguments of nominal predicates occur inside the NP
  - Others are introduced by support verbs
  - Especially light verbs "X made an argument", "Y took a nap"

Semantic Role Labeling

Conclusion

Semantic Role Labeling

- A level of shallow semantics for representing events and their participants
  - Intermediate between parses and full semantics
- Two common architectures, for various languages
  - FrameNet: frame-specific roles
  - PropBank: Proto-roles
- Current systems extract by
  - parsing sentence
  - Finding predicates in the sentence
    - For each one, classify each parse tree constituent