Comparison of Utterance Generation Methods for Artificial Second Language Tutor

Taku Nakamura†, Rafal Rzepka††, Kenji Araki††, Kentaro Inui†

†Tohoku University, ††Hokkaido University

ESL learning

- Japanese ESL (English as a Second Language) learners have **limited opportunities to use English** [1]

- A possible solution: **artificial language tutor**

  Chat system as a

  - language acquisition supporter
  - conversational partner

- [1] Paul Doyon. Shyness in the Japanese EFL class: Why it is a problem, what it is, what causes it, and what to do about it. The Language Teacher, 24(1):11–16, 2000.

Artificial language tutor

- Final goal
  - Build an attractive artificial tutor for language learning
    - An ideal example


User: *"I'm college student".*


System: *"I'm* **a** *college student, I see!*
I'm interested in learning. What is your major?"

- Conversational contents need to be **controlled** and **related** to user utterances

Difficulties in artificial language tutors

- Responses must be linguistically correct

- Balance between interesting and harmful responses
  - Knowledge from the Web or users can interests learners, but can contain improper or discriminative expressions

This work

- Proposes prototype methods of generating responses;
  - Affected by leaners' utterances
  - Based on templates and comparatively reliable knowledge resources
- Compares generated utterances with those of traditional chat systems
- Shares findings on user preferences for correction styles

System for language learning

- CSIEC (Computer Simulation in Educational Communication) system [2]
  - Multiple functions for English learning
  - Includes a chatting partner (chatbot)
  - Based on databases of knowledge
  - Shortage of topics in system utterances

- [2] Jiyou Jia. CSIEC: A computer assisted English learning chatbot based on textual knowledge and reason- ing. Knowledge-Based Systems, 22(4):249–255, 2009.

Providing conversational topics in our approach

- Assumption:

  Input-related words /phrases can offer chatting topics

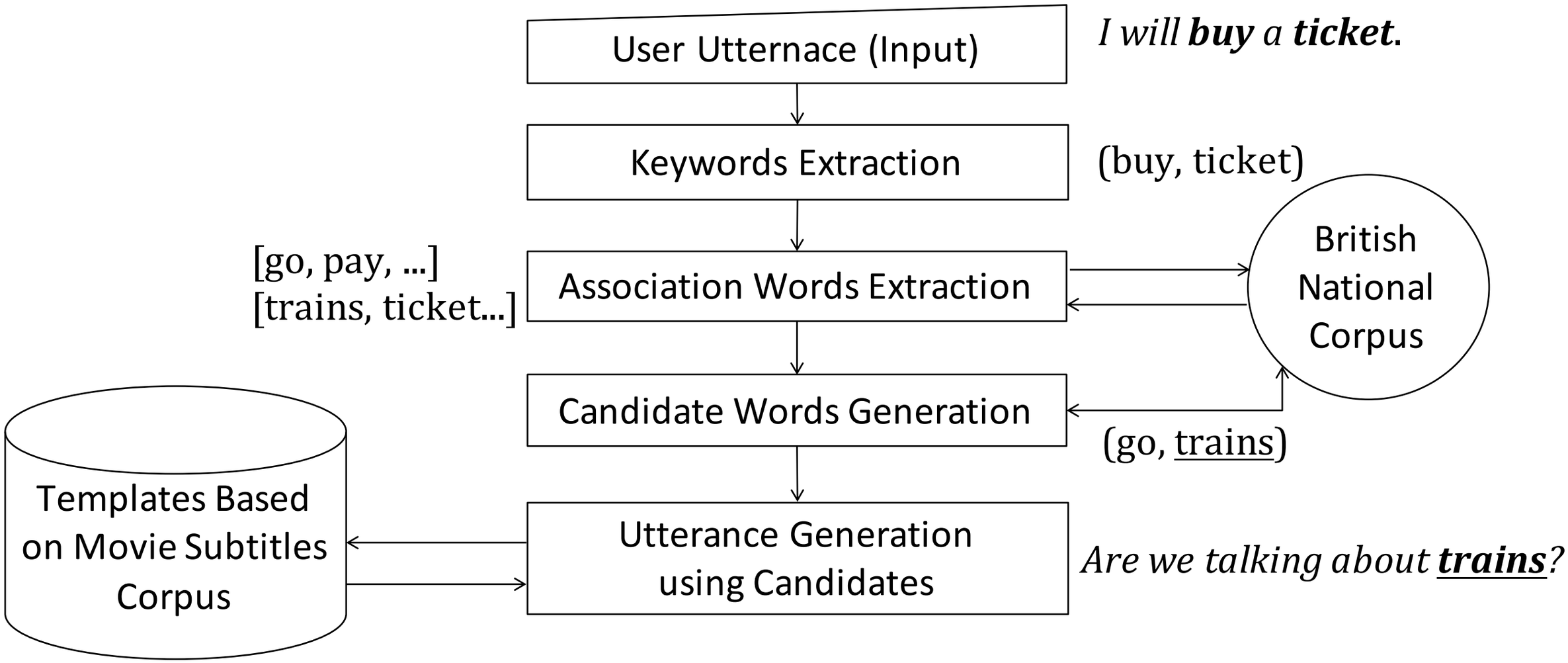
  e.g. ‘movies’ → ‘actors’, ‘watching’, ‘eating popcorn’

- Proposed methods use
  - Word associations by co-occurrence frequency
  - Relations in an ontology (ConceptNet)

CoAPM overview

- Co-occurring Action Phrases-based Method

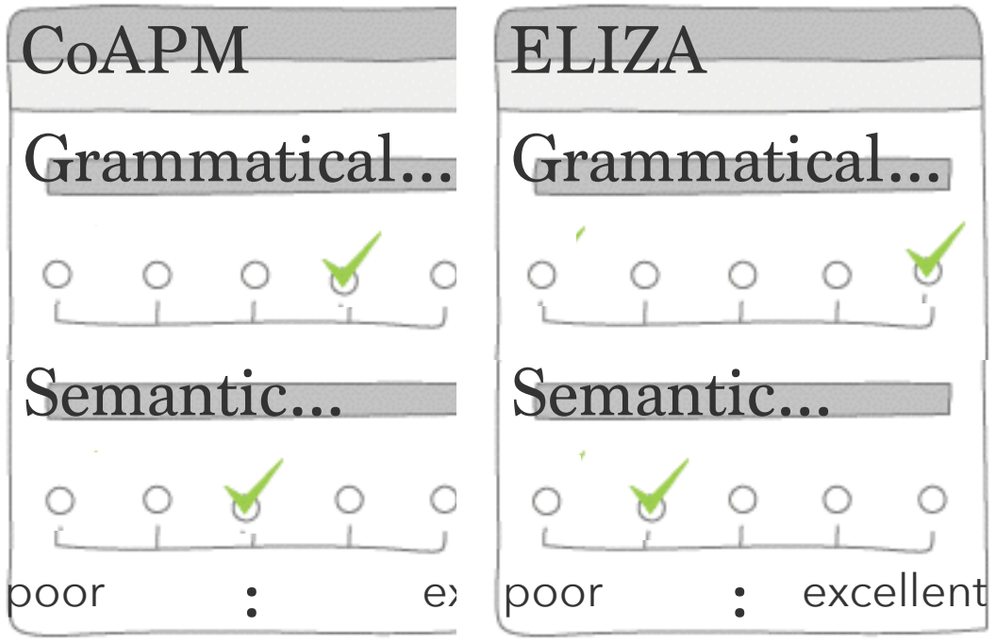
  uses **verbs** and **nouns** frequently co-occurring with those in an input utterance in sentences in the corpus



CoAPM evaluation

- Compared utterances from CoAPM and ELIZA* responded to 20 inputs in a questionnaire
  - To see how learners react to each response

Input 1: *"So, very huge and we can do the all sports."*
CoAPM: *"Are we talking about school ?"*
ELIZA: *"Can you elaborate on that?"*
:



*python implementation of ELIZA by Jez Higgins (http://www.jezuk.co.uk/cgi- bin/view/ software/eliza).

Evaluation settings

- **Input utterances**: selected from test takers' utterance in the English interview test transcription

  The NICT JLE Corpus (https://alaginrc.nict.go.jp/nict_jle/index_E.html)

- **Evaluators**: Five Japanese college students

  (four males in their early 20s, one male in his early 30s)

- **Criteria**
  - Grammatical naturalness
  - Semantic naturalness
  - Motivation to keep studying
- Assessed in 5-point scale (1: poor ⟷ 5: excellent)

CoAPM evaluation results

- Average scores

	CoAPM	ELIZA
Grammatical naturalness	3.50 (1.25)	**3.74** (1.45)
Semantic naturalness	2.20 (1.43)	**2.25** (1.49)
Motivation to keep studying	2.17 (1.37)	**2.39** (1.46)

- ( ) : Standard deviations of scores among evaluators

- Original criteria were given in Japanese language

Example erroneous outputs and analysis

- Input: *"In free time, I like to read books."*
  - CoAPM: *"Does chapter read?"*
  - ELIZA: *"Very interesting."*

Input-related but linguistically incorrect due to insufficient templates

- Input: *"What did you watch?"* [*watch*: verb]
  - CoAPM: *"Are we talking about watch?"* [*watch*: noun]
  - ELIZA: *"Please consider whether you can answer your own question."*

Incorrect POS analysis; should be treat as a verb

- Observations:
  - **Confirming style** of ELIZA is preferred in a short run conversation for language acquisition purpose
  - **Follow-up questions** are important for tutoring task

Preliminary survey of error correction methods

- For future English error suggesting function
- Questionnaire on preferences for correction methods in dialogue
- Presented three types of correction examples to an erroneous input: *“I spend time listening music.”*

- Subjects: The same five Japanese college students

Preliminary survey of error correction methods

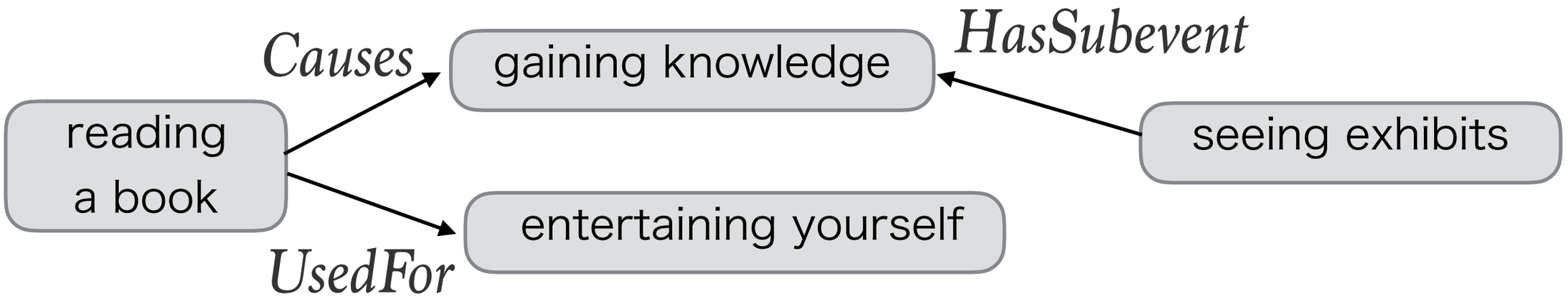
- Corrections in each method to the erroneous input: *"I spend time listening music."* and the questionnaire results

Correction methods	Examples	Respondents
Explicit-correction	*"No, listening to"*	2 / 5 (40%)
Recast	*"listening to"*	2 / 5 (40%)
Prompt	*"listening..."*	1 / 5 (20%)

- **"Explicit-correction"** and **"Recast"** were preferred in this small survey

Utilization of ontology and phrases in a specific form

- ConceptNet: general human knowledge resource
  - Helps to deal with ongoing topics

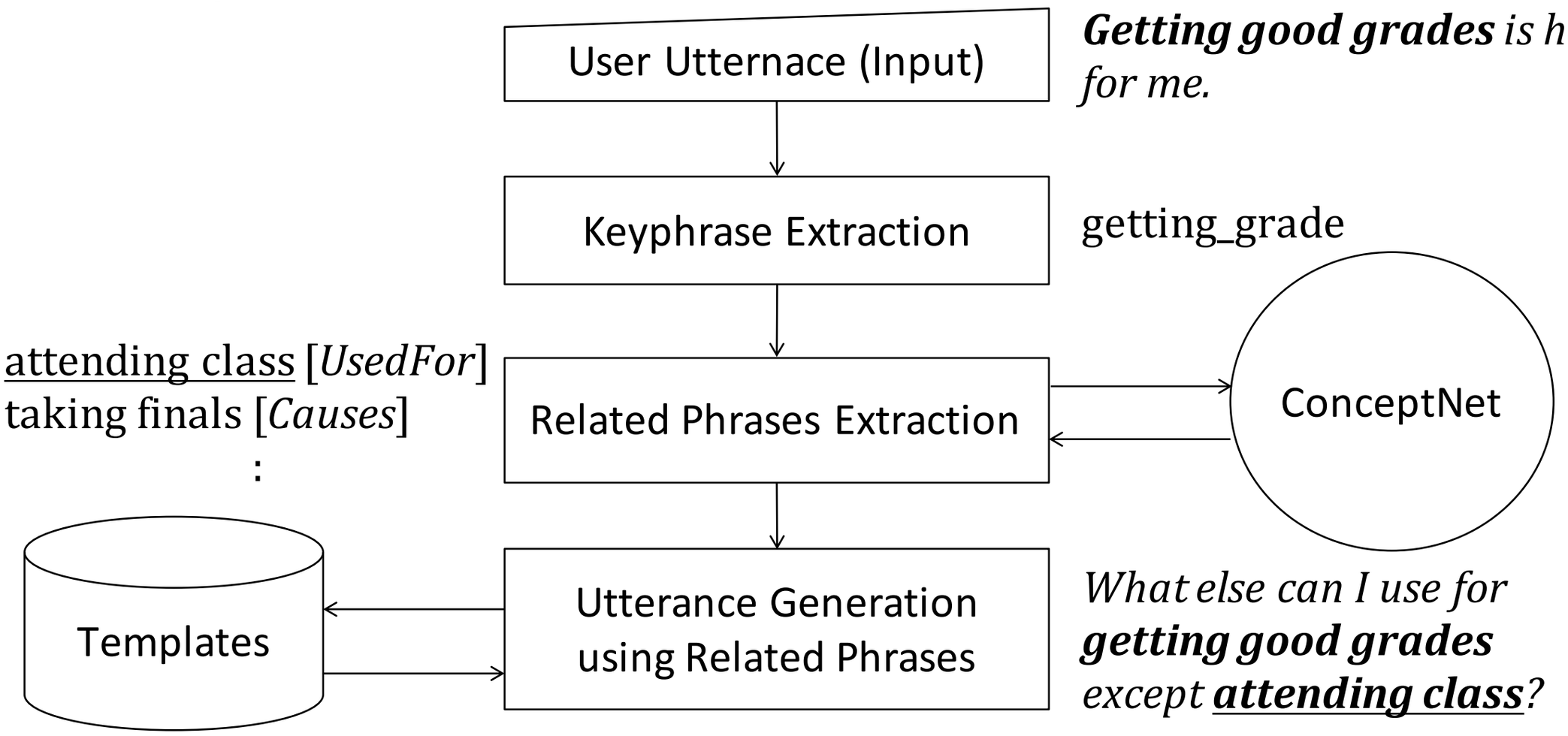


- Utilize ***relations*** and **natural language expressions**
- Handle <u>"- ing (gerund form of verb) + noun"</u> phrases
  - Focus on utterance contents

RAPM overview

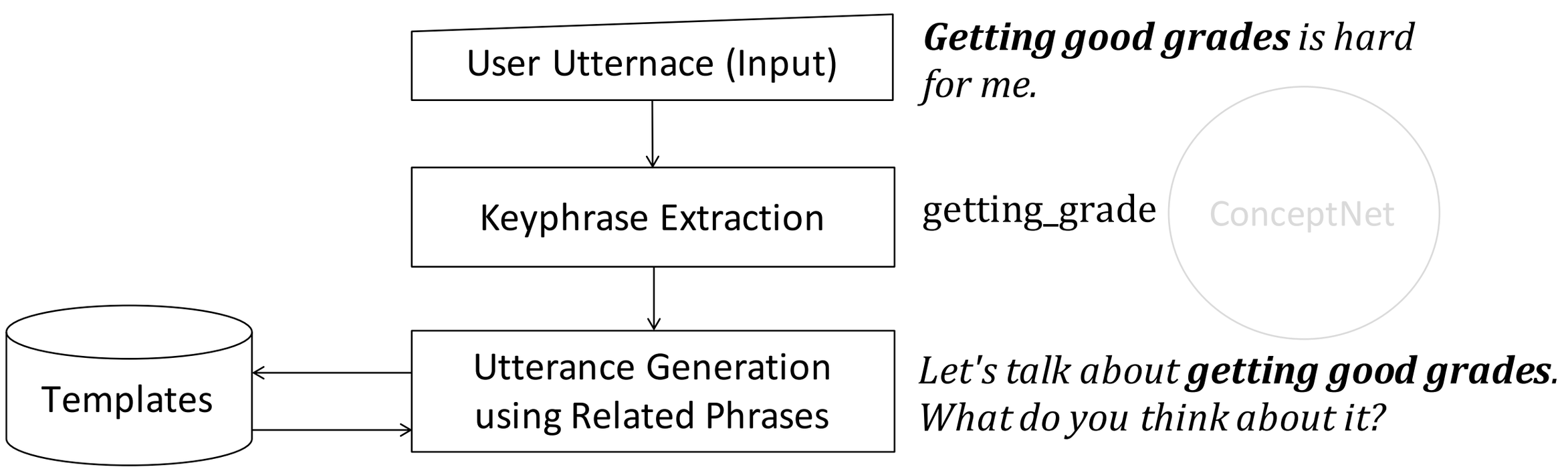
- Related Action-Phrases based Method

  uses [*relations*] between "-ing + noun" form phrases in ConceptNet



CiAPM overview

- Cited Action Phrases-based method
  - To assess effectiveness of repeating approach
  - Uses “-ing + noun” phrases in the input utterance



Templates for RAPM and CiAPM

- **Templates for any relation** - RAPM-NONREL and CiAPM
  - "Talking about [V-ing N (related phrase)]... What is your opinion on that topic?"
  - "Speaking of that, what do you think about [V-ing N (related phrase)]?"
- **Templates for specific relations** - RAPM-REL
  - relation: *RelatedTo*

"Often [V'-ing N' (phrase from input)] and [V-ing N (related phrase)] are a good combination. What do you think?"

  - relation: *HasProperty*

"What about [V-ing N (related phrase)] while [V'-ing N' (phrase from input)]?"

([V-ing N] denotes "-ing (verb in gerund form) + noun" phrase)

RAPM / CiAPM evaluation

- Compared utterances from RAPM (-NONREL / - REL), CiAPM and ELIZA, ALICEBOT responded to 10 inputs in a questionnaire

- **Input utterances**: randomly selected from test takers' utterance (containing at least one "-ing + noun" phrase) in the English interview test transcription (The NICT JLE Corpus)
- **Evaluators**: Six Japanese college students

  (three undergraduates and three graduates in their 20s, science majors)

Evaluation criteria

(A) Will to continue the conversation

(B) Semantical naturalness of dialogue

(C) Appropriateness in English conversation practice

(D) Vocabulary richness

(E) Knowledge richness

(F) Human-likeness of the system

- 5 - point scale (1: poor ⟷ 5: excellent)
- Original criteria were given in Japanese language

RAPM / CiAPM evaluation results

- Average scores (and standard deviations among evaluators)

	(A) Continuity	**(B)** Semantic	**(C)** Practice	**(D)** Vocabulary	**(E)** Knowledge	**(F)** Humanity
CiAPM	**3.10** (1.09)	**3.12** (1.11)	**3.36** (1.09)	2.98* (0.94)	**3.16** (0.94)	**3.10** (0.95)
RAPM-NONREL	3.08 (1.12)	2.82 (1.09)	3.16 (1.14)	**3.20*** (0.91)	3.02 0.95)	2.88 (1.08)
RAPM-REL	3.08 (1.21)	2.70 (1.17)	3.12 (1.20)	3.18* (0.89)	2.98 (1.05)	2.96 (1.23)
ELIZA	2.36 (1.19)	2.92 (1.28)	2.56 (1.21)	2.30 (1.16)	2.54 (1.16)	2.78 (1.21)
ALICEBOT	2.80 (1.21)	2.66 (1.29)	2.94 (1.31)	2.66 (1.14)	2.84 (1.09)	2.92 (1.32)

*Statistically significant difference between the model and ELIZA scores

Example outputs and analysis (1)

- Input: *"And a woman is playing piano."*
  - CiAPM:
    "Let's talk about playing piano. What do you think about it?"
  - RAPM-NONREL:
    "Learning something about music... why some people love it and some don't?"
  - RAPM-REL:
    "I guess you were learning something about music - can you tell me how you did that?"
  - ELIZA:
    "Let's change focus a bit... Tell me about your family."
  - ALICEBOT:
    "Aren't all woman playing piano ?"

Input-affected by ConceptNet relations ⇨ Responses with richer vocabulary

Example outputs and analysis (2)

- Input: *"And they enjoyed eating delicious food and alcohol."*
  - CiAPM:
    "Let's talk about eating delicious food. What do you think about it?"
  - RAPM-NONREL:
    "Talking about diminishing your own hunger... What is your opinion on that topic?"
  - RAPM-REL:
    "I guess you were diminishing your own hunger - can you tell me how you did that?"

Input-affected but topic preference dependents on users

- Conversational topics need to be chosen considering the context and learners' preferences / language levels

Conclusion

- For a future artificial language tutor,
  - Proposed input-affected utterance generation methods
  - Assessed how leaners react with utterances generated by the methods and traditional chatbots

- ELIZA-like confirming style was preferred in the task
- Our approach showed richer vocabulary
- Conversational topics must be carefully selected

Future work

- Combine our methods with vocabulary acquisition systems with a language level estimator [3]
- Incorporate
  - Personality modeling
  - Context processing
  - Functioning spelling and grammar error detection

- Experiment on tutor's autonomy in choosing topics
  - Consider approaches to restrict potentially harmful expressions

• [3] Michal Mazur. A Study on English Language Tutoring System Using Code-Switching Based Second Language Vocabulary Acquisition Method. PhD thesis, Hokkaido University, 2016

References

- Paul Doyon. Shyness in the Japanese EFL class: Why it is a problem, what it is, what causes it, and what to do about it. The Language Teacher, 24(1):11–16, 2000.
- Jiyou Jia. CSIEC: A computer assisted English learning chatbot based on textual knowledge and reason- ing. Knowledge-Based Systems, 22(4):249–255, 2009.
- Shawn Loewen. Error correction in the second language classroom. Clear News, 11(12):1–7, 2007.
- Diane J Tedick. Research on error correction and implications for classroom. ACIE Newsletter, 1986.
- Shinsuke Higuchi, Rafal Rzepka, and Kenji Araki. A casual conversation system using modality and word associations retrieved from the Web. In Proceedings of the Conference on Empirical Methods in Nat- ural Language Processing, EMNLP ‘08, pages 382–390, Stroudsburg, PA, USA, 2008. Association for Computational Linguistics.
- Michal Mazur. A Study on English Language Tutoring System Using Code-Switching Based Second Language Vocabulary Acquisition Method. PhD thesis, Hokkaido University, 2016