< INLG 2006, Jul. 15th, 2006 >

Adj-to-V Paraphrasing in Japanese
Based on Lexical Constraints of Verbs

Atsushi Fujita   @ Nagoya Univ.
Naruaki Masuno   @ IBM Eng. & Tech.
Satoshi Sato   @ Nagoya Univ.
Takehito Utsuro   @ Univ. of Tsukuba

Paraphrases

- ■ "Different expressions that convey the same meaning"

  `The software is available on the Web.`
  `You can get the software from the Web.` …
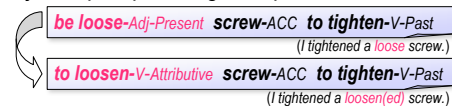 `The tool can be downloaded from a Website.`

- ■ Application
  - Generation
    - Text simplification [Carroll et al., 1999] [Inui et al., 2003]
    - Pre- and post-editing for MT [Shirai et al., 1995]
  - Recognition
    - QA [Hermjakob et al., 2002] [Takahashi, 2005]
    - Multi-document summarization [Barzilay et al., 2003]

2

Approaches to handle paraphrases

- ■ Collecting synonymous expressions
  - from thesauri, parallel/non-parallel corpus, Web, etc.
  - Lexical paraphrases ("burst into tears" ⇔ "cry")
- ■ Describing transfer rules
  - transformational grammar, synchronous grammar, etc.
  - Syntactic paraphrases (passivization, topicalization)
- ■ Combining transfer rules and lexical resources
  - Lexical relation (LFs in MTT: e.g. v(attractive) = attract)
  - Lexical constrains (aspectual property, agentivity, etc.)
  - Compositional paraphrases (alternation, category-shifting)

3

Phenomena we focus on

- ■ Adj-to-V paraphrasing in Japanese

  *be loose-*Adj-Present ***screw-***ACC ***to tighten-***V-Past
  (*I tightened a loose screw.*)

  *to loosen-*V-Attributive ***screw-***ACC ***to tighten-***V-Past
  (*I tightened a loosen(ed) screw.*)

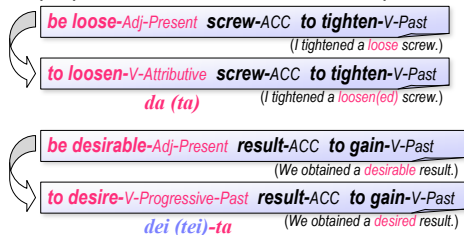 - A challenge to inter-categorial paraphrasing
    - Bridging gaps between different syntactic categories
  - Reasonable material for extending lexical semantics
    - Utilizing frameworks for verbs, at first

4

Issue

- ■ Add proper verbal suffixes to ensure equivalency

  *be loose-*Adj-Present ***screw-***ACC ***to tighten-***V-Past
  (*I tightened a loose screw.*)

  *to loosen-*V-Attributive ***screw-***ACC ***to tighten-***V-Past
  da (ta)   (*I tightened a loosen(ed) screw.*)

  *be desirable-*Adj-Present ***result-***ACC ***to gain-***V-Past
  (*We obtained a desirable result.*)

  *to desire-*V-Progressive-Past ***result-***ACC ***to gain-***V-Past
  dei (tei)-ta   (*We obtained a desired result.*)

  - What sorts of suffixes are appended?
  - What determines them?

5

Distribution of verbal suffixes

- ■ Closed example collection C_{ad-c}
  - 91 Adj tokens
  - 148 candidate Vs

  - Choice points
    - Voice: re/φ
    - Aspect: *tei*/φ
    - Tense: *ru/ta*
  - Majority
    - *ta*
    - *re- ru*

Verbal suffixes	# of Adj-V pairs		
	Total	Vi	Vt
ru	9	7	2
tei-ru	5	3	2
re-ru	14	0	14
re-tei-ru	2	0	2
ta	57	44	13
tei-ta	2	0	2
re-ta	6	0	6
re-tei-ta	0	0	0
ta and *tei-ru*	4	4	0
ta and *ru*	1	1	0
non-paraphrasable	48	13	35

6

Task

- Determining verbal suffixes
  - Input
    - Pair of Adj (source) and V (target)
    - Head of noun phrase modified by the Adj/V

      > *loose* ⇔ *loosen*

      > *screw*
  - Output
    - Verbal suffixes for the V
      - candidates are those in **Table 1**

        > *da (ta)* (Attributive)

  > **be loose**-*Adj-Present* **screw-**ACC **to tighten-***V-Past*
  > (*I tightened a loose screw.*)
  > **to loosen-***V-Attributive* **screw-**ACC **to tighten-***V-Past*
  > *da (ta)*   (*I tightened a loosen(ed) screw.*)

Method: verbal features

- Verb generation ⇒ Utilizing lexical constraints
  - 7 linguistic tests derived from LCS (**Table 2, p.43**)
    [Kageyama, 1996] [Kato et al., 2005] [Takeuchi et al., 2006]
    - Transitive or intransitive (V_a)
    - Aspectual property (5 tests: V_b-V_f)
    - Agentivity (1 test: V_g)
  - Examples of manual examination result
    - *nozomu* (to desire): V_a: "yes (transitive)", V_b: "yes", V_c: "yes", V_d: "no", V_e: "progressive", V_f: "no", V_g: "yes"
    - *wasureru* (to forget): V_a: "yes (transitive)", V_b: "no", V_c: "yes", V_d: "no", V_e: "perfective", V_f: "yes", V_g: "no"

Method: handmade rule-set

- The rule-set consists of 8 rules (**Table 3, p.43**)
  - Each labels either of verbal suffixes (**Table 1**)
  - Conditioned with conjunction of feature-value
    - 7 linguistic test results
    - Some clues are employed in addition
      - *N*: Semantic class of the noun [NIJL, 2004]
        - ✓ e.g. *neji* (screw): *material 1 (1.4151)*
      - *D*: Affix pattern
        - ✓ e.g. *yurui* (be loose) ⇔ *yurumu* (to loosen): A_i-V_mu
      - *C*: Clause or not
  - Otherwise non-paraphrasable

Experimental results (Table 4)

Verbal suffixes		C_{ad-c}		C_{ad-o}	
		Recall	Precision	Recall	Precision
ta	Vi	42/44	42/63	18/18	18/29
	Vt	3/13	3/3	1/6	1/1
re-ru	Vt	12/14	12/19	7/13	7/11
ru	Vi	3/7	3/6	0/1	0/5
	Vt	0/2	0/0	0/1	0/0
tei-ru	Vi	0/3	0/0	0/6	0/0
	Vt	1/2	1/7	2/4	2/5
ta and *tei-ru*	Vi	2/4	2/2	0/1	0/1
No rule		for 11 inputs		for 6 inputs	
Total (**RULE**)		63/100 (63%)	63/100 (63%)	29/56 (52%)	29/53 (55%)
BL		57/100 (57%)	57/148 (39%)	24/56 (43%)	24/83 (29%)

Observations

- Necessary conditions for Vi+"*ta*" and "*re-ru*" are covered
- But those rules do not give sufficient conditions
- Others are not well-analyzed yet

Conclusion

- What we do:
  - Adj-to-V paraphrasing
    - As a case study of inter-categorial paraphrasing
  - Determining verbal suffixes in Adj-to-V paraphrasing
    - To ensure equivalence in aspectual meaning
    - Aspectual and agentive properties are useful
      - Majority of the test collection is explained (R, P=.52-.63)
      - (Beats ML and SLM (for our small data-set))
    - (Over-generation & SLM-based filtering work well)

Future work

- Further experiments on Adj-to-V paraphrasing
  - with a larger dataset
  - Explore lexical properties of adjectives and contextual clues
  - Re-design the task: features → aspect → suffixes
- Other types of inter-categorial paraphrasing
  - Predicative and adverbial usages of adjectives to verbs
  - V-to-Adj (reverse)
  - Adj-to-N, N-to-Adj, etc.
- Re-formalization as a problem of lexical choice
  - In what condition, V is preferable to Adj, Vi is to Vt