One Tree is not Enough

Cross-lingual Accumulative Structure Transfer for Semantic Indeterminacy

Patrick Ziering♠ Lonneke van der Plas♢

♠Institute for Natural Language Processing, University of Stuttgart, Germany

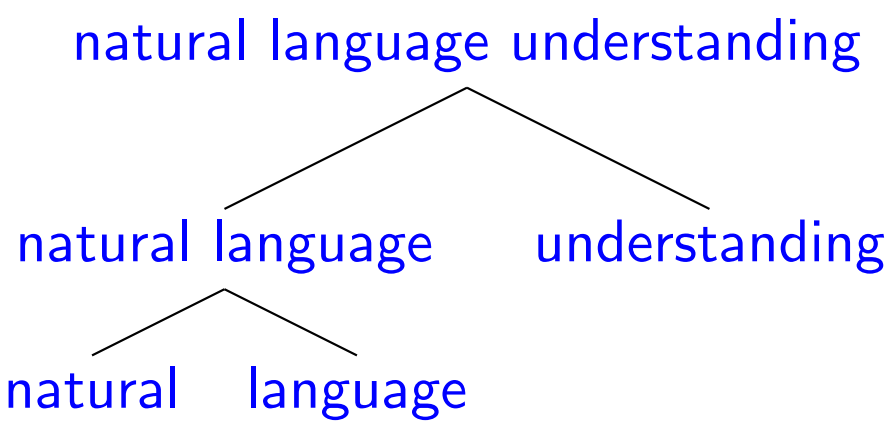
♢Institute of Linguistics, University of Malta, Malta

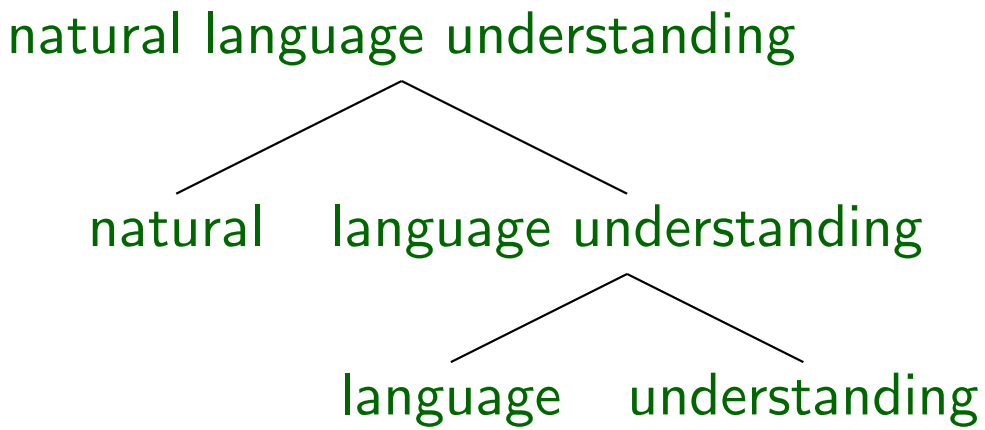
7th September 2015
RANLP 2015

Introduction

Parsing Complex Nominals

- Revealing the internal structure of noun compounds
- Important for natural language understanding





Semantic Indeterminacy

Sometimes: virtually no difference in meaning between structures

- PP-attachment [Hindle and Rooth, 1993]
  - $\Rightarrow$ *They mined the roads along the coast*

Semantic Indeterminacy

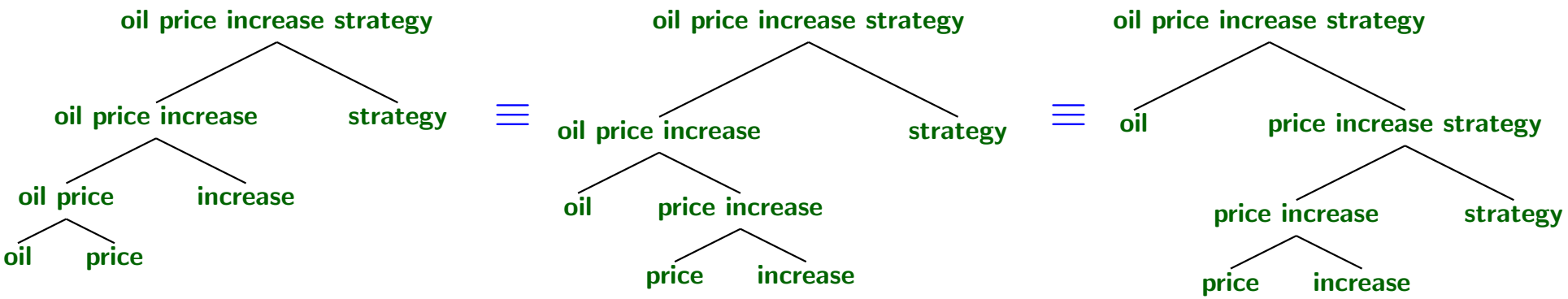
Sometimes: virtually no difference in meaning between structures

- PP-attachment [Hindle and Rooth, 1993]
  - $\Rightarrow$ *They mined the roads along the coast*
- Parsing of 4NCs (e.g., **oil price increase strategy**):

Semantic Indeterminacy

Sometimes: virtually no difference in meaning between structures

- PP-attachment [Hindle and Rooth, 1993]
  - ⇒ *They mined the roads along the coast*
- Parsing of 4NCs (e.g., **oil price increase strategy**):



Semantic Indeterminacy

- Established phenomenon in previous work
- Semantic indeterminacy is most often discarded in syntactic analysis
  - Vadas [2009]
  - Lauer [1995]

Semantic Indeterminacy

- Established phenomenon in previous work
- Semantic indeterminacy is most often discarded in syntactic analysis
  - Vadas [2009]
  - Lauer [1995]
- Incorporating semantic indeterminacy is important for NLP

→ All possible antecedents in anaphora resolution:
oil price, *price increase*, *oil price increase*,
price increase strategy, . . .

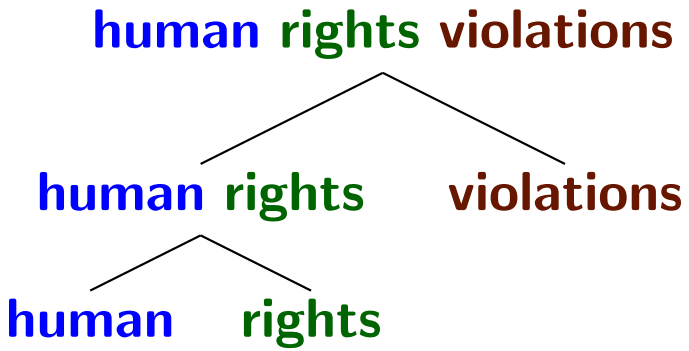
Cross-lingual Structure Transfer (CST)

- Behaghel's [1909] First Law: *Elements which belong close together intellectually will also be placed close together*
- Spelling variations for noun compound translations
- **human rights violations**
  - ⇒ **Verletzungen** *der* **Menschenrechte**

Cross-lingual Structure Transfer (CST)

- Behaghel's [1909] First Law: *Elements which belong close together intellectually will also be placed close together*
- Spelling variations for noun compound translations
- **human rights violations**
  - ⇒ **Verletzungen** *der* **Menschenrechte**

⇒ **human** and **rights** belong closer together:



Evidence for Semantic Indeterminacy in Parallel Corpora

Noun compound translations lead to structure variations

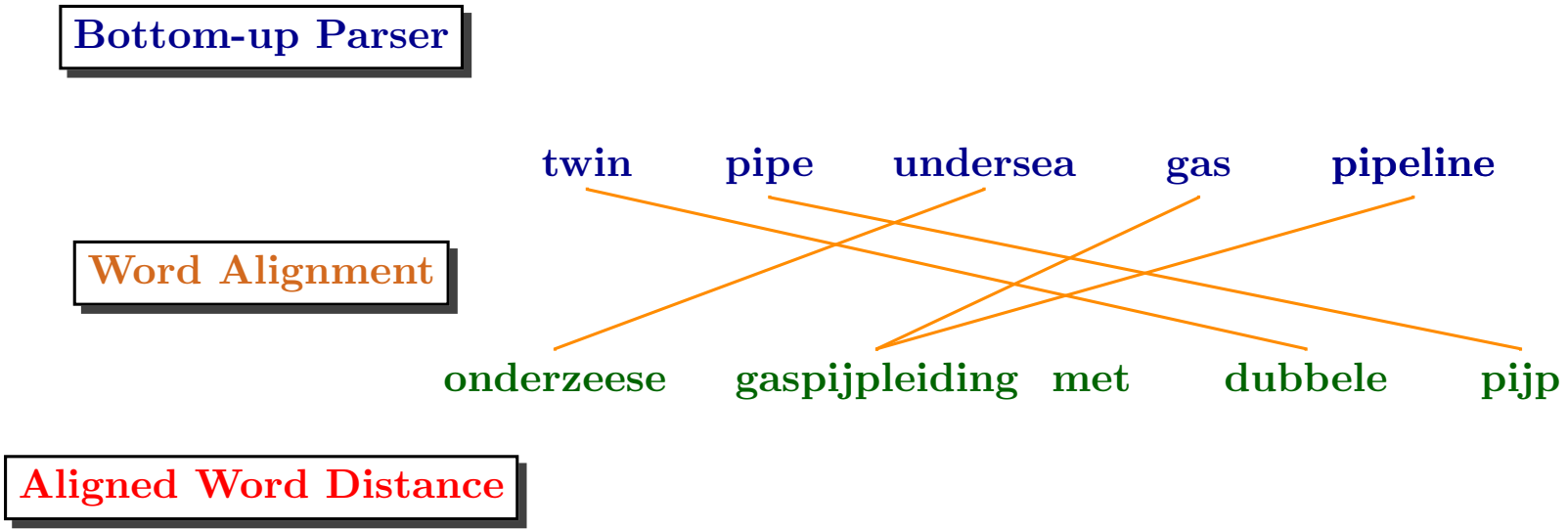
- *tobacco advertising ban*
  - *Werbeverbot für Tabakerzeugnisse* $\Rightarrow$ RIGHT
  - *forbuddet mod tobaksreklamer* $\Rightarrow$ LEFT

These variations are visible in particular across languages

A monolingual perspective suffers from coventional language use

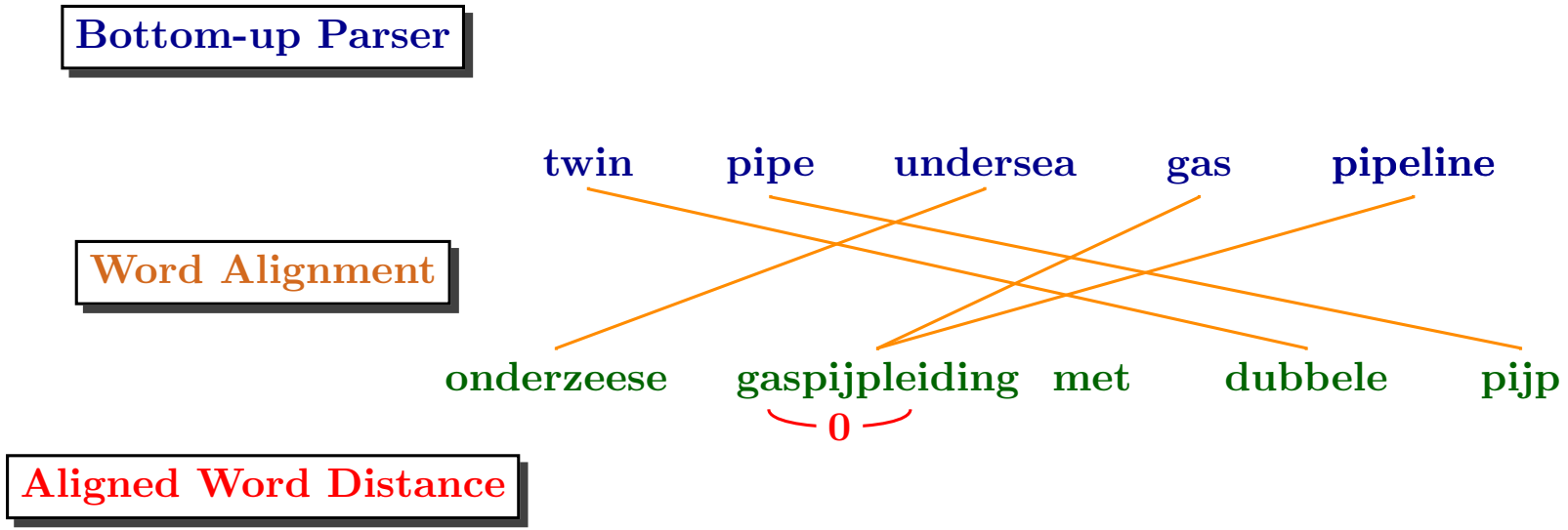
Language	Bigram	freq
	tobacco advertising	**205**
	advertising ban	84
	Tabakwerbung	31
	Werbeverbot	**96**

Language-Isolated Deterministic Structure Transfer



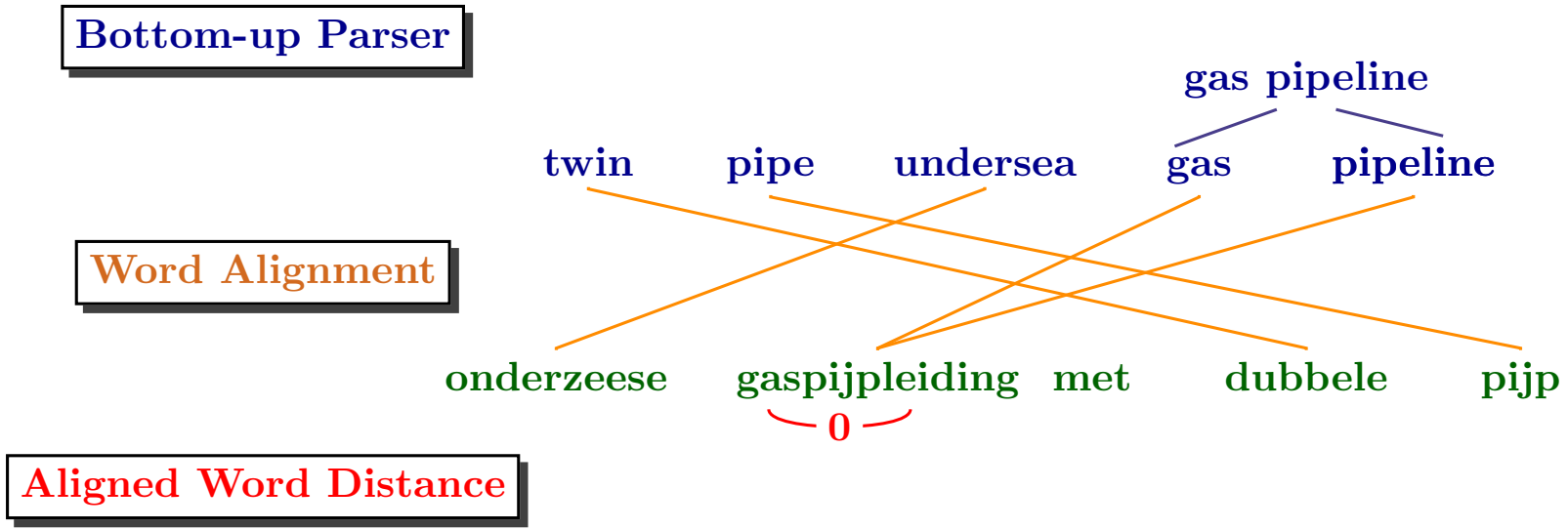
[Ziering and Van der Plas, 2015]

Language-Isolated Deterministic Structure Transfer



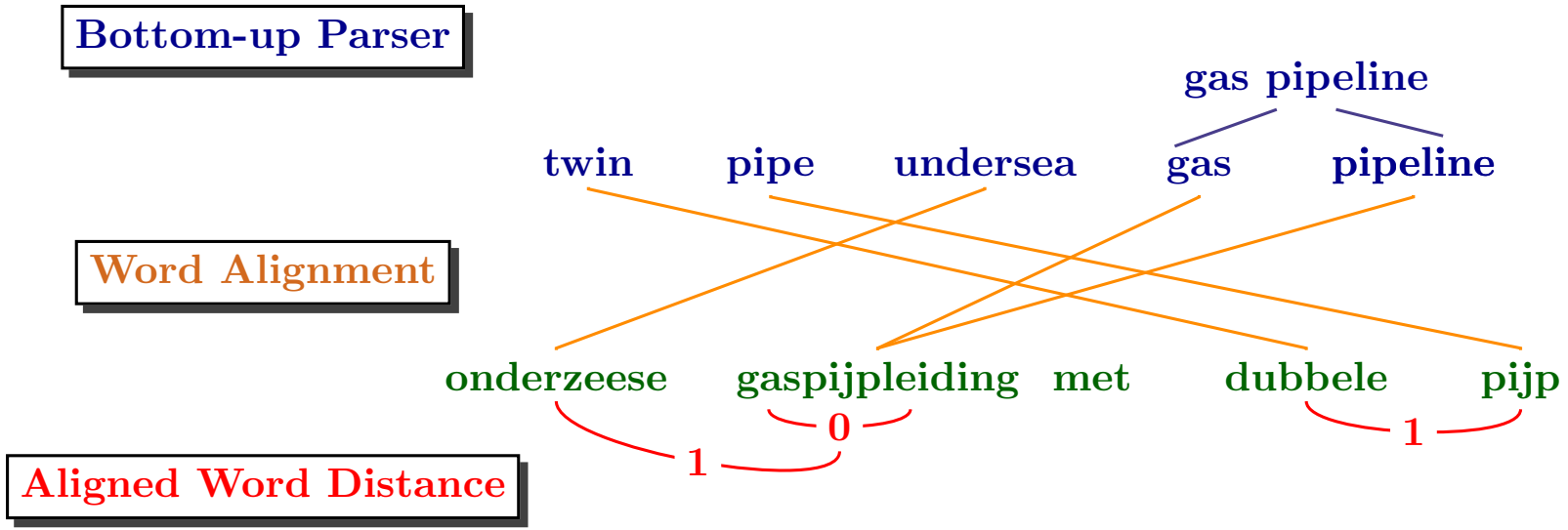
[Ziering and Van der Plas, 2015]

Language-Isolated Deterministic Structure Transfer



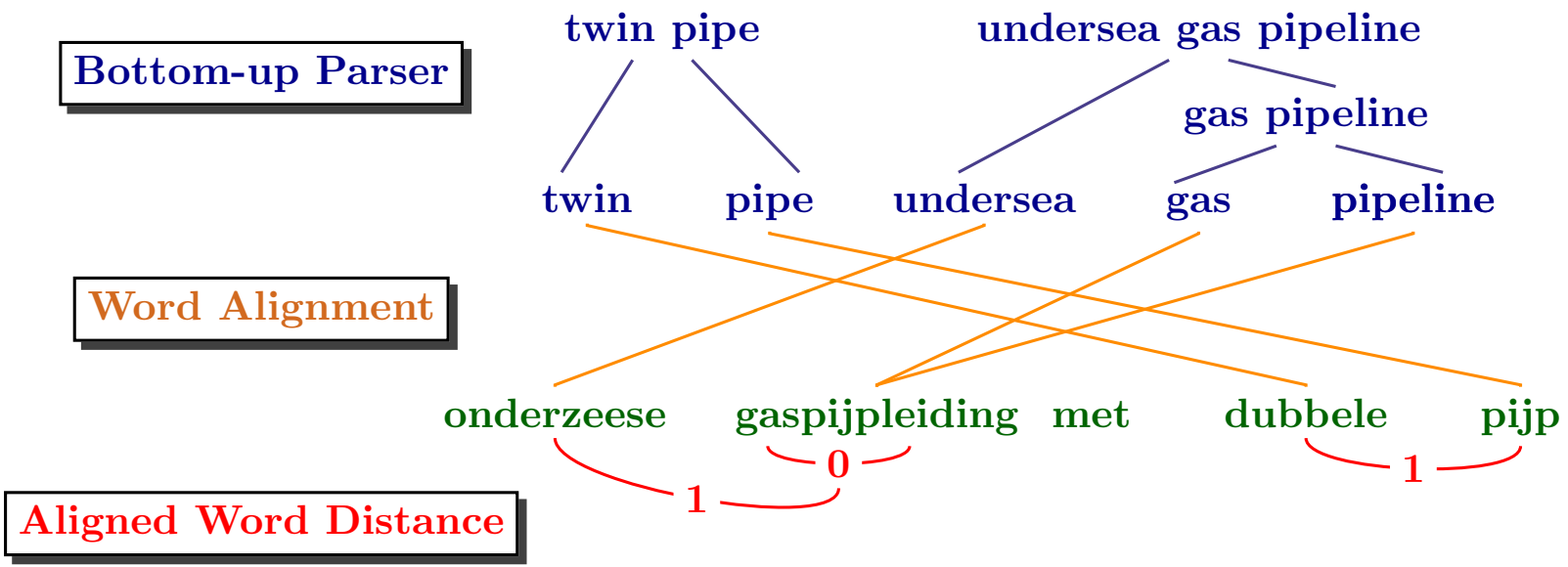
[Ziering and Van der Plas, 2015]

Language-Isolated Deterministic Structure Transfer



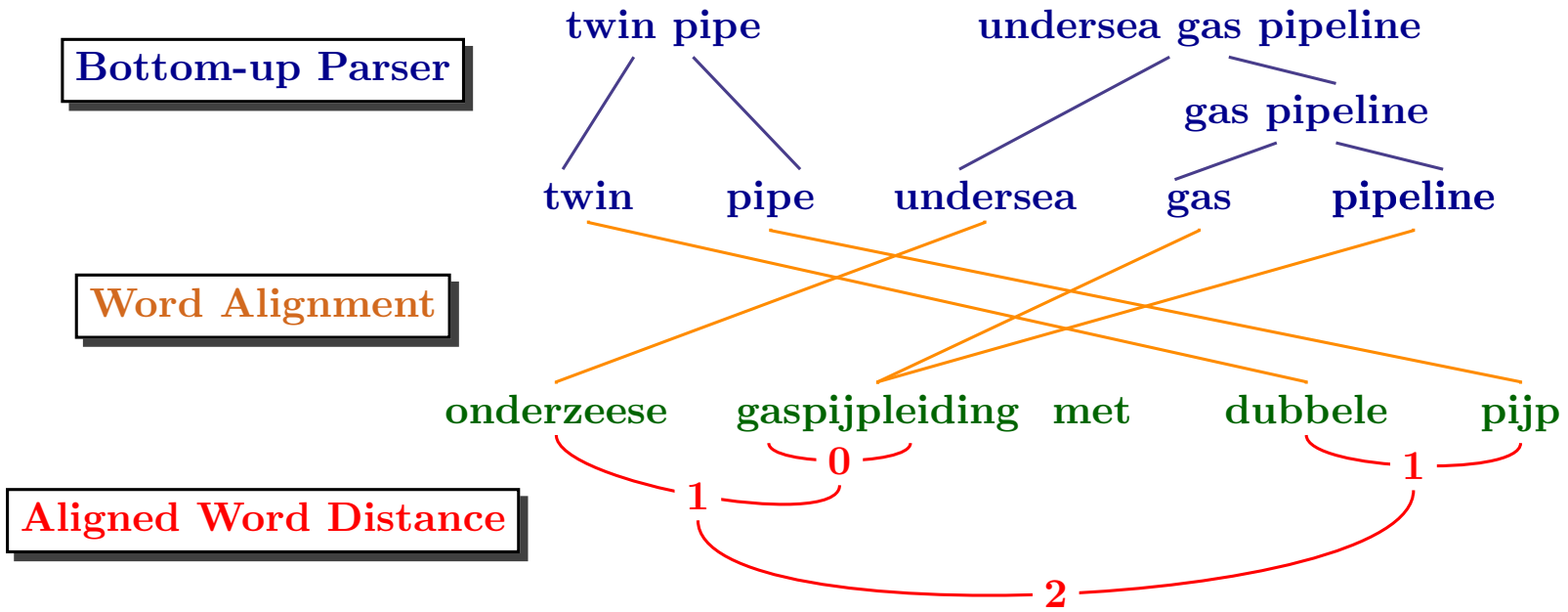
[Ziering and Van der Plas, 2015]

Language-Isolated Deterministic Structure Transfer



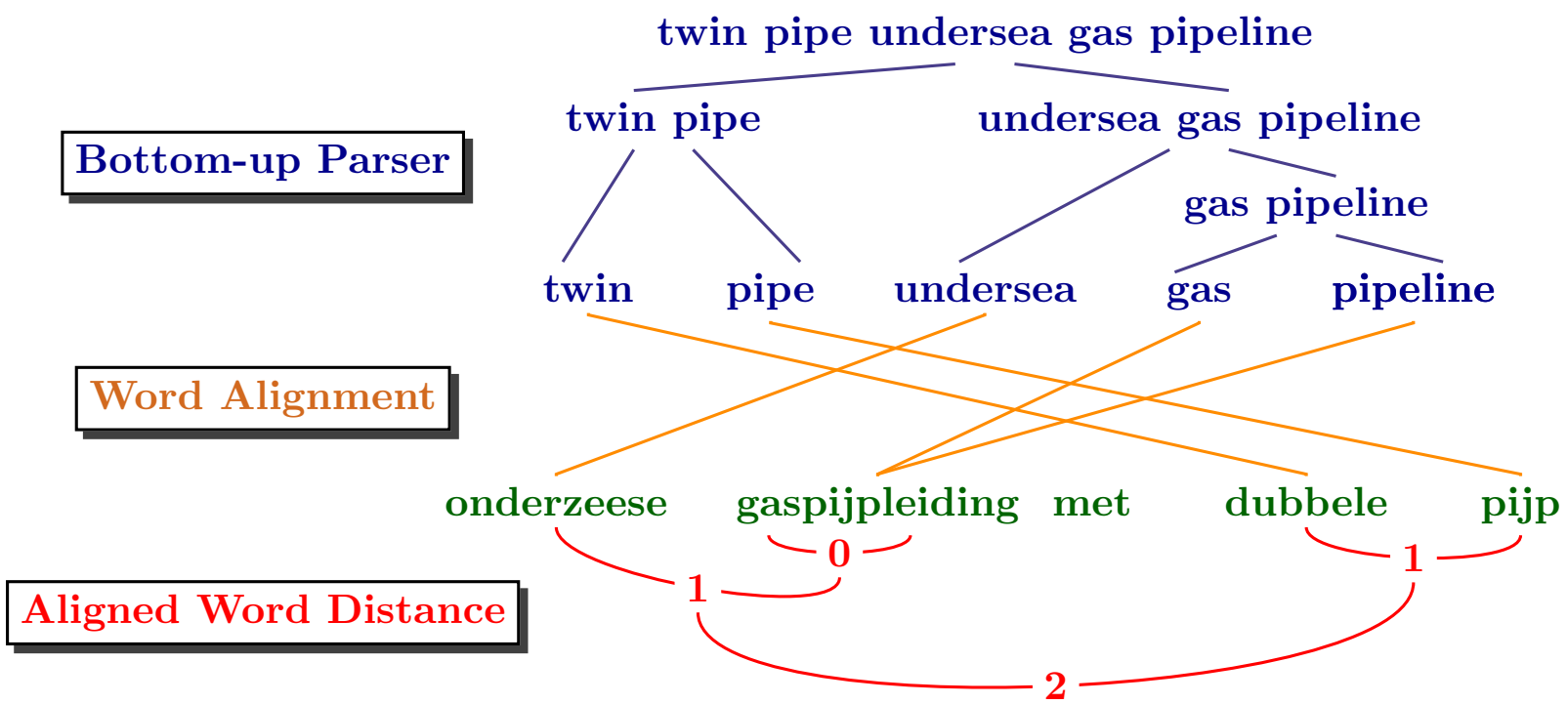
[Ziering and Van der Plas, 2015]

Language-Isolated Deterministic Structure Transfer



[Ziering and Van der Plas, 2015]

Language-Isolated Deterministic Structure Transfer



[Ziering and Van der Plas, 2015]

Language-Isolated Deterministic Structure Transfer

Limitations of LIDST:

(A) Deterministic output for each individual language
→ No output, if overlapping constituents have the same AWD

(B) Deterministic output for language ensemble using majority vote
→ Cannot handle semantic indeterminacy

Tree Accumulative Methods

Tree Accumulative Methods

Motivation:

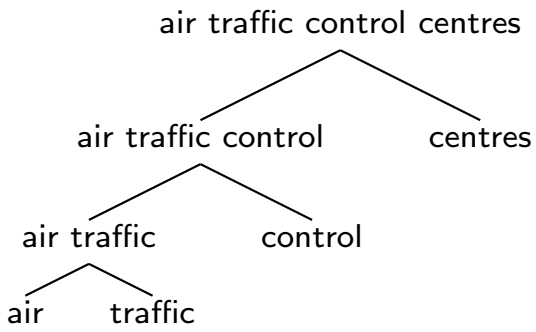
(A) Translations with which overlapping constituents have the same AWD can still provide partial results
$\Rightarrow$ Non-deterministic classification criterion

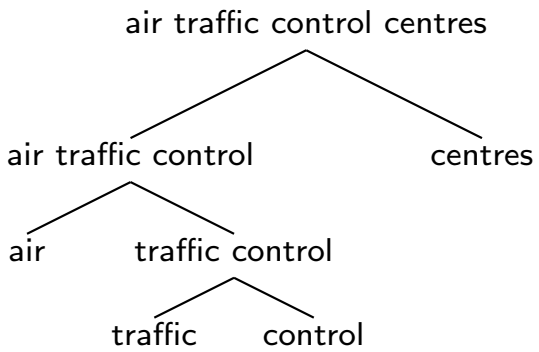
(B) Semantic indeterminacy can only be captured non-deterministically
$\Rightarrow$ Cross-lingual tree accumulation

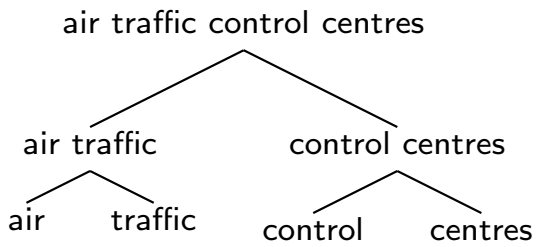
Full Tree Accumulative Structure Transfer (FAST)

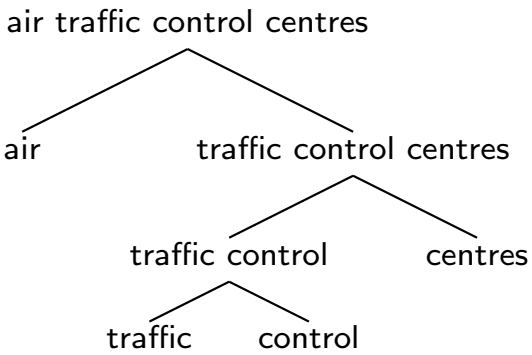
For a given language *l* and *k*-partite noun compound *k*NC:

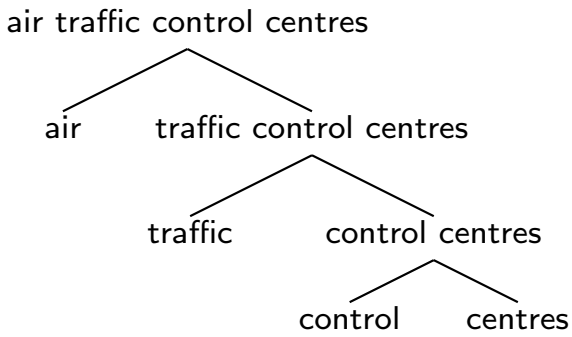
1. We create all binary trees.









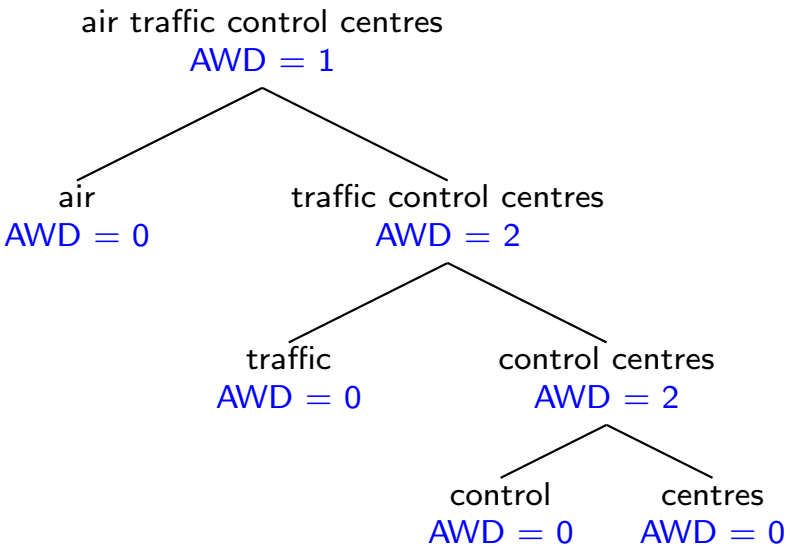


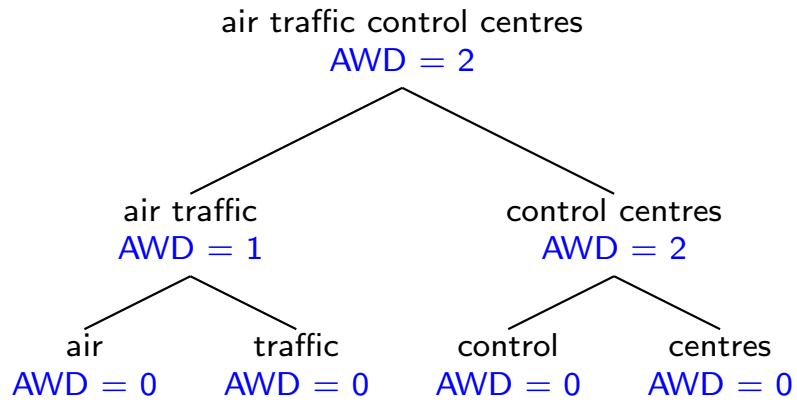
Full Tree Accumulative Structure Transfer (FAST)

2. We annotate all tree nodes N_i with the AWD of their children

$$N_i.AWD = \begin{cases} \text{leaf}(N_i) & \mapsto 0 \\ \text{else} & \mapsto AWD(N_i.\text{L}, N_i.\text{R}) \end{cases}$$

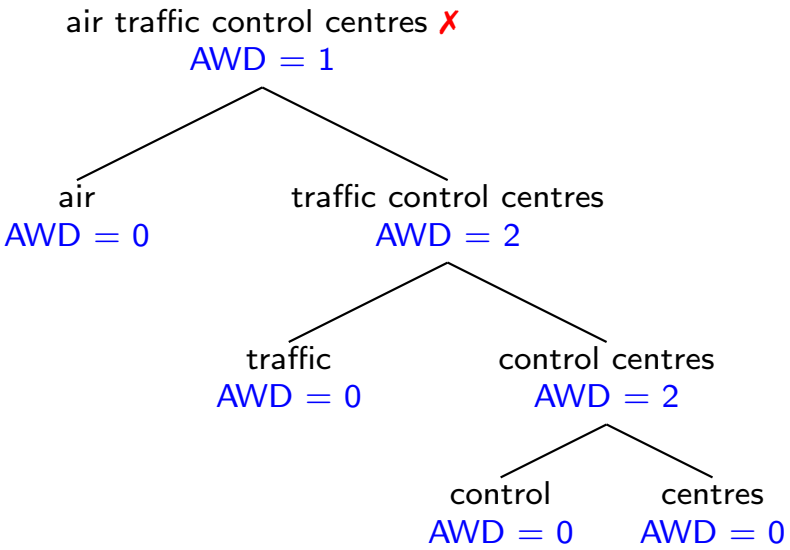
- *air traffic control centres*
  $\rightarrow$ **centres** *de* **contrôle** *du* **trafic aérien**

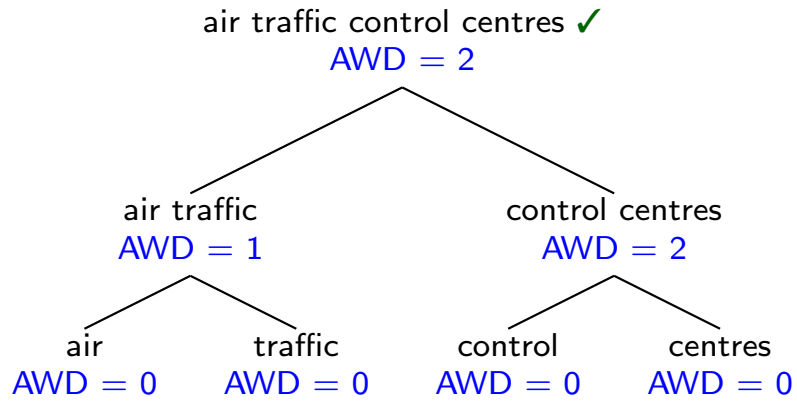




Full Tree Accumulative Structure Transfer (FAST)

3. Tree validation:
   A tree is valid, if the AWD annotation is monotonically decreasing top down.





$\Rightarrow$ Only valid trees are returned

Full Tree Accumulative Structure Transfer (FAST)

4. Tree accumulation of all valid trees from all languages
5. Ranking of all trees by frequency

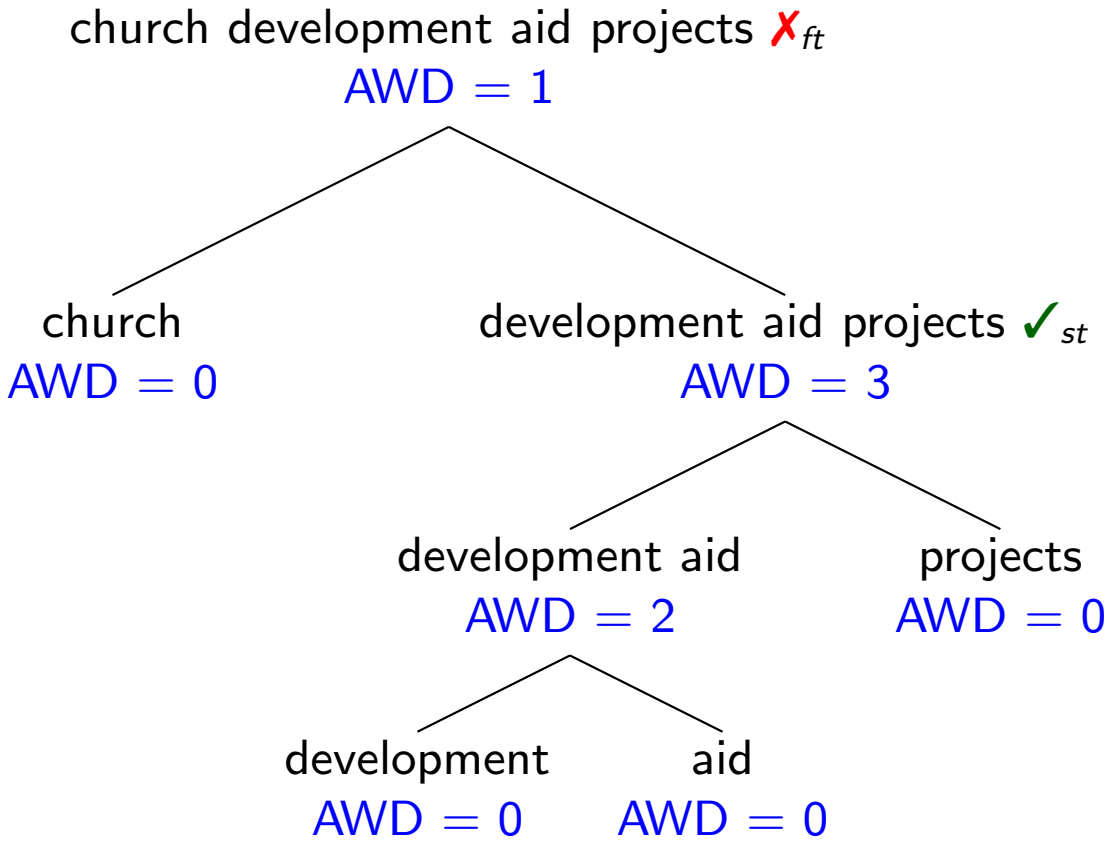
Rank	Structure	Frequency
1	[*air traffic*] [*control centres*]	13
1	[[*air traffic*] *control*] *centres*	13
2	[*air* [*traffic control*]] *centres*	10

Subtree Accumulative Structure Transfer (SAST)

Sometimes an invalid full tree still contains a valid subtree

- *church development aid projects*

  → **progetti ecclesiastici** *di* **aiuti** *allo* **sviluppo**



Subtree Accumulative Structure Transfer (SAST)

- For all languages $l \in L$:
  1. Creation of all binary trees
  2. Tree annotation with AWD
  3. Subtree validation and accumulation
- Subtrees are assigned a subtree score (*sts*):

$$sts(st) = \frac{freq(st.\text{valid})}{|L| \cdot C_{\Delta}}$$

- All full trees are assigned a full tree score (*fts*):

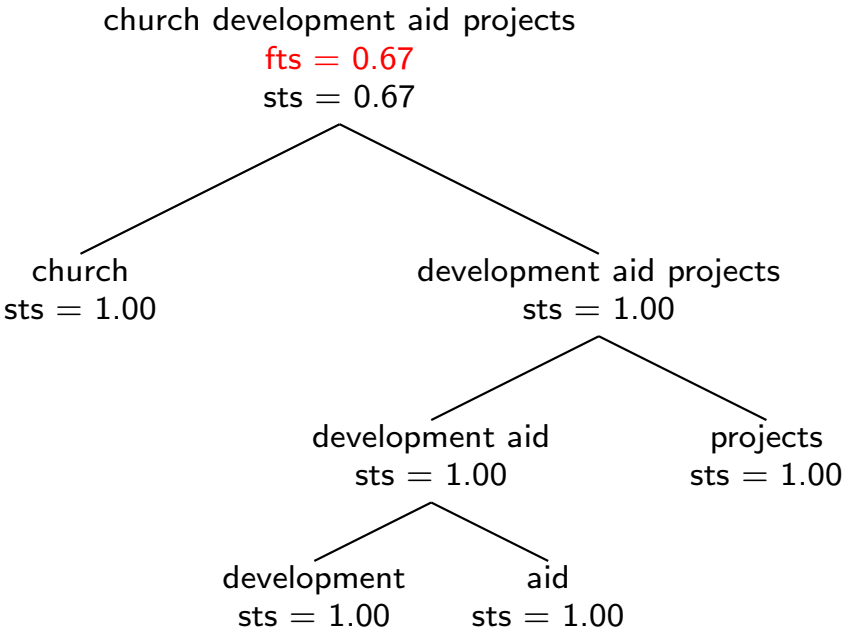
$$fts(ft) = \prod_{st \in ft} sts(st)$$

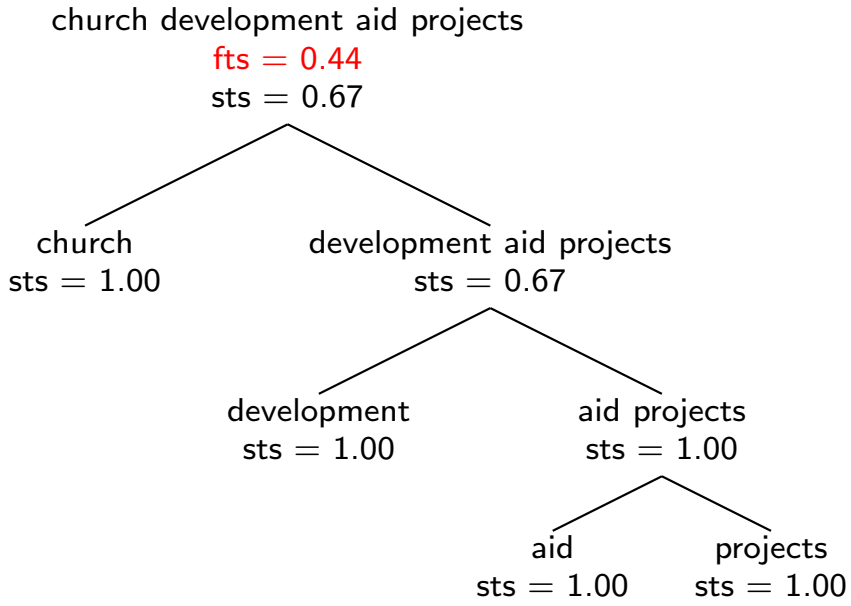
- Full tree ranking according to *fts*

Difference between FAST and SAST

Given the language ensemble {  ,  ,  }:

- SAST ranks the left tree higher by exploiting the valid subtree derived from the Italian translation





Experiments

Dataset

- Noun Compound Database [Ziering and Van der Plas, 2014]
  - OPUS Europarl corpus [Tiedemann, 2012]
  - 10 languages out of 3 families:

Germanic	Romance	Hellenic
English	**French**	**Greek**
Danish	**Italian**	
Dutch	**Portuguese**	
German	**Spanish**	
Swedish		

- Extraction of 3NCs and 4NCs by PoS patterns

24,848	3NC tokens	(16,565 types)
1468	4NC tokens	(1257 types)

Gold Standard

- 3NC test set [Ziering and Van der Plas, 2015]
  - 278 LEFT- or RIGHT-branched 3NC tokens
  - 120 cases of semantic indeterminacy
- 4NC test set
  - 50 4NC tokens
  - Annotation guidelines of Vadas and Curran [2007]
  - Two trained annotators
    - Single tree $(1, \ldots, 5)$
    - Semantic Indeterminacy $[i; \ldots; j]$
  - Single trees $\cup$ Semantic Indeterminacy $\rightarrow$ 33 4NC tokens

Structure Retrieval

- How well does the system ranking fit to the set of gold trees?
- R-Precision [Buckley and Voorhees, 2000]

$$\text{R-Prec}(k\text{NC}) = \frac{|\text{top-R}(\textit{sys trees}) \cap \textit{gold trees}|}{|\text{top-R}(\textit{sys trees})|}$$

$\rightarrow$ Mean R-Precision (MRP) as macro average

Models in Comparison

LIDST Language-Isolated Deterministic Structure Transfer

LINDST Language-Isolated Non-Deterministic Structure Transfer

- Expansion of LIDST
- Frequency ranking instead of majority vote

CHANCE Creates a random tree ranking

FREQ Creates a tree ranking according to structure pattern frequencies

Results

- MRP on test set of 3NCs and 4NCs

System	MRP
FAST	93.7%
SAST	94.0%
LIDST	92.6%
LINDST	92.0%
FREQ	84.6%
CHANCE	62.5%

Results

- MRP on test set of 3NCs and 4NCs

System	MRP
FAST	93.7%
SAST	94.0%
LIDST	92.6%
LINDST	92.0%
FREQ	84.6%
CHANCE	62.5%

- All CST systems outperform the baselines

Results

- MRP on test set of 3NCs and 4NCs

System	MRP
FAST	93.7%
SAST	94.0%
LIDST	92.6%
LINDST	92.0%
FREQ	84.6%
CHANCE	62.5%

- All CST systems outperform the baselines
- FAST and SAST outperform LIDST and LINDST, but differences are small

Results

- Evaluation on test set of 4NCs

System	MRP
FAST	**70.0%**
SAST	69.5%
LIDST	54.5%‡
LINDST	62.9%‡
FREQ	60.1%
CHANCE	32.0%

Results

- Evaluation on test set of 4NCs

System	MRP
FAST	**70.0%**
SAST	69.5%
LIDST	54.5%‡
LINDST	62.9%‡
FREQ	60.1%
CHANCE	32.0%

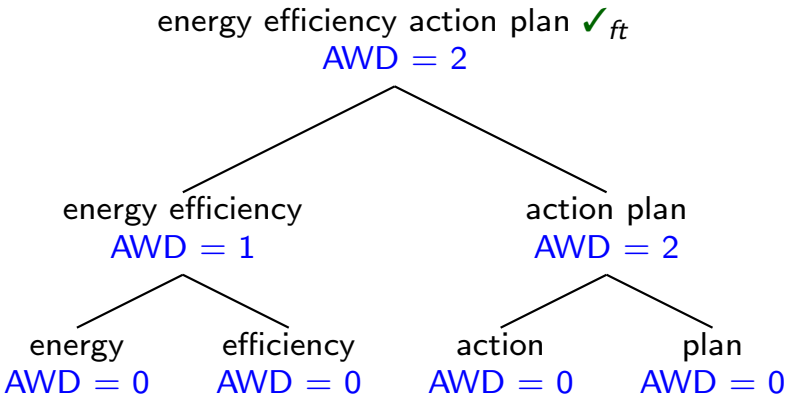
- FAST and SAST significantly outperform LI(N)DST in MRP

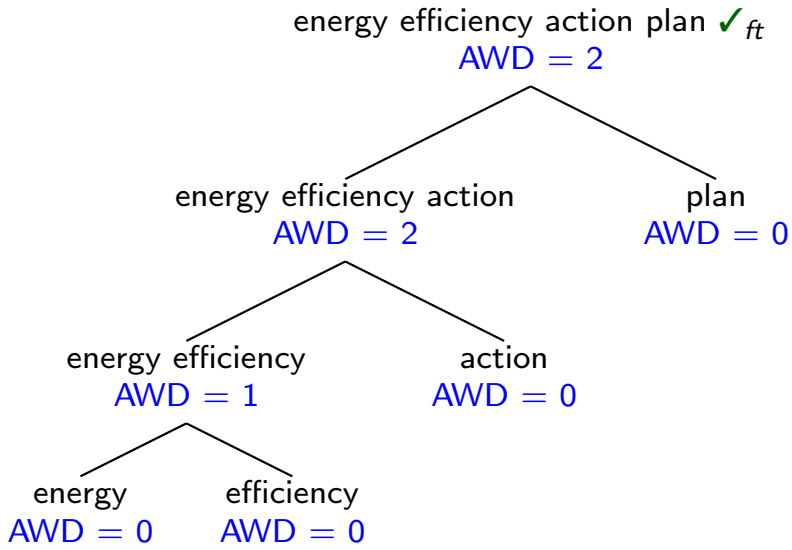
Discussion

Discussion

- Accumulative CST also means a benefit in a deterministic take
  → partial evidence from several languages can be combined

- *energy efficiency action plan*
  → **plan** *de* **acción** *de* **eficiencia energética**

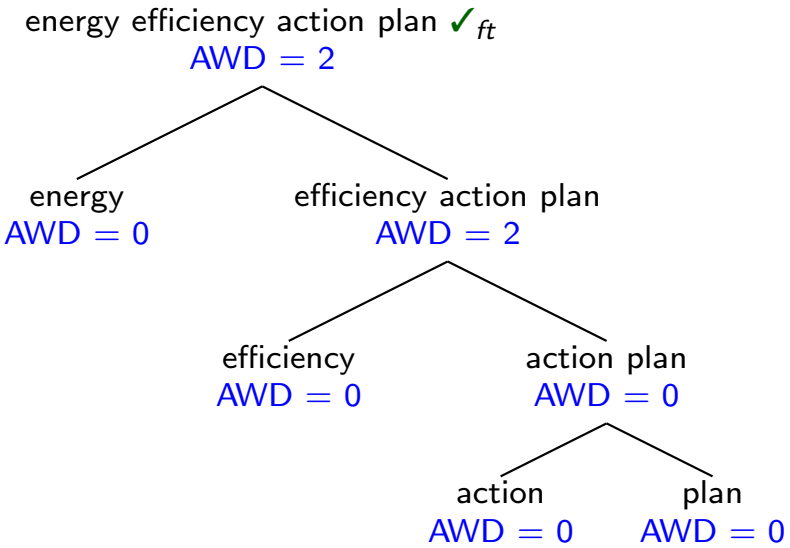


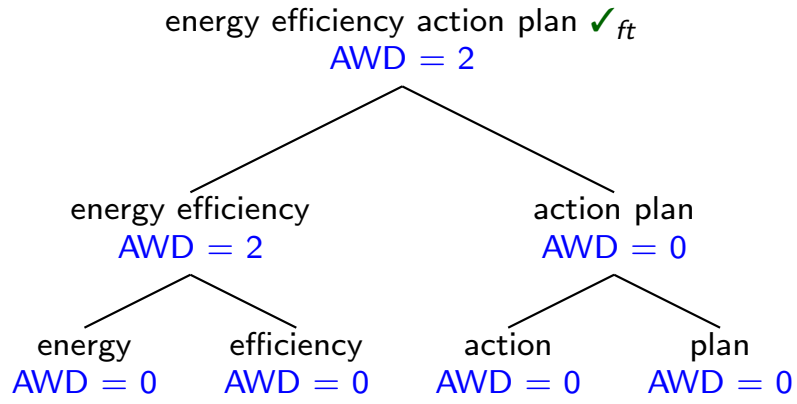


Discussion

- Accumulative CST also means a benefit in a deterministic take
  → partial evidence from several languages can be combined

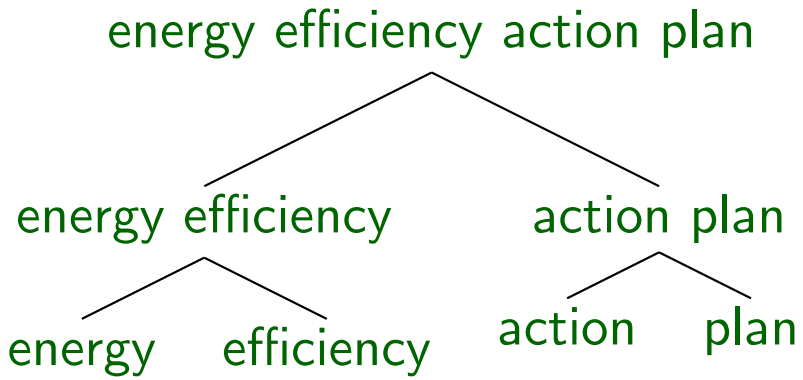
- *energy efficiency action plan*
  → **Aktionsplan** *zur* **Effizienz** *von* **Energie**



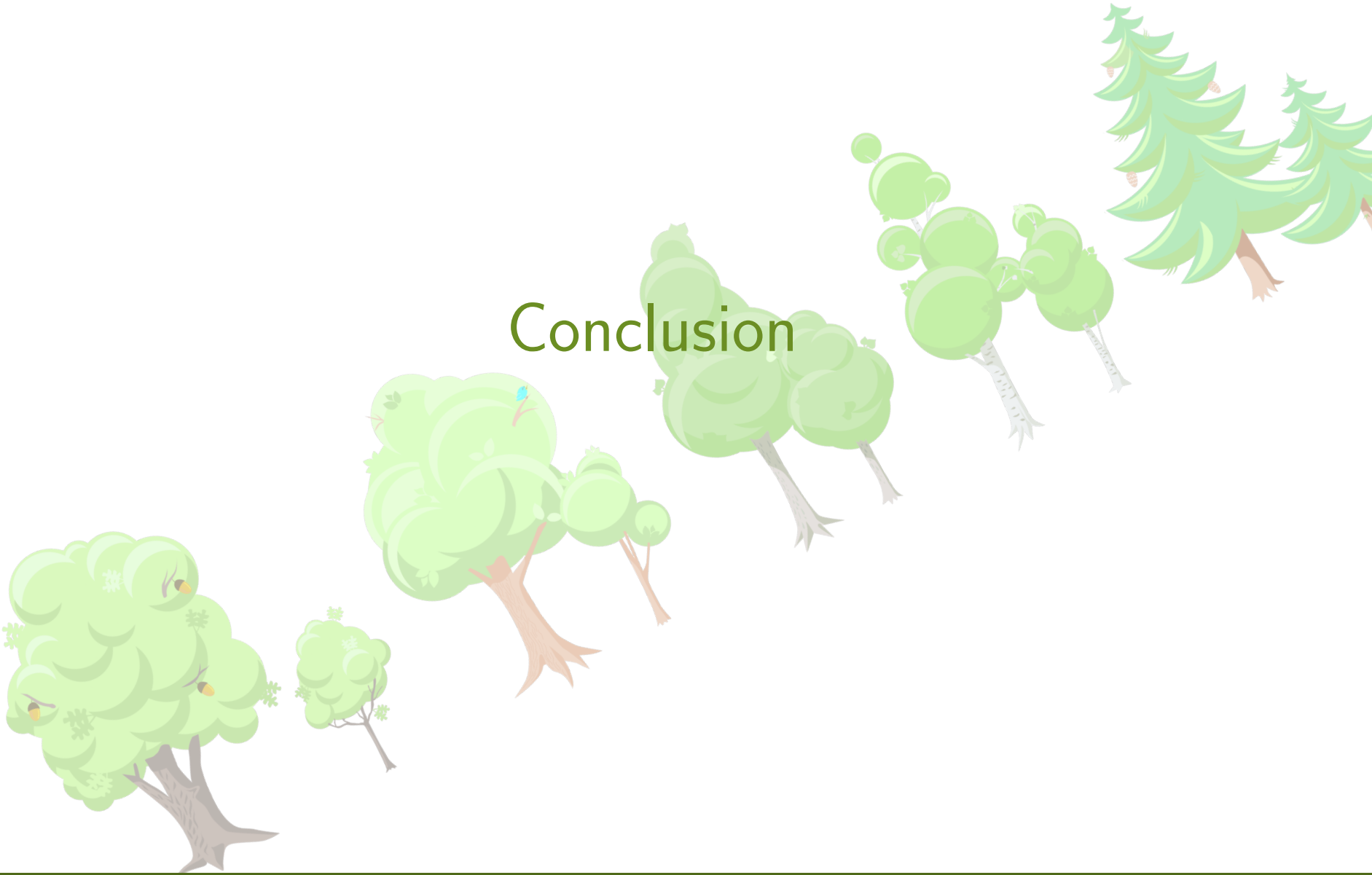


Discussion

- Accumulative CST also means a benefit in a deterministic take
  $\rightarrow$ partial evidence from several languages can be combined
- Both languages fail to provide a single deterministic output
- In contrast:
  the cross-lingual tree accumulation provides the correct tree:



Conclusion



Conclusion

- Two models of accumulative cross-lingual structure transfer
  1. FAST
  2. SAST

  → Outperform previous CST approach significantly on 4NCs

- Non-deterministic CST is suitable for semantic indeterminacy

- Cross-lingual tree accumulation combines partial results in a deterministic take

Full Results Table on 4NC Test Set

System	MRP	MP@1	MR@1	MP@2	MR@2
FAST	**70.0%**	**72.7%**	**47.5%**	60.6%	74.2%
SAST	69.5%	69.7%	44.4%	**63.6%**	**78.8%**
LIDST	54.5%‡	69.7%	44.4%	47.0%‡	59.1%‡
LINDST	62.9%‡	69.7%	44.4%	54.5%†	66.7% †
FREQ	60.1%	63.6%	38.4%	56.1%	65.2%
CHANCE	32.0%	39.4%	23.7%	33.3 %	42.4 %

- FAST and SAST significantly outperform LI(N)DST in MRP

MP/R@k: Macro average of Precision/Recall at k

Full Results Table on 4NC Test Set

System	MRP	MP@1	MR@1	MP@2	MR@2
FAST	**70.0%**	**72.7%**	**47.5%**	60.6%	74.2%
SAST	69.5%	69.7%	44.4%	**63.6%**	**78.8%**
LIDST	54.5%‡	69.7%	44.4%	47.0%‡	59.1%‡
LINDST	62.9%‡	69.7%	44.4%	54.5%†	66.7% †
FREQ	60.1%	63.6%	38.4%	56.1%	65.2%
CHANCE	32.0%	39.4%	23.7%	33.3 %	42.4 %

- FAST and SAST significantly outperform LI(N)DST in MRP
- No difference between the CST systems in MP/R@1

MP/R@k: Macro average of Precision/Recall at k

Full Results Table on 4NC Test Set

System	MRP	MP@1	MR@1	MP@2	MR@2
FAST	**70.0%**	**72.7%**	**47.5%**	60.6%	74.2%
SAST	69.5%	69.7%	44.4%	**63.6%**	**78.8%**
LIDST	54.5%‡	69.7%	44.4%	47.0%‡	59.1%‡
LINDST	62.9%‡	69.7%	44.4%	54.5%†	66.7% †
FREQ	60.1%	63.6%	38.4%	56.1%	65.2%
CHANCE	32.0%	39.4%	23.7%	33.3 %	42.4 %

- FAST and SAST significantly outperform LI(N)DST in MRP
- No difference between the CST systems in MP/R@1
- FAST and SAST significantly outperform LIDST in MP/R@2

MP/R@k: Macro average of Precision/Recall at k