

Analyzing Post-dialogue Comments by Speakers – How Do Humans Personalize Their Utterances in Dialogue? –

Toru Hirano* Ryuichiro Higashinaka Yoshihiro Matsuo

NTT Media Intelligence Laboratories, Nippon Telegraph and Telephone Corporation

1-1 Hikarinooka, Yokosuka-Shi, Kanagawa, 239-0847, Japan

{hirano.tohru,higashinaka.ryuichiro,matsuo.yoshihiro}@lab.ntt.co.jp

Abstract

We have been studying methods to personalize system utterances for users in casual conversations. We know that personalization is important, but no well-established way to personalize system utterances for users has been proposed. In this paper, we report the results of our experiment that examined how humans personalize utterances when speaking to each other in casual conversations. In particular, we elicited post-dialogue comments from speakers and analyzed the comments to determine what they thought about the dialogues while they engaged in them. In addition, by analyzing the effectiveness of their thoughts, we found that dialogue strategies for personalization related to “topic elaboration”, “topic changing” and “tempo” significantly increased the satisfaction with regard to the dialogues.

1 Introduction

Recent research on dialogue agents has focused on casual conversations or chats (Bickmore and Picard, 2005; Ritter et al., 2011; Wong et al., 2012; Meguro et al., 2014; Higashinaka et al., 2014) because chat-oriented conversational agents are useful for entertainment or counseling purposes. For chat-oriented conversational agents, it is important to personalize their utterances to increase user satisfaction (Sugo and Hagiwara, 2014). Several methods to personalize system utterances using user information extracted from dialogues have been proposed (Sugo and Hagiwara, 2014; Kim et al., 2014; Kobayashi and Hagiwara, 2016). Although we know that personalization is important,

*Presently, the author is with Nippon Telegraph and Telephone East Corporation.

no well-established way to personalize system utterances for users has been proposed.

In this paper, we report the results of our experiment that examined how humans personalize their utterances when speaking to each other in casual conversations. In particular, to analyze what speakers aimed to convey in dialogues (called **dialogue strategy**), we collected post-dialogue comments by interviewing speakers individually about what they thought about the dialogues after a one-to-one text-based chat. In the interview, we recorded what the speaker said and later made a transcript of the recorded voice for analysis. We manually analyzed the post-dialogue comments to break the dialogue strategies for personalization down into patterns.

In the experiment, we extracted 252 dialogue strategies for personalization from 2,498 utterances. Then, we broke them down into 39 unified dialogue strategies with 10 categories. In addition, by analyzing the effectiveness of the dialogue strategies in relation to the satisfaction of speakers with regard to the dialogues, we found that using the dialogue strategies in the “topic elaboration”, “topic changing”, and “tempo” categories of chat-oriented conversational agents would be expected to increase user satisfaction.

2 Related Work

ELIZA (Weizenbaum, 1966) and ALICE (Wallace, 2004) are chat-oriented conversational agents that have the capability to personalize system utterances for users. For example, these agents can use the user’s name or show that they remember the user’s preferences by filling slots of utterance templates with user information extracted from previous utterances.

There are several studies on personalizing system utterances using user information extracted

from dialogues (Sugo and Hagiwara, 2014; Kim et al., 2014; Kobyashi and Hagiwara, 2016). They used the same approach as that of ELIZA and ALICE to show that the agents remember user information. In addition, they selected system utterances that had the most similar vectors to the user’s interest, which were represented by word vectors of previous utterances. This way is often used in information search (Shen et al., 2005; Qiu and Cho, 2006) and recommendation (Ardissono et al., 2004; Jiang et al., 2011).

Some commercial chat-oriented conversational agents have a function for personalizing system utterances for a user. For instance, an application called “Caraf”¹ operates simultaneously with car navigation systems and preferentially guides the registered user in accordance with his/her favorite brands for banks, gas stations, convenience stores, and so on. A dialogue API called “TrueTALK”² provides information related to the user’s likes and tastes, e.g. it provides concert information for the user’s favorite singers when the user says “I have free time”. A social robot called “Jibo”³ can learn the user’s preferences to personalize system utterances by selecting topics related to the user’s preferences.

From these studies, it can be seen that there have been many attempts to personalize system utterances. However, as far as we know, there is no thorough research about ways to personalize utterances in dialogues.

3 Collecting Post-dialogue Comments

3.1 Procedure

To analyze dialogue strategies of speakers, we collected *post-dialogue comments* by interviewing experimental participants individually about what they thought about the dialogue after a one-to-one *text-based chat*. In the interview, to elicit spontaneous comments from the speakers, what the speaker said was recorded and was later manually transcribed. After the interview, each participant filled out a *questionnaire about satisfaction*.

For experimental participants, we recruited 4 advanced-level speakers of text-based chat, who use text-based chat on business, and 30 normal

¹<http://www.fujitsu-ten.co.jp/eclipse/product/wifi/caraf1/index.html>

²http://www.jetrun.co.jp/curation/truetalk_lp.html

³<https://www.jibo.com/>

	Total	Average
Text-based chat	2,457	27.3
Post-dialogue comments	4,986	55.4

Table 1: Number of utterances for 90 dialogues.

speakers who are good at typing and are open to having a conversation with a stranger. The male-female ratio of the experimental participants is 1:1, and most of the participants were in their 20s or 30s. They were paid for their participation.

Text-based Chat

30 normal speakers took part in 3 dialogue sessions each, talking to one of the 4 advanced-level speakers, who was the same gender as the normal speaker. The normal speaker always talked to the same advanced-level speaker.

Normal and advanced-level speakers performed text-based chat in different rooms. In preparation, to get used to the chat operation, the participants first performed an example dialogue session with the experiment manager.

Each dialogue session lasted for ten minutes. The participants were instructed to enjoy the chat with their partner.

Post-dialogue Comments

Just after text-based chat, we collected the post-dialogue comments by interviewing participants separately about what they thought about each of the utterances in the dialogue. We recorded the interview and later manually transcribed it and aligned it to utterances in the text-based chat.

Each interview session lasted for seven minutes. Normal and advanced-level speakers were interviewed in different rooms. At the beginning of each interview session, the participants were given the instruction by text to comment about each utterance in the dialogue they had just engaged in by considering the following points.

- What did you think when you saw your partner’s utterance/reaction?
- What intention did you have when you replied to your partner’s utterance?

Questionnaire about Satisfaction

After the interview, each participant filled out a questionnaire about satisfaction asking for his/her

No.	Speaker: Utterance
1	A: Do you like driving cars?
2	B: Yes, I do. Do you drive a car?
3	A: I don't have a driving license. My world would probably expand if I could drive a car!
4	B: Taking trains or airplanes expands your world more than driving a car.
5	A: As I recall, my friend from Gunma told me about the number of cars per capita in Gunma.
6	B: Yup, yup! it's an obscure area.
...	...
10	B: In fact, I am living in an inconvenient place now, too.
11	A: Really?
12	B: On the outskirts of Kanagawa.
...	...

Table 2: Example of text-based chat.

subjective evaluation of the dialogue on a five-point Likert scale, where 1 is “very dissatisfied”, 2 is “somewhat dissatisfied”, 3 is “neither satisfied nor dissatisfied”, 4 is “somewhat satisfied”, and 5 is “very satisfied”.

3.2 Collected Data

In total, we collected 2,457 utterances (27.3 utterances per dialogue) in text-based chat and 4,986 utterances (55.4 utterances per dialogue) in post-dialogue comments for 90 dialogues as shown in Table 1.

Table 2 and Table 3 show examples of collected text-based chat and post-dialogue comments. “Target” in Table 3 means the corresponding ID (we call target) of the utterance in the text-based chat. For example, from the post-dialogue comment “She remembered that I said I am from Gunma and she said the number of cars per capita in Gunma...” whose target is 5, it can be seen that the partner selected a topic related to both current topics: “car” and the speaker’s hometown. Also, from the post-dialogue comment whose target is 11 and 12, it can be seen that the speakers decided to talk about specific things, which is easy for the partners to understand.

Figure 1 shows 180 (90 dialogues \times 2 participants) answers to a questionnaire about satisfaction, and the average score was 3.87 points.

Target	Post-dialogue comments
1	In line 1, a question related to the topic of the previous dialogue session has been asked!
2, 3	It is my favorite topic. But, just in case, I asked if she likes driving cars in line 2. In line 3, she replied that she does not drive a car, and I was disappointed.
5	She remembered that I said I am from Gunma, and she said that the number of cars per capita in Gunma... I became excited!
...	...
11, 12	In line 11, it is thoughtful of her to be surprised, and in line 12, to be more specific, I said “On the outskirts of Kanagawa”.
12	Therefore, I think it was easy to understand , and it became easy to imagine.
...	...

Table 3: Example of post-dialogue comments by speaker B.

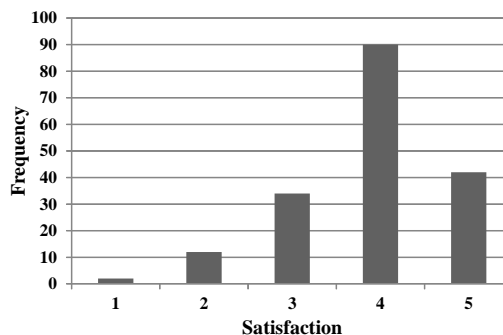


Figure 1: Satisfaction of participants.

4 Analyzing Post-dialogue Comments

4.1 Analysis Procedure

We analyzed the post-dialogue comments for what speakers thought about the dialogues while they engaged in them. The analysis was done as follows: Step 1) we read the post-dialogue comments and manually extracted the dialogue strategies for personalizing the utterances, Step 2) we annotated the extracted dialogue strategies with categories, and Step 3) we unified similar dialogue strategies within each category. In the analysis, we focused

Category	Description
Topic Changing	Strategies about when or how to change topics.
Topic Selection	Strategies about selecting next topic when changing topics.
Topic Elaboration	Strategies about elaborating on current topic.
Topic General	Strategies related to overall topics in dialogues.
Attitude	Strategies about stating one’s opinions and interests.
Expression	Strategies about expressions in utterances.
Tempo	Strategies about tempo of dialogues.
Role	Strategies about roles, speakers or listeners, in dialogues.
Discourse	Strategies about flows in discourses.
Others	Other strategies.

Table 4: Categories of dialogue strategies and their descriptions.

on the comments; the content of the text-based chat was not used.

In this paper, we used 2,498 utterances of post-dialogue comments for 45 dialogues. To analyze inter-annotator agreements, two annotators individually performed the following three steps.

Step 1: Extracting Dialogue Strategies from Post-dialogue Comments

The annotators were instructed to read utterances in post-dialogue comments and find what speakers thought about personalization. When the annotators found such a thought, they annotated the utterances with a summarized text (i.e., dialogue strategy) of the thinking behind the utterances, such as “using the partner’s name” or “talking about topics related to the partner’s hobby”. Otherwise, they annotated the utterances with “no”.

For instance, from the example of post-dialogue comments shown in Table 3, the dialogue strategies “selecting topics related to both the current and previous hometown of the partner” and “bringing up a specific topic” would be extracted. The former strategy would let the partner talk about a familiar topic, and the latter would let the partner easily imagine the topic.

Step 2: Annotating Dialogue Strategies with Categories

To annotate dialogue strategies with categories, we manually defined the 10 categories shown in Table 4 by summarizing the dialogue strategies extracted at Step 1.

There are 4 categories related to topics, such as “topic changing”, which consists of strategies about when or how to change topics, and “topic selection”, which consists of strategies about se-

lecting the next topic when changing topics. Apart from the categories related to “topics”, there are 6 categories, such as “attitude”, which consists of strategies about stating one’s opinions and interests, and “role”, which consists of strategies about speakers or listeners in dialogues.

The annotators were instructed to annotate dialogue strategies extracted at Step 1 with one category from the ten categories shown in Table 4. For instance, the dialogue strategies “selecting topics related to both the current and previous hometown of the partner” and “bringing up a specific topic” would be annotated with the “topic elaboration” and “topic general” categories, respectively.

Step 3: Unifying Similar Dialogue Strategies within Each Category

In dialogue strategies annotated with the same category at Step 2, there may be some strategies that are similar to each other. Therefore, we combine similar dialogue strategies.

The annotators were instructed to unify similar dialogue strategies by generalizing them even though they have different details. For example, the dialogue strategies “talking about topics related to partner’s hobby” and “talking about topics related to partner’s hometown” would be unified to “talking about topics related to partner’s information”.

The unified dialogue strategies induced individually by the two annotators were later compared by the two annotators to see if they correspond to each other. If similar unified dialogue strategies were found, they were given the same identifiers for matching.

4.2 Results

Inter-annotator Agreement

From 2,498 utterances, annotator A extracted 252 dialogue strategies for personalization. The dialogue strategies were unified into 39 kinds of dialogue strategies. Annotator B extracted 303 dialogue strategies and the dialogue strategies were unified into 41 kinds of dialogue strategies. Both annotators annotated 211 utterances with dialogue strategies and 2,154 utterances with no specific strategy at Step 1. At Step 2, both annotators annotated 187 dialogue strategies with the same categories. At Step 3, we found that 156 dialogue strategies out of the 187 dialogue strategies were under the same unified dialogue strategies.

As for the agreement of the extracted dialogue strategies, *precision* is 51.5% (156/303), *recall* is 61.9% (156/252), and *F-measure* is 0.56. These values indicate how annotator B extracts the same unified dialogue strategies as annotator A and are calculated by the following formulae:

$$\begin{aligned} Precision &= \frac{C}{B}, \\ Recall &= \frac{C}{A}, \\ F\text{-measure} &= \frac{2 \cdot precision \cdot recall}{precision + recall}, \end{aligned}$$

where C represents the number of dialogue strategies annotated with the same unified dialogue strategy by both annotators, A represents the total number of extracted dialogue strategies by annotator A, and B represents the total number of extracted dialogue strategies by annotator B.

The *accuracy* of the inter-annotator agreement of annotating 2,498 utterances in post-dialogue comments with unified dialogue strategies, that is the results of Step 1 + 2 + 3, is 92.4% (Cohen’s $\kappa = 0.64$) (Cohen, 1960). Here, the accuracy is calculated by the following formula:

$$Accuracy = \frac{M}{T}$$

where M represents the number of utterances that are annotated with the same unified dialogue strategies or “no” by both annotators, and T represents the total number of utterances used for the analysis. Because κ is more than 0.6, we can say the agreement is substantial. Table 5 shows the inter-annotator agreement for each step in the annotation.

	<i>Accuracy</i>	κ
Step 1	94.7 (2,365/2,498)	0.73
Step 1 + 2	93.7 (2,341/2,498)	0.69
Step 1 + 2 + 3	92.4 (2,310/2,498)	0.64

Table 5: Inter-annotator agreement of 2,498 utterances in post-dialogue comments.

Dialogue Strategies for Personalization

Table 6 shows the results of annotator A; there are 39 kinds of dialogue strategies with annotated categories. It also shows the frequency of each unified dialogue strategy. Note that almost all the dialogue strategies for personalization presented here have not been used in any previous studies. Here, we explain some of the dialogue strategies for personalization in detail.

From this table, we can see that the most frequent dialogue strategies were “telling partner that I am interested in the current topic, too” and “showing empathy for the opinion of the partner” in the “attitude” category, which consists of dialogue strategies for letting the partner talk comfortably in a dialogue. Dialogue strategies in the “attitude” category are mainly used by the conversational participants when they were listening, and there are strategies, such as giving back-channel feedback and showing that I am impressed with the story of the partner, that can be performed by giving praise to the partner.

One of the second most frequent dialogue strategies was “bringing up a specific topic” in the “topic general” category, which is a dialogue strategy for letting the partner speak easily by providing topics that are easy to imagine. For instance, providing a specific topic, “Tigers”, would let the partner speak more easily than an unspecific topic such as “baseball”. In this “topic general” category, there is also a strategy “bringing up several specific topics”, which is similar to the previous strategy “bringing up a specific topic”. This strategy has another purpose, which is to increase the probability that the partner would be interested in one of the topics by providing several specific topics.

With a frequency equal to the dialogue strategy “bringing up a specific topic”, we can see the dialogue strategy “selecting topics related to partner information” in the “topic selection” category, which is a dialogue strategy for letting the partners

Category	Dialogue Strategy	Frequency
Topic Changing	Changing topics when partner does not know about current topic.	■ 5
	Changing topics when only I talked a lot.	■ 3
	Changing topics when my replies seemed to be unexpected.	1
	Changing topics when partner paused for long time in dialogue.	1
	Changing topics by talking about next topic in current conversation.	1
Topic Selection	Selecting topics related to partner's information.	■■■■■ 22
	Selecting topics related to inferred partner information.	■■■■■ 11
	Selecting topics related to common experiences with partner.	■ 4
	Asking question that partner asked me before.	■ 2
	Selecting topics of similar experiences to one partner talked about.	■ 2
Topic Elaboration	Selecting topics related to both current topic and partner information.	■ 5
	Selecting topics related to both current topic and inferred partner info.	■ 2
	Asking about past experiences of topic after talking about present one.	1
Topic General	Bringing up specific topic.	■■■■■ 22
	Bringing up several specific topics.	■■■■■ 13
	Not talking about too local topics.	■■■ 8
	Bringing up topic that seems to be common topic.	■■■ 6
	Bringing up topic in way that makes partner ask questions.	■ 3
	Answering only questions that partner would ask again.	■ 2
	Answering question and bringing up conversable topic.	■ 2
	Asking questions that seem to be easy for partner to answer.	1
Attitude	Telling partner that I am interested in current topic, too.	■■■■■ 24
	Showing empathy for partner.	■■■■■ 24
	Showing that I am impressed with story of partner.	■■■ 6
	Giving back-channel feedback.	■ 2
	Not saying anything negative.	1
Expression	Using emotional terms.	■■■■■ 16
	Using friendly and frank expressions.	■■■■■ 10
	Using expressions that partner used.	■ 2
	Using partner's name.	■ 2
	Using terms for sharing feelings.	1
	Exaggerating story.	1
Tempo	Keeping dialogue fast-paced.	■■■■■ 14
	Keeping pace with tempo of partner.	■■■ 7
Role	Both participants in the conversation speaking one after another.	■■■ 7
	Changing roles, speaker or listener, depending on partner.	■ 5
Discourse	Talking about partner after talking about myself.	■■■■■ 10
Others	Asking open questions because partner likes talking a lot.	■ 2
	Asking "why" questions.	1

Table 6: Unified dialogue strategies to personalize utterances in dialogue extracted by annotator A.

speak easily by providing topics related to the partner. Also, we can see the strategy of selecting the topic by using information of the partner inferred from the dialogues and not selecting a totally new topic when changing topics in the dialogue. These strategies are the ones used in the related work. In this category, there is the other dialogue strategy of selecting topics related to common experiences with the partner.

There are dialogue strategies about elaborating on the current topic in the “topic elaboration” category. In this category, the most frequent strategy was “selecting topics related to both the current topic and partner information”. For example, as a simple way to elaborate on the topic “car”, we can select topics about “car parts”, such as tire or handle, or “automakers”, such as Toyota or Honda, as elaboration topics. However, this strategy selects “car life in the countryside” by considering where the partners are from and which topics are familiar to the partner.

As moderately high frequency dialogue strategies, there were strategies using “emotional terms” and “friendly and frank expressions” in the “expression” category. These dialogue strategies are to let the partner feel comfortable by using expressions for talking with one’s friends or families. In this category, there are other strategies such as not only “using the partner’s name”, which is used in related work, but also “using the expressions that the partner used” to take advantage of being close to the partners.

Effectiveness of Category of Dialogue Strategies for Satisfaction of Participants

We analyzed the effectiveness of the category of dialogue strategies in relation to the satisfaction of participants with regard to the dialogues. For each category of dialogue strategy, we split the dialogues into two classes. One is the dialogues whose utterances in post-dialogue comments are annotated with a category, and the other is those whose utterances in post-dialogue comments are not annotated with that category. Then, we calculated the average satisfaction score of the dialogues in the two classes. For the statistical significance test, we used two-tailed tests with Welch’s t-test (Welch, 1947).

Table 7 shows the results. The satisfaction of dialogues annotated with the category “topic elaboration”, “topic changing”, and “tempo” are significantly higher than that of other categories.

Category	Annotated	Not annotated
Topic Changing	4.20*	3.79
Topic Selection	3.90	3.85
Topic Elaboration	4.29**	3.80
Topic General	3.92	3.80
Attitude	3.90	3.84
Expression	3.89	3.87
Tempo	4.19*	3.71
Role	3.90	3.84
Discourse	3.75	3.90
Others	3.25	3.91

Table 7: Average satisfaction scores of dialogues whose utterances are annotated or not annotated with category. Superscript * next to annotated scores indicates that score is statistically better than not annotated score. ** means $p < 0.01$; * means $p < 0.05$. For statistical test, we used two-tailed Welch’s t-test.

The “topic elaboration” and “tempo” categories increased the satisfaction score by 0.48 points and the “topic changing” category by 0.41 points. This means that the personalization using the dialogue strategies in these categories would be expected to increase the user satisfaction.

4.3 Discussion

By analyzing the post-dialogue comments, extracting dialogue strategies for personalization and breaking them down into patterns worked to some extent. In particular, the extracted dialogue strategies were not only the ones in the “topic selection” category, which have been used in related work, but also the ones in the other categories. In addition, by analyzing the effectiveness of the dialogue strategies in relation to the satisfaction of speakers with regard to dialogues, we found that using the dialogue strategies in the “topic elaboration”, “topic changing”, and “tempo” categories with conversational agents would be expected to increase the user satisfaction.

However, some issues remain about the coverage of dialogue strategies for personalization because the dialogue strategy “showing that the agent remembers user information directly”, which is used in related work (e.g. saying “As I recall, you like driving a car, don’t you?”), was not extracted in our analysis. In this paper, we

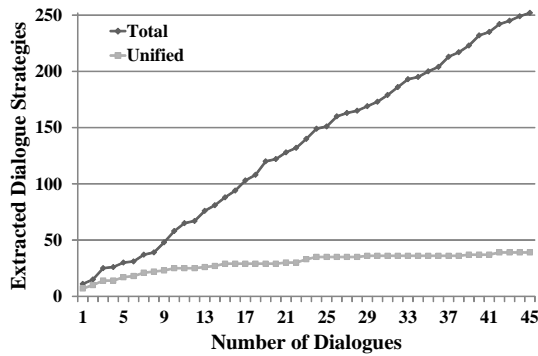


Figure 2: Number of extracted dialogue strategies.

collected all the post-dialogue comments within a day, so dialogue strategies that appear in the long term were not extracted.

It is difficult to collect new dialogue strategies for personalization efficiently by increasing the number of the post-dialogue comments because the increasing rate of unified dialogue strategies are rather low as shown in Figure 2, which shows the number of extracted total and unified dialogue strategies extracted from the post-dialogue comments.

From these points, to collect the post-dialogue comments, the periods of collecting data, such as within a few days, weeks or months, and devising a new means for collecting dialogue strategies should be considered.

5 Summary and Future Work

In this paper, we reported the results of our experiment that examined how humans personalize utterances when speaking to each other in casual conversations. In particular, we solicited post-dialogue comments from speakers and analyzed the comments to find out what they thought about the dialogues while they engaged in them.

In the experiment, we extracted 252 dialogue strategies for personalization from 2,498 utterances. Then, we broke them down into 39 unified dialogue strategies with 10 categories. In addition, we found that using the dialogue strategies in the “topic elaboration”, “topic changing”, and “tempo” categories of chat-oriented conversational agents would be expected to increase user satisfaction.

As future work, we would like to implement the dialogue strategies extracted in the analysis, especially the dialogue strategies in the above three

categories, on chat-oriented dialogue systems to check if they actually increase user satisfaction.

References

- Liliana Ardissono, Cristina Gena, Pietro Torasso, Fabio Bellifemine, Angelo Difino, and Barbara Negro. 2004. User modeling and recommendation techniques for personalized electronic program guides. In *Personalized Digital Television - Targeting Programs to Individual Viewers*, volume 6 of *Human - Computer Interaction Series*, pages 3–26.
- Timothy W. Bickmore and Rosalind W. Picard. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2):293–327.
- J. Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.
- Ryuichiro Higashinaka, Kenji Imamura, Toyomi Meguro, Chiaki Miyazaki, Nozomi Kobayashi, Hiroaki Sugiyama, Toru Hirano, Toshiro Makino, and Yoshihiro Matsuo. 2014. Towards an open domain conversational system fully based on natural language processing. In *Proceedings of the 25th International Conference on Computational Linguistics*, pages 928–939.
- Yechun Jiang, Jianxun Liu, Mingdong Tang, and Xiaoqing Liu. 2011. An effective web service recommendation method based on personalized collaborative filtering. In *Proceedings of the 2011 IEEE International Conference on Web Services*, pages 211–218.
- Yonghee Kim, Jeesoo Bang, Junhwi Choi, Seonghan Ryu, Sangjun Koo, and Gary Geunbae Lee. 2014. Acquisition and use of long-term memory for personalized dialog systems. In *Proceedings of the 2014 Workshop on Multimodal Analyses enabling Artificial Agents in Human-Machine Interaction*.
- Shunya Kobayashi and Masafumi Hagiwara. 2016. Non-task-oriented dialogue system considering user’s preference and human relations (in Japanese). *Transactions of the Japanese Society for Artificial Intelligence*, 31(1):DSF-A_1–10.
- Toyomi Meguro, Yasuhiro Minami, Ryuichiro Higashinaka, and Kohji Dohsaka. 2014. Learning to control listening-oriented dialogue using partially observable markov decision processes. *ACM Transactions on Speech and Language Processing*, 10(4):15:1–15:20.
- Feng Qiu and Junghoo Cho. 2006. Automatic identification of user interest for personalized search. In *Proceedings of the 15th International Conference on World Wide Web*, pages 727–736.

- Alan Ritter, Colin Cherry, and William B. Dolan. 2011. Data-driven response generation in social media. In *Proceedings of the Conference on Empirical Methods in Natural Language*, pages 583–593.
- Xuehua Shen, Bin Tan, and ChengXiang Zhai. 2005. Implicit user modeling for personalized search. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, pages 824–831.
- Kensuke Sugo and Masafumi Hagiwara. 2014. A dialogue system with knowledge acquisition ability from user’s utterance (in Japanese). *Transactions of Japan Society of Kansei Engineering*, 13(4):519–526.
- Richard S. Wallace. 2004. *The Anatomy of A.L.I.C.E.* ALICE Artificial Intelligence Foundation, Inc.
- Joseph Weizenbaum. 1966. ELIZA — a computer program for the study of natural language communication between man and machine. *Communications of the Association for Computing Machinery*, 9:36–45.
- B. L. Welch. 1947. The generalization of ‘student’s’ problem when several different population variances are involved. *Biometrika*, 34(1/2):28–35.
- Wilson Wong, Lawrence Cavedon, John Thangarajah, and Lin Padgham. 2012. Strategies for mixed-initiative conversation management using question-answer pairs. In *Proceedings of the 24th International Conference on Computational Linguistics*, pages 2821–2834.