

Overview

Procedural text

1. 両手鍋で油を熱する。
セロリと青ねぎとニンニクを加え、
1分ほど炒める。
2. ブイヨンと水とマカロニと胡椒を加えて、
パスタが柔らかくなるまで煮る。
3. 刻んだセージをまぶす。

Today's
Talk

Flow graph [Mori+, LREC14]



Cooking robot

[Bollini+, ISER13]



www.denso.co.jp

Smart kitchen

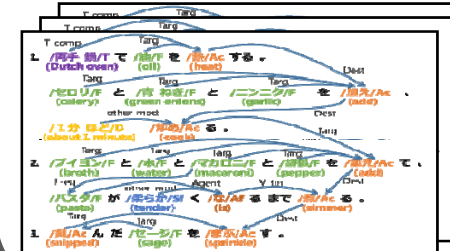
[Hashimoto+, IPMU08]



Cut garlic
Then add it

Intelligent search

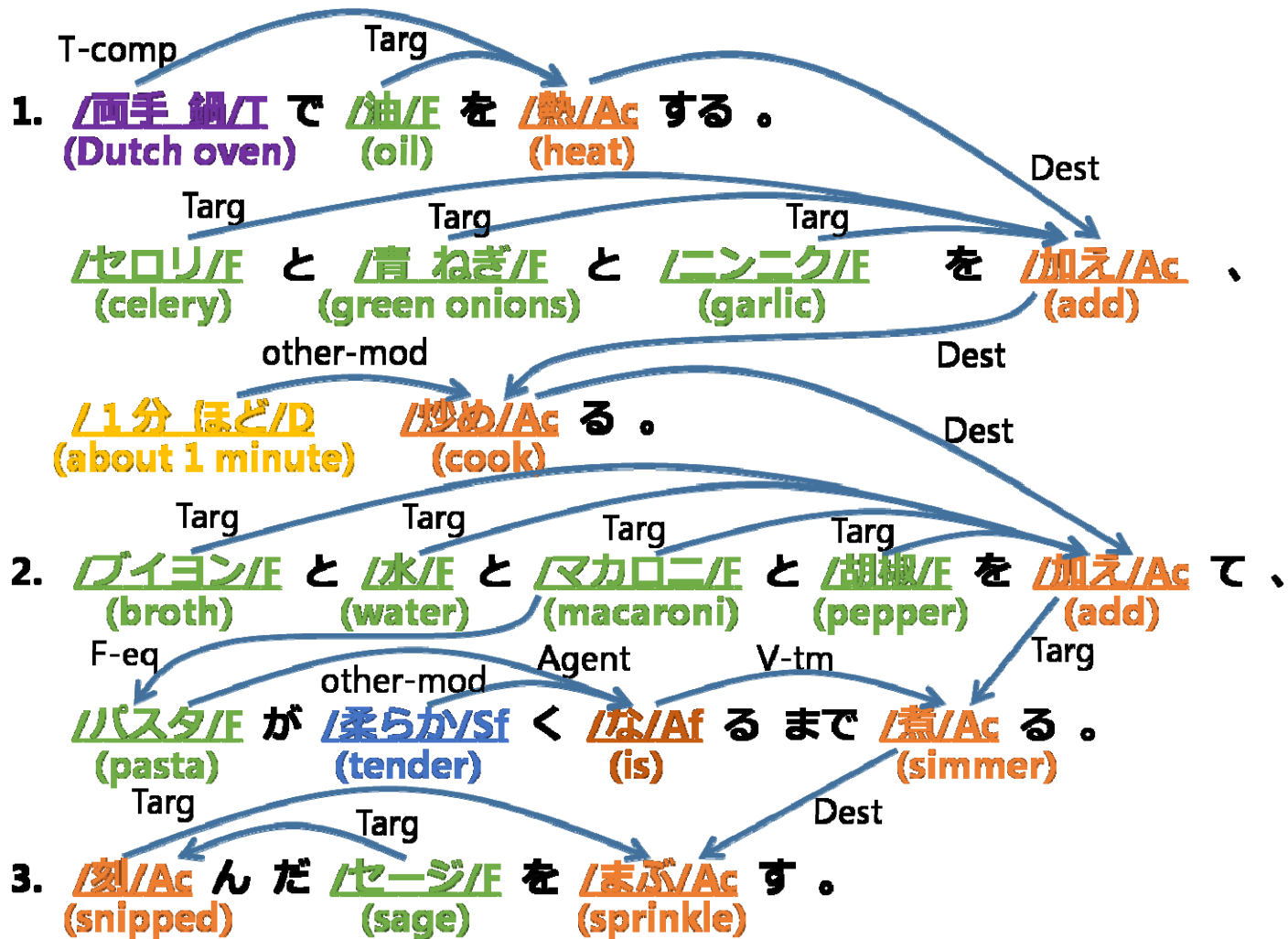
[Yamakata+, SocInfo13]



FLOW GRAPHS FOR PROCEDURAL TEXT

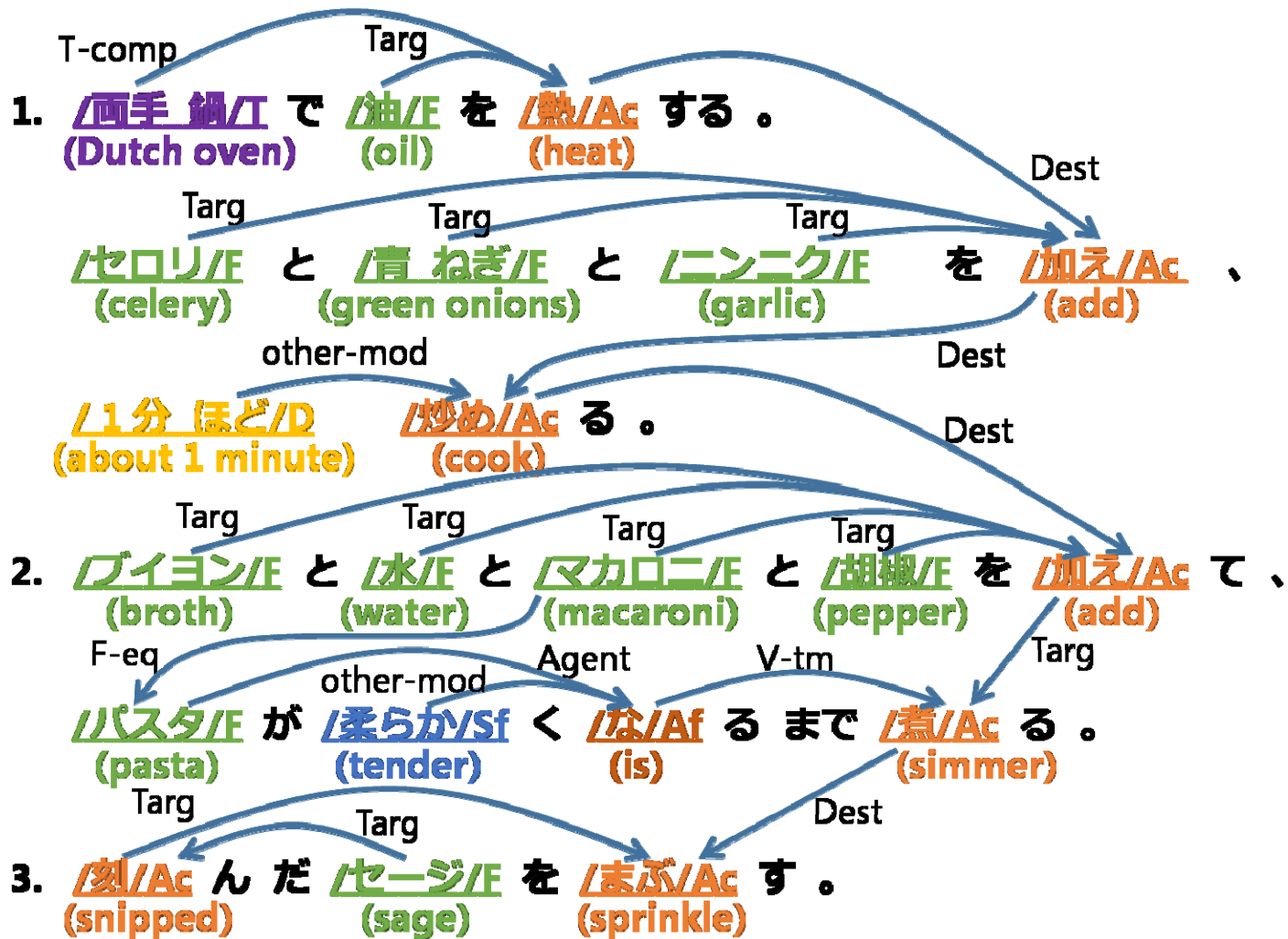
Flow Graph

- Direct acyclic graph (DAG) representing work flow



Vertices

- Concepts = Important terms (word sequence)

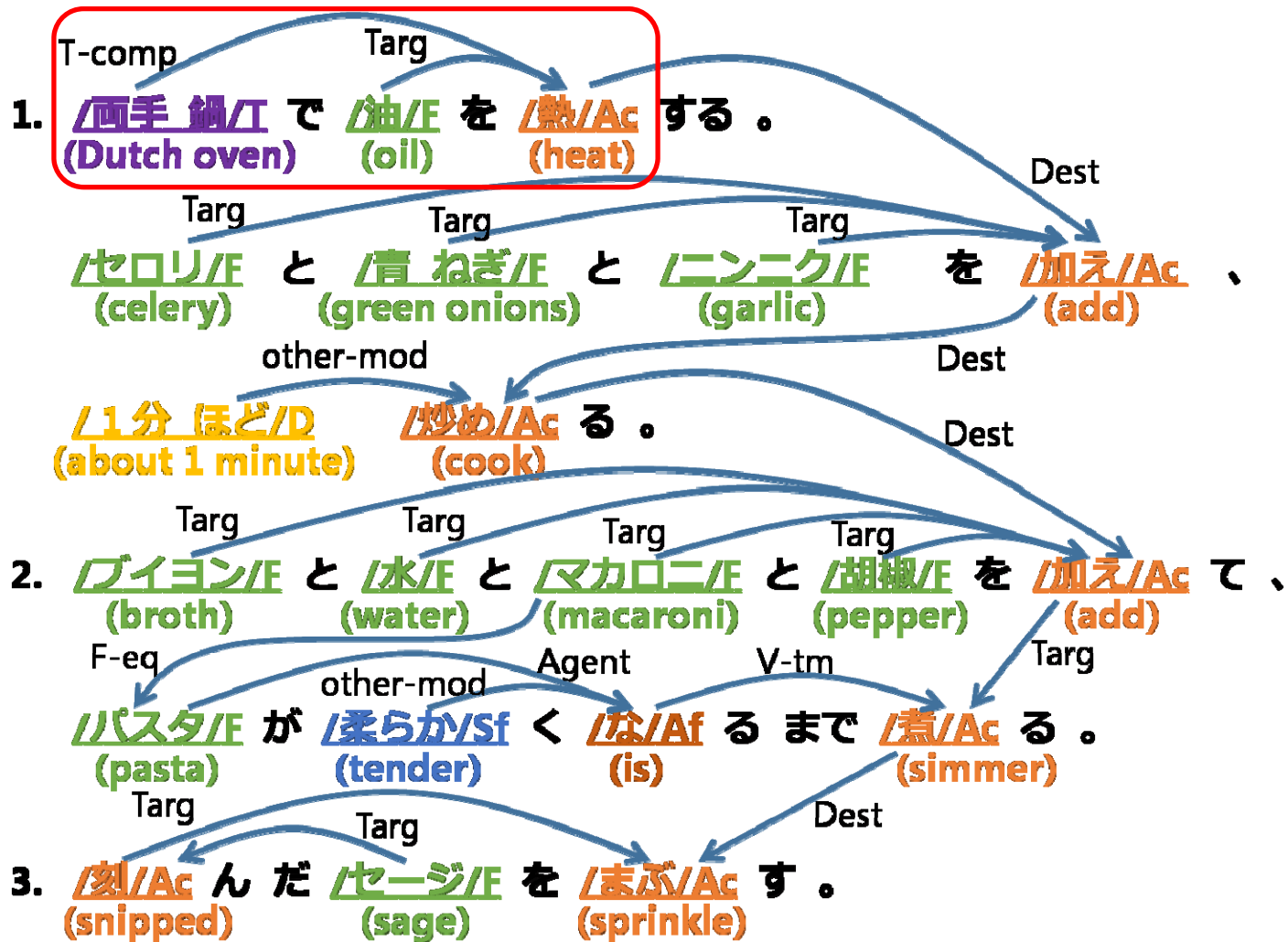


Vertex Labels

Tag	Meaning	Remarks
F	Food	Eatables including intermediate products,
T	Tool	Knife, container, etc.
D	Duration	Duration for cooking
Q	Quantity	Quantity of food
Ac	Action by chef	Verbs representing chef's actions (stem only)
Af	Action by food	Verbs representing food's actions (stem only)
Sf	Food state	Food's initial or intermediate states
St	Tool state	Tool's initial or intermediate states

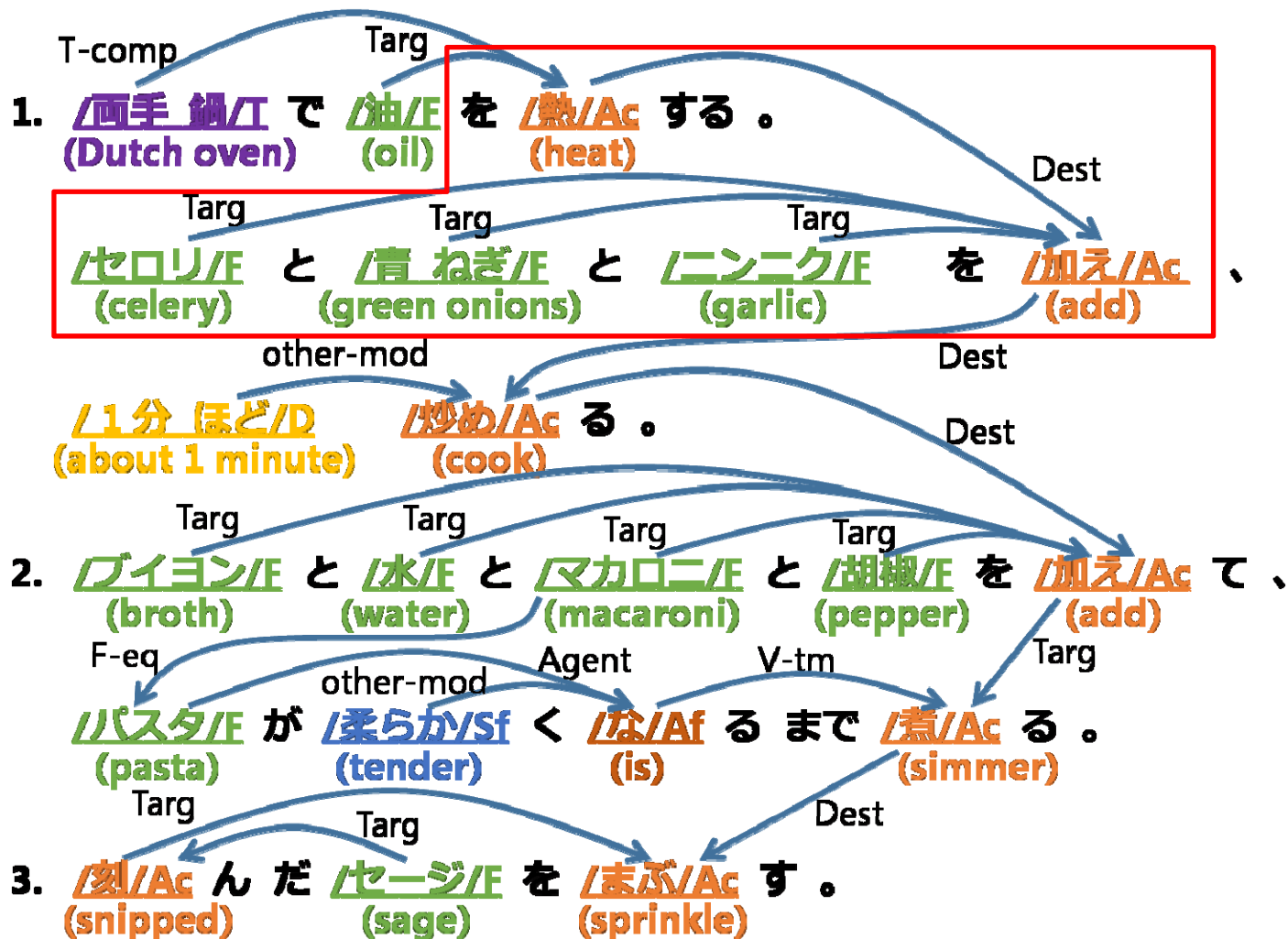
Arcs

- Relationships between two important terms



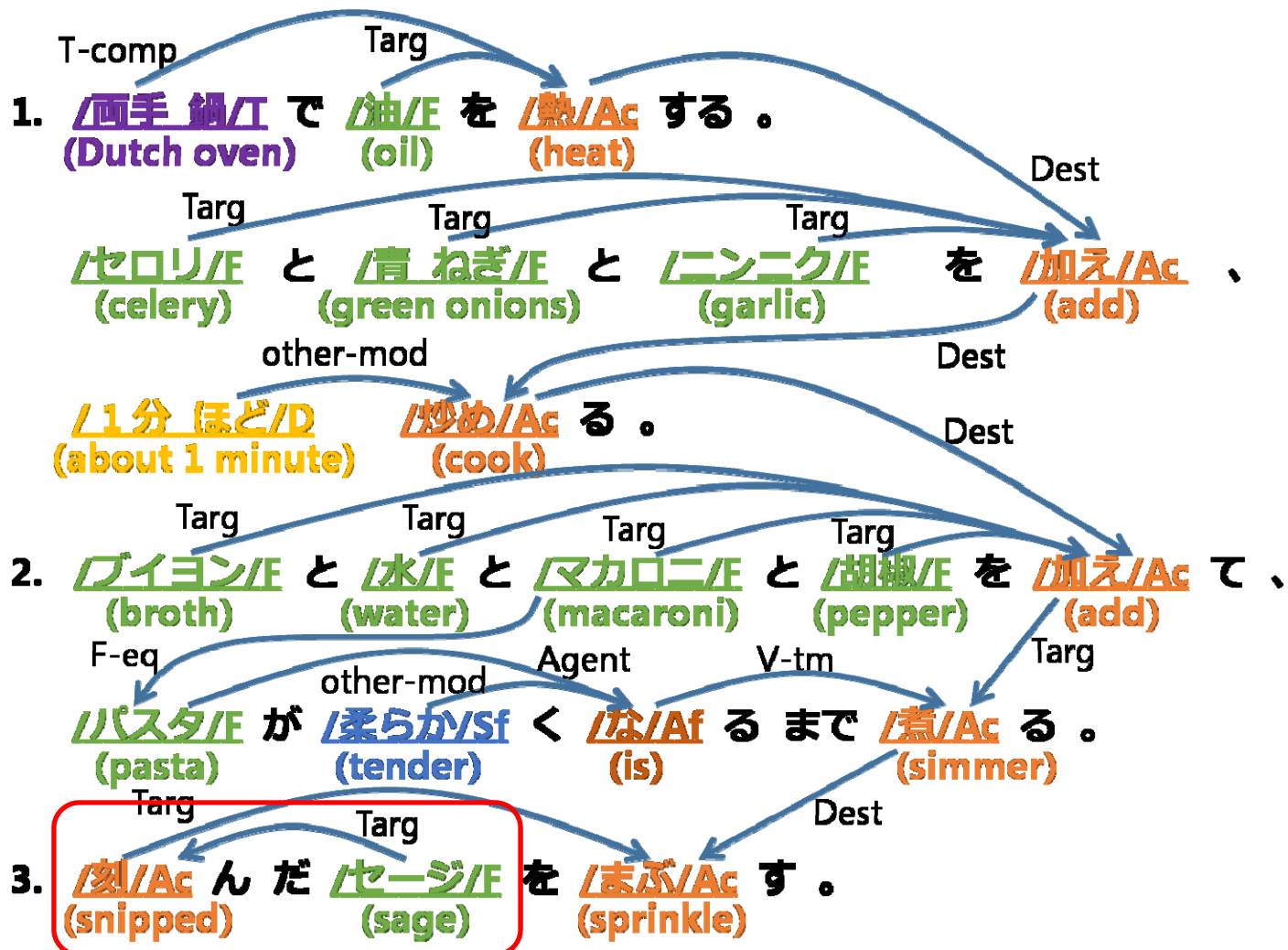
Arcs

- Relationships between two important terms



Arcs

- Relationships between two important terms



Arc Labels

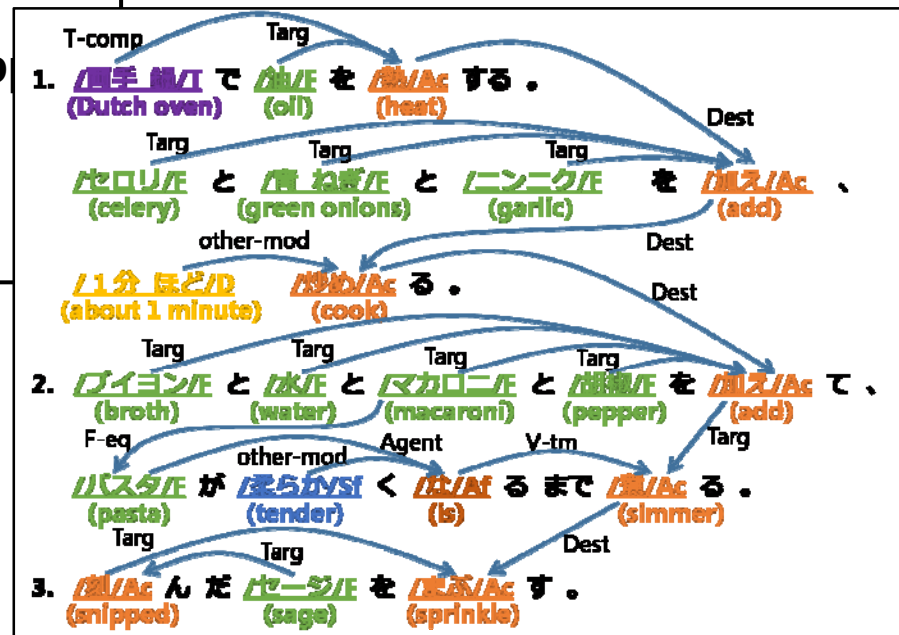
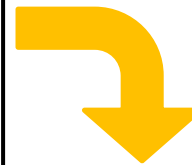
Label	Meaning	Remark
Agent	Action agent	Relationship with actions (Ac or Af)
Targ	Action target	
Dest	Action destination	
T-comp	Tool complement	Tool used in an action
F-comp	Food complement	Food used as a tool
F-eq	Food equality	Identical food
F-part-of	Food part-of	Refer to a part of a food
F-set	Food set	Refer to a set of foods
T-eq	Tool equality	Identical tool
T-part-of	Tool part-of	Refer to a part of a tool
A-eq	Action equality	Identical action (Ac, Af)
V-tm	Head verb for timing, etc.	
other-mod	Other relationships	

PROCEDURAL TEXT UNDERSTANDING

Procedural Text Understanding

1. 両手鍋で油を熱する。
 (In a Dutch oven, heat oil.)
 セロリと青ねぎとニンニクを加え、
 (Add celery, green onions, and garlic.)
 1分ほど炒める。
 (Cook for about 1 minute.)
2. ブイヨンと水とマカロニと胡椒を加えて、
 パスタが柔らかくなるまで煮る。
 (Add broth, water, macaroni, and pep
 and simmer until the pasta is tender.)
3. 刻んだセージをまぶす。
 (Sprinkle the snapped sage.)

1. Word segmentation
2. Concept identification
3. Flow graph estimation



Procedural Text

1. 両手鍋で油を熱する。

(In a Dutch oven, heat oil.)

セロリと青ねぎとニンニクを加え、1分ほど炒める。

(Add celery, green onions, and garlic. Cook for about 1 minute.)

2. ブイヨンと水とマカロニと胡椒を加えて、

(Add broth, water, macaroni, and pepper,

パスタが柔らかくなるまで煮る。

and simmer until the pasta is tender.)

3. 刻んだセージをまぶす。

(Sprinkle the snipped sage.)

Step 1. Word Segmentation

Result:

1. **両手鍋で油を熱する。**
(In a Dutch oven, heat oil.)

セロリと青ねぎとニンニクを加え、1分ほど炒める。
(Add celery, green onions, and garlic. Cook for about 1 minute.)

2. **ブイヨンと水とマカロニと胡椒を加えて、**
(Add broth, water, macaroni, and pepper,

パスタが柔らかくなるまで煮る。
and simmer until the pasta is tender.)

3. **刻んだセージをまぶす。**
(Sprinkle the snipped sage.)

Step 1. Word Segmentation

- Pointwise solution [Neubig+, ACL11]
 - Binary classification problem at each point
 - Trainable from partially segmented sentences
 - KyTea: <http://www.phontron.com/kytea/>

両手鍋で油を熱する。

? ? ? ? ? ? ? ? ?
N Y Y Y Y Y Y N Y



両手鍋で油を熱する。

Step 2. Concept Identification

Result:

1. /両手 鍋/T で /油/F を /熱/Ac する。
(Dutch oven) (oil) (heat)

/セロリ/F と /青ねぎ/F と /ニンニク/F を /加え/Ac 、
(celery) (green onions) (garlic) (add)

/1分 ほど/D /炒め/Ac する。
(about 1 minute) (cook)

2. /ブイヨン/F と /水/F と /マカロニ/F と /胡椒/F を /加え/Ac て、
(broth) (water) (macaroni) (pepper) (add)

/パスタ/F が /柔らか/Sf くなるまで /煮/Ac する。
(pasta) (tender) (is) (simmer)

3. /刻/Ac んだ /セージ/F を /まぶ/Ac する。
(snipped) (sage) (sprinkle)

Step 2. Concept identification

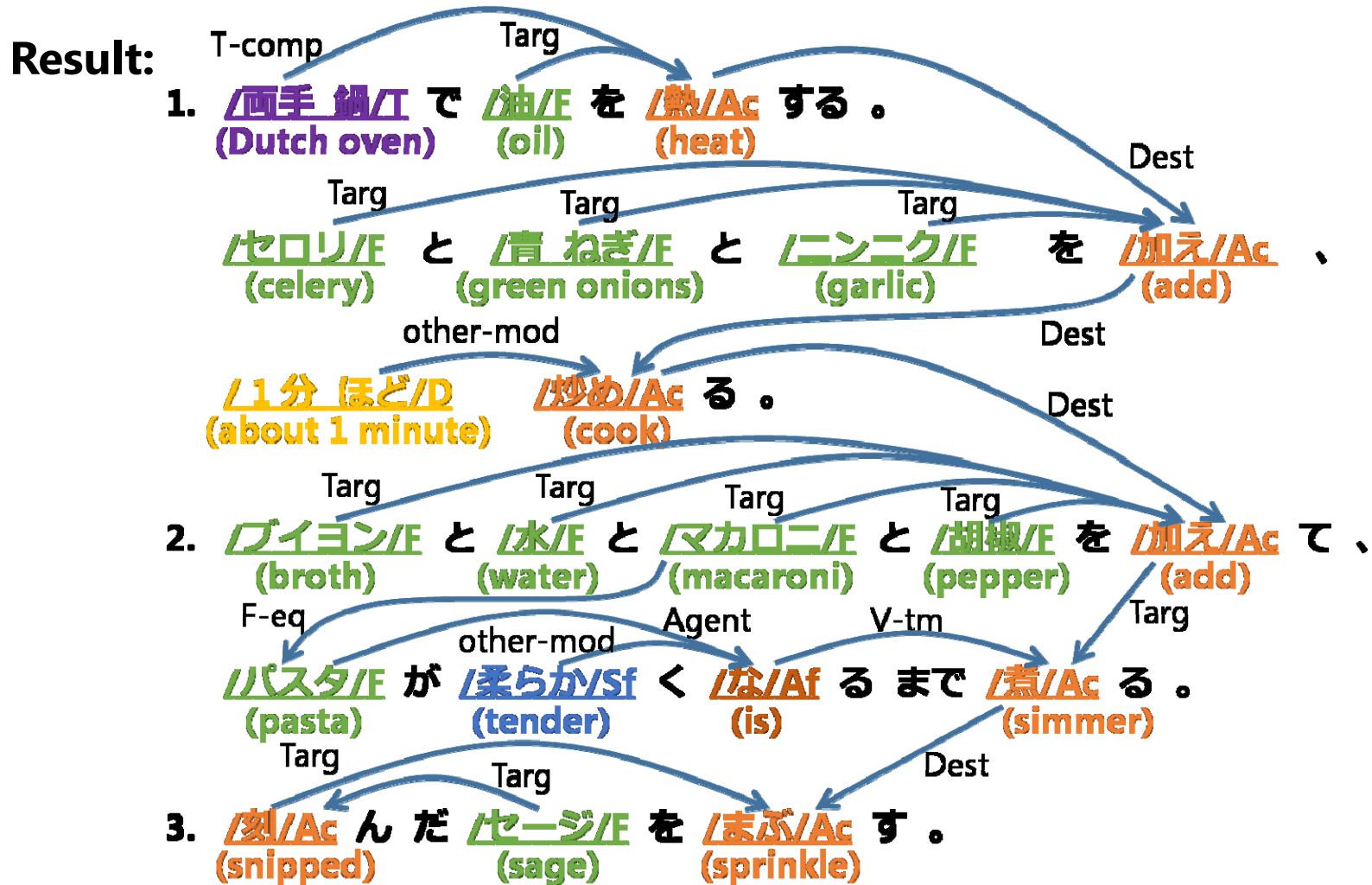
- Same as NER [Mori+, CwC12]
 1. BIOES tagging for each word with a score
 2. Search for the best sequence

– PWNER: <http://plata.ar.media.kyoto-u.ac.jp/tool/>

prob.	両手	鍋	で	油	を	熱	する	。
T-B	0.9	0.2	0	0	0	0	0	0
T-I	0.1	0.8	0	0	0	0	0	0
F-B	0	0	0	0.8	0	0	0	0
...								
Other	0	0	1	0	1	0	0	0

Step 3. Flow Graph Estimation

Parsing of a document, not a sentence



Step 3. Flow Graph Estimation

1. Maximum Spanning Tree search

- Logistic regression (LR) scores
- 1st order

2. Arc addition

- DAG constrains
- LR score \leq penalty(n)

Step 3-1. Tree Estimation

MST using scores by Logistic Regression

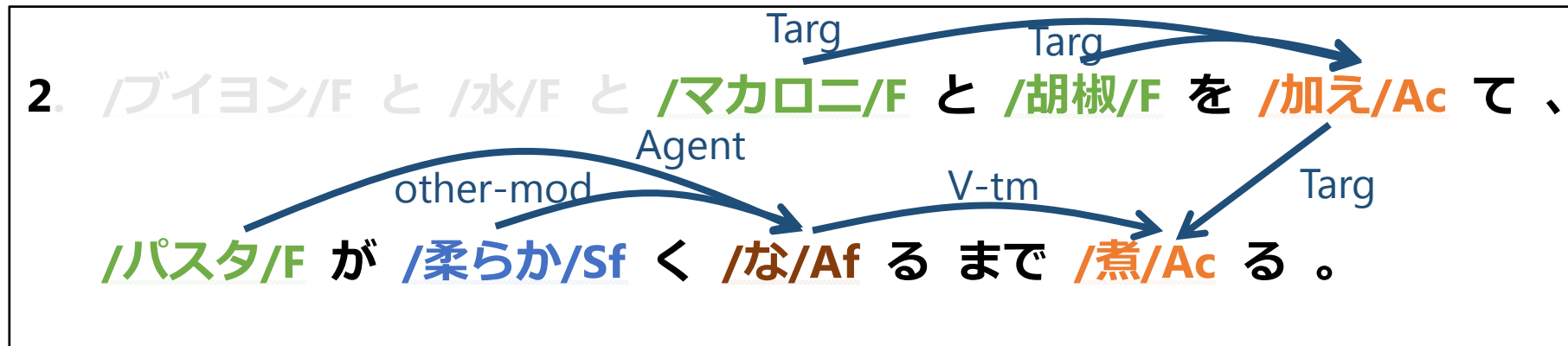
2. /ブイヨン/F と /水/F と /マカロニ/F と /胡椒/F を /加え/Ac て、
/パスタ/F が /柔らか/sf く /な/Af る まで /煮/Ac る。

↓ LR score + **MST**

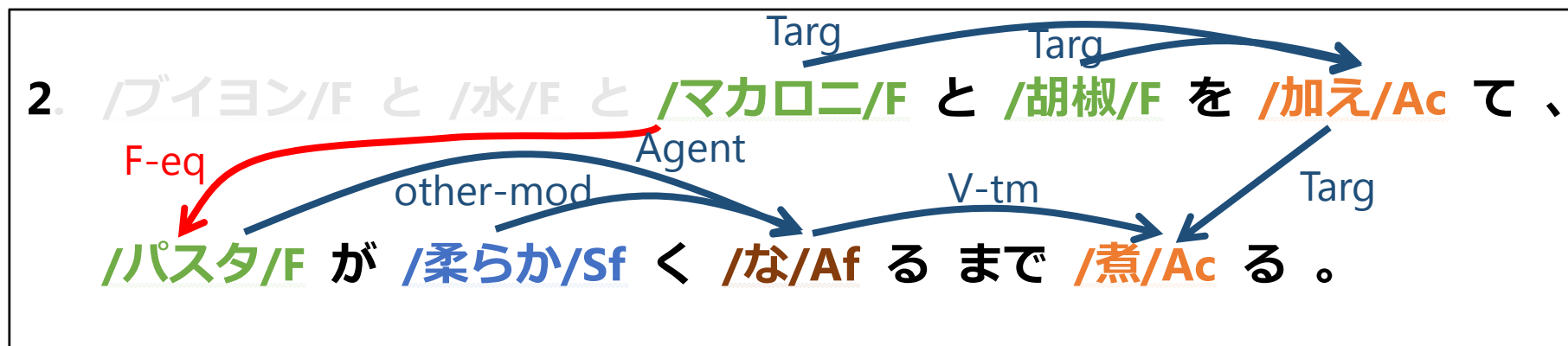
2. /ブイヨン/F と /水/F と /マカロニ/F と /胡椒/F を /加え/Ac て、
/パスタ/F が /柔らか/sf く /な/Af る まで /煮/Ac る。

Step 3-2. Arc Addition

Arc addition under some constraints



↓ **LR score** > $\xi / \lambda e^{-\lambda n}$ Penalty for n -th arc addition



Features of a Labeled Arc

... で /油/F を **u** /熱/Ac する。
(... heat oil.)

... と /ニンニク/F を **v** /加え/Ac 、 /1分 ほど/D /炒め/Ac る。
(Add ... and garlic. Cook for about 1 minute.)

Combination of the followings:

- 3 surrounding words of **u** or **v**
- function words between **u** and **v**
- arc label
- word sequence, pronunciation, and concept tag of **u** or **v**
- whether **u** and **v** in the same sentence or not
- whether **u** and **v** in the same step or not
- whether an Ac exists between **u** and **v** or not

Naïve MST Parser Application

- Only concepts are considered.

1. /両手 鍋/T で /油/F を /熱/Ac する。
(Dutch oven) (oil) (heat)

/セロリ/F と /青 ねぎ/F と /ニンニク/F を /加え/Ac 、
(celery) (green onions) (garlic) (add)

/1分 ほど/D /炒め/Ac する。
(about 1 minute) (cook)

2. /ブイヨン/F と /水/F と /マカロニ/F と /胡椒/F を /加え/Ac て、
(broth) (water) (macaroni) (pepper) (add)

/パスタ/F が /柔らか/Sf < /な/Af る まで /煮/Ac する。
(pasta) (tender) (is) (simmer)

3. /刻/Ac んだ /セージ/F を /まぶ/Ac す。
(snipped) (sage) (sprinkle)

Procedural Text Understanding

- Words not covered by concepts are clues.

1. /両手 鍋/T で /油/F を /熱/Ac する。
(Dutch oven) (oil) (heat)

/セロリ/F と /青 ねぎ/F と /ニンニク/F を /加え/Ac 、
(celery) (green onions) (garlic) (add)

/1分 ほど/D /炒め/Ac する。
(about 1 minute) (cook)

2. /ブイヨン/F と /水/F と /マカロニ/F と /胡椒/F を /加え/Ac て、
(broth) (water) (macaroni) (pepper) (add)

/パスタ/F が /柔らか/Sf < /な/Af る まで /煮/Ac する。
(pasta) (tender) (is) (simmer)

3. /刻/Ac んだ /セージ/F を /まぶ/Ac す。
(snipped) (sage) (sprinkle)

EVALUATION

Experimental Settings

- Criterion: F-measure
 - Harmonic mean of precision and recall
 - Unit: labeled arcs
- Recipe flow graph corpus (r-FG corpus)
 - Randomly downloaded from CookPad
 - Manually annotated with flow graphs
 - Available at <http://plata.ar.media.kyoto-u.ac.jp/data/recipe/>

#Recipes	#Sent.	#Vertices	#Words
200	1,303	7,268	25,446

→ 10-fold cross validation

Flow Graph Estimation

- Baseline: naïve application of an MST parser
 - Input: a concept sequence (cf. word sequences)
 - MST + arc addition
 - Ignore words not covered by the concepts

Method	Precision	Recall	F-measure
Baseline	65.1	61.5	63.2
Proposed	73.5	69.1	71.2

- Words other than concepts are important for flow graph estimation.

Text Parsing on a Raw Text

Task	Input	F
Word segmentation (WS)	Raw Recipe Texts	98.6
Concept Identification (CI)	Gold WS results	90.7
Flow Graph Estimation (FGE)	Gold WS/CI results	72.1
WS + CI + FGE	Raw Recipe Texts	51.6

- WS
 - Enough high and well solved
- CI
 - Less accurate than WS
 - Needs improvement or more resource
- FGE
 - Least accurate, most difficult
 - Needs more resource or more sophisticated techniques

CONCLUDING REMARKS

Conclusion

Procedural text

1. 両手鍋で油を熱する。
セロリと青ねぎとニンニクを加え、
1分ほど炒める。
2. ブイヨンと水とマカロニと胡椒を加えて、
パスタが柔らかくなるまで煮る。
3. 刻んだセージをまぶす。

Today's
Talk

Flow graph [Mori+, LREC14]



• Framework for

procedural text understanding

1. Word segmentation (for Japanese or Chinese)
2. Concept identification
3. Flow graph construction (most difficult)
 - Referring to words other than concepts as clues

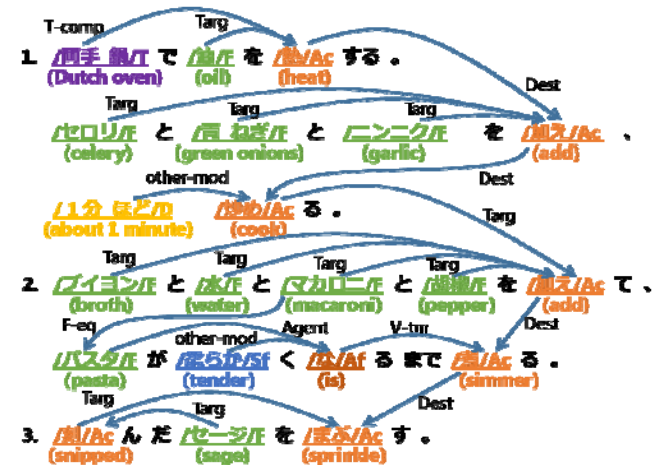
Applications

Procedural text

1. 両手鍋で油を熱する。
セロリと青ねぎとニンニクを加え、
1分ほど炒める。
2. ブイヨンと水とマカロニと胡椒を加えて、
パスタが柔らかくなるまで煮る。
3. 刻んだセージをまぶす。

Today's
Talk

Flow graph [Mori+, LREC14]



Cooking robot

[Bollini+, ISER13]



www.denso.co.jp

Smart kitchen

[Hashimoto+, IPMU08]



Cut garlic
Then add it

Intelligent search

[Yamakata+, SocInfo13]

